166h

THE PERFORMANCE OF SOCIAL RESPONSIBLE INVESTING FROM RETAIL INVESTORS' PERSPECTIVE: INTERNATIONAL EVIDENCE

Guillermo Badía

Faculty of Business and Economics, University of Zaragoza, Gran vía no. 2, 50005, Zaragoza, Spain

Luis Ferruz

Faculty of Business and Economics, University of Zaragoza, Gran vía no. 2, 50005, Zaragoza, Spain

Maria do Céu Cortez

School of Economics and Management, University of Minho, Gualtar 4719-057, Braga, Portugal

Área temática: H) Responsabilidad Social Corporativa

B) Valoración y Finanzas

***Keywords*:** Socially responsible investing; Retail investor; Performance evaluation; Market states.

THE PERFORMANCE OF SOCIAL RESPONSIBLE INVESTING FROM RETAIL INVESTORS' PERSPECTIVE: INTERNATIONAL EVIDENCE

Abstract

This paper investigates the performance of portfolios constructed by socially responsible (SR) retail investors compared to conventional investments. We provide evidence of SRI financial performance at the worldwide level as well as at the regional level, for 5 regions (Americas, Europe except UK, United Kingdom, Pacific and Emerging markets). Furthermore, we analyse the impact of different market states on the financial performance. Over the period 2005 to 2014, our results show that SRI portfolios statistically outperform the conventional investments. During bear market periods the financial performance is neutral for both portfolios, whereas during bull market periods SRI portfolios statistically outperform the conventional investment. We document that this outperformance is related to a positive and statistically significant exposure to the size and value risk factors. At the regional level, the results show statistical differences in the financial performances among regional portfolios. These results point out country-specific factors may affect the relationship between corporate social and financial performance.

Keywords: Socially responsible investing; Retail investor; Performance evaluation; Market states

1. Introduction

Socially Responsible Investing (SRI) interest has increased significantly over the last decades in both academic research (Scholtens, 2015) and practitioner investors' universe (Ferruz et al., 2012; Ooi and Lajbcygier, 2013; Duuren et al., 2016). Investors have are increasingly willing to incorporate into their investment decisions not only financial criteria (returns and risk), but also the non-financial attributes of SRI (Benson and Humphrey 2008; Nicolosi et al. 2014).

The growth of SRI, as Nilsson (2015) notes, is taking place despite some scepticism on its effects, such as a limited set of SRI investment options and loss of portfolio diversification. Increasingly, recent studies, such as Leite and Cortez (2016) and Rehman et al. (2016), point out that there are no differences between the financial performance generated by SRI and conventional investments. Through a meta-analysis, Revelli and Viviani (2015) find that corporate social responsibility attributes in investment portfolio do not affect portfolio financial performance in relation to conventional investments. Fatemi et al. (2015) and Ramanathan (2016) even show that SRI strategies are more profitable than conventional investments. In a review paper on the relationship between corporate social performance (CSP) and corporate financial performance (CFP) between 1996-2015, Javed et al. (2016) indicate that most studies find a positive relationship, in line with other previous meta-analyses, such as Orlitzky et al. (2003) and bibliographic reviews such as Lu et al. (2014).

However, we note that most studies are conducted from the perspective of institutional investors' investment decisions and not from the perspective of retail investors who wish to construct SR portfolios, nonetheless, as Benijts (2010) notes, there has been a considerable increase in the popularity of SRI among retail investors. Nilsson (2015) highlights that retail investors choose to devote, at least part of their funds to investments that include some kind of social or environmental concerns, and that they have become an important factor in shaping SRI. In fact, according to the 2016 Global Sustainable Investment Review it has been a feature of the SRI market in most of the regions that professional institutional investors dominate the market, but interest by retail investors in SRI is continuing to grow. Indeed, the relative proportion of retail SRI investments in Canada, Europe and the United States increased from 13 percent in 2014 to 26 percent at the start of 2016. Over a third of SRI assets in the United States were retail. There are, at least, two issues that are relevant for retail investors. First, in many

studies on SRI obtain data on socially responsible stocks from proprietary and expensive databases. We argue that access to information sources is more limited and restricted for retail investors than for institutional investors. Retail investors have little choice but to use open sources of information which are free to access; whereas institutional investors have access to expensive information sources and databases. Several studies consider the perspective of retail investors in building SR strategies and use free and accessible sustainability information, but we note that they are mainly focused on US and UK equity markets (e.g. Brammer et al., 2009; Edmans, 2011; Filbeck, 2013; Brzeszczyński and McIntosh, 2014; Carvalho and Areal, 2016).

Second, most studies on SRI focus on products such as funds, pension funds or indices (e.g. Statman, 2006; Schröder, 2007; Renneboog et al., 2008; Cortez et al., 2012; Managi et al., 2012; Lean and Nguyen, 2014). In fact, Osthoff (2015) and Leite and Cortez (2016) highlight that most SRI studies focus on the performance of SRI mutual funds. Retail investors may be interested in investing in actively managed SRI mutual funds. However, as Auer and Schumacher (2016) points out, evaluating the impact of incorporating social screens by analysing the performance of mutual funds has limitations. A major problem is that there is some evidence that the label 'socially responsible' might be more a marketing strategy, thus not assuring investors that a SRI fund is really socially responsible. The issue of whether SRI funds are simply conventional funds in disguise has been debated in the literature. For instance, Wimmer (2013) shows that the social level of SRI funds largely disappears after two years. In turn, Utz and Wimmer (2014) find that that, on average, SRI funds do not hold more ethical stocks than conventional funds and that a mutual fund being classified as SRI does not in any assure exclusion of socially controversial firms. The findings of Humphrey et al (2016) reinforce that argument that SRI funds and conventional funds are not so different after all and Statman and Glushkov (2016) even find evidence of closet SRI funds, which are conventional funds that avoid investing in unethical stocks. In this context, retail investors may find it very difficult to know the extent which a SRI fund is really considering social criteria in their selection process. By constructing SR portfolios, retail investors can be more confident the companies that are included in their portfolios are indeed reflecting their social concerns. Furthermore, in countries where mutual funds are marketed by commercial banks, their interests may lead SR private investors towards products that are not suited for their social concerns. Banks

are predominantly oriented to maximize profits and not the interests of depositors and investors. In practice, to achieve the bank's objectives, sometimes insufficient attention is given to the approach to clients. Graafland and Van de Ven (2011), for instance, document that during the credit crisis, in some cases banks did not behave according to the moral standards they set themselves and claim that commercial practices and ethical values of financial professionals played a relevant role in the global financial crisis. Van Hoorn (2015) points out that the financial services industry sometimes provides an environment highly conducive to unethical behaviour.

Considering that worldwide evidence regarding the possibility of SRI retail investors to yield positive financial performance is scarce and the problems socially conscious investors may face when trying to select true SRI funds, we focus on a retail investor's perspective by constructing portfolios on the basis of social criteria. It is important to mention that currently, the technological developments in trading systems have reduced transaction costs and commissions, encouraging retail investors to trade and leading to an increase in the trading volume and liquidity (Butt and Virk, 2017).

The objective of this study is to analyse the performance of portfolios constructed by SR retail investors compared to conventional investments. Following Nilsson's (2015) concerns that SR retail investors need easy-to-use tools on social information, we construct portfolios with stocks listed on the Global-100 'Global-100 Most Sustainable Corporations in the World' list (Global-100, hereafter), which is freely available to the public. We argue that it is essential to use information sources that any retail investor may access, in order to set up an investment portfolio that follows socially responsible investment criteria.

Brzeszczyński and McIntosh (2014) analyse the performance of portfolios of UK stocks listed on the Global-100 and find that the performance of the UK-SRI portfolios is not significantly different from the performance of the market indexes. We contribute to the literature by extending the portfolio performance evaluation of portfolios constructed on the basis of free and available SR information to a worldwide context. This analysis is relevant considering that the patterns of development of SRI are not homogenous across countries (Neher and Hebb, 2015). Furthermore, Hörisch et al. (2015) indicate that country-specific factors tend to affect the relationship between corporate social and financial performance. Our analysis uses all SRI Global-100 stocks without limiting the study to any specific country.

Additionally, we analyse the impact of different market states on the financial performance of SRI portfolios. Recent research shows that the performance of SR equity funds (Nofsinger and Varma, 2014; Becchetti et al., 2015), SR fixed-income funds (Silva and Cortez, 2016), and SR stocks (Brzeszczyński and McIntosh, 2014; Carvalho and Areal, 2016) is sensitive to different market states (e.g., expansion and recession periods). We use a conditional model in line with Carvalho and Areal (2016) and Leite and Cortez (2016) in order to consider for time-varying risk. This model allows both risk and performance to change over different market regimes.

Although this analysis is conducted from a retail investor perspective, nonetheless, of course, institutional investors can take into account the results and conclusions reached in this empirical study for constructing their SRI strategies.

The structure of the paper is as follows: Section 2 presents a short overview of the relevant literature. Section 3 describes the data and Section 4 presents the research methods used. Section 5 contains and details the empirical results and Section 6 summarizes our main findings and offers some concluding remarks.

2. Prior literature

The current SRI literature is large on products such as investment funds, pension funds and indices, and provides conflicting evidence (see for instance Bauer et al., 2005; Geczy et al., 2005; Statman, 2006; Schröder, 2007; Renneboog et al., 2008; Capelle-Blancard and Monjon, 2014; Statman and Glushkov, 2016; Belghitar et al., 2017; Brière et al., 2017). However, previous research which be useful to retail investors for constructing stock SR portfolios by using free and available SRI information is scarce, and focusing mainly on the US and the UK markets. For instance, Filbeck et al. (2009) analyse performance of portfolios composed by the 100 Best Corporate Citizens published by *Business Ethics* magazine over the period 2000-2007. Specifically, they study the stock price reaction to the press releases and the long-term return performance of the portfolios. On the one hand, they find significant positive abnormal returns for stocks that are new listed to the annual listing on the press release date of the survey. On the other hand, they document that the top 100 stocks outperform the S&P 500 over longer holding periods, though their result do not identify statistical differences between SRI and conventional investments. Brammer et al. (2009), using the same list analyzing SR portfolio performance over the period 2000-2004, find that over the year following the announcement, stocks of the 100 Best Corporate Citizens yield negative abnormal

returns. Nevertheless, they identify that can be as a consequence of stock features and when allowing for these firm characteristics, the poor performance of the highly rated firms declines. Moreover, they find that companies in the top 100 but outside the S&P 500 can provide considerable positive abnormal returns.

Edmans (2011) analyses portfolios constructed on the basis of the 100 Best Companies to Work For in America in order to test the relationship between employee satisfaction and long-run stock returns. He shows that companies with stronger employee satisfaction not only had higher risk-adjusted returns in the stock market but also exhibited both higher earnings announcement returns and higher long-term earnings surprises. He reveals positive abnormal returns between 1984 and 2005. Especially Edmans (2011) presents comprehensive evidence that the stock market does not entirely value the intangible assets that companies create through strong relations with their employees. In this sense, Fulmer et al. (2003) also investigate the link between employee relations and firms' performance using the 100 Best Companies to Work For in America. They argue that stronger employee satisfaction could affect positively the stock performance. Over the period 1995 to 2000, they find that the financial performance of stocks in the list was better than a matched peers sample and, generally, statistically significant. Similarly, Filbeck and Preece (2003) document that stocks in this list outperform a matched sample portfolio statistically over the period 1987 to 1999. Carvalho and Areal (2016) investigate portfolios of stocks listed on the 100 Best Companies to Work for in America but in times of financial crises and find that their financial performance and systematic risk remain unaffected in bear markets.

In turn, Anginer and Statman (2010) analyse performance of portfolios composed by the *Fortune* magazine's annual list of America's Most Admired Companies testing the relation between reputation and subsequent returns. Over the period 1983 to 2007, they document that low-ranked stocks outperform high-ranked stocks, and that stocks of firms moving up the reputation scale lagged stocks of firms moving down the scale. Filbeck et al. (2013) investigate whether the fact of being listed on different public surveys of exceptional firms (*Fortune*'s ''Most Admired Companies'' and ''100 Best Companies to Work For,'' *Business Ethics* ''Best Corporate Citizens,'' and *Working Mother*'s ''100 Best Companies for Working Mothers'') adds value to a portfolio and find that Most Admired Companies and Best Corporate Citizens rankings are the most influential.

Outside the US market, Brzeszczyński and McIntosh (2014) investigate whether UK stocks listed on the Global-100 yield higher returns than the FTSE100 and FTSE4Good indices for the period 2000-2010. They find that the returns of the UK-SRI portfolios are higher than the returns of both the FTSE100 and the FTSE4GOOD indexes; but that the difference in returns is not statistically significant.

The studies mentioned above suffer from some limitations. First, all of them are country-specific studies (US and UK). Given the social and demographic country-specific factors (Bauer and Smeets, 2015), SRI financial performance to retail investors should be documented and compared in more regions. Second, except for Brzeszczyński and McIntosh (2014) and Carvalho and Areal (2016), none of previous studies analyse the market state effect on financial performance. Recent research on the performance of SR equity funds, SR fixed-income funds and SR stocks find that portfolio performance is market state dependant. Third, Brzeszczyński and McIntosh (2014), who use UK-stocks of Global-100, use traditional portfolio performance measures and, given the well-known limitations of these methodologies, their results should be interpreted with caution. Nevertheless, they also analyse the ability of Fama and French (1993) three-factor and Carhart (1997) four-factor models to explain performance of UK-SRI portfolios and find that returns of the UK-SRI portfolios cannot be consistently explained by conventional factors other than the market factor. In spite of the fact that they use these models, they do not analyse the statistical difference between UK-SRI alpha portfolios and conventional investment alpha portfolios.

3. Data

In this study, stocks perceived as socially responsible are those that are included in the Global 100. This list starts in February 2005 and provides an annual list of the 100 most sustainable businesses in the world. It is managed by Corporate Knights, who also provides indexing solutions and market-beating portfolios. Global-100 firms are considered to be SR because they demonstrate a greater capacity for proper management of the three factors covered by SRI in their industries: environmental, social and governance (ESG) criteria.

We identify and analyse stocks included in the Global-100 from January 2005 to December 2014. Monthly discrete returns of all stocks are computed on the basis of the total return series (in US dollars) collected from Thomson Reuters database. To evaluate the long-term performance of SRI portfolios, we use the calendar-time portfolio

approach (as in Carvalho and Areal, 2016). This approach involves creating an equally-weighted portfolio of the stocks included in the Global-100 list in each year. Portfolios are rebalanced annually at the end of the month in which a new list is announced: each January before the World Economic Forum in Davos. The list is published on www.global100.org and can be consulted easily and free of charge by any investor interested in building SR investment strategies. Thus, SRI criteria can easily be included in investment decisions without having to implement a complex firm selection process (e.g. screening and engagement).

This paper examines international SR portfolios constructed on the basis of the list. From 2005 to 2014, 26 countries are represented in the sample. Table 1 shows the country stock allocation of the Global-100 during full sample period. We can observe how the UK and the US are the most weighted countries in the sample, 19.40% and 16.72% respectively. In this sense, it appears justified that previous research had focused on these markets; however, the percentage of countries such as Japan (12.54%), Canada (6.27%), and Australia (5.67%), among others, highlights the relevance of analyzing the SRI phenomenon to retail investors on other countries. Furthermore, it is worthwhile noting that the highest percentage (32.54%) of companies is from continental Europe country firms. In spite of the fact that other countries are less represented, it is also interesting to analyse them, since, for instance, the sample collected firms of emerging markets such as Brazil, India, South Korea or Taiwan, reflecting the fact that firms engaging in SRI practices are not restricted to developed markets.

Table 1. Country stock allocation

This table presents the country stock allocation of the Global-100 lists during full sample period. (January 2005 to December 2014). Figures are represented in percentage (%) of the total number of stocks. The *Continental Europe Countries* encompass the percentage of European countries excluding UK.

Country	%	Country	%
Australia	5.67	Japan	12.54
Austria	0.90	Netherlands	1.79
Belgium	1.19	Norway	1.79
Brazil	2.09	Portugal	0.60
Canada	6.27	Singapore	1.79
Denmark	1.79	South Africa	0.60
Finland	2.69	South Korea	0.30
France	5.97	Spain	2.39
Germany	5.07	Sweden	4.18
Hong Kong	0.60	Switzerland	2.69
India	0.90	Taiwan	0.60
Ireland	0.30	United Kingdom	19.40
Italy	1.19	United States	16.72
Continental Europe Countries	*32.54*		

The SRI portfolio financial performance is evaluated relative to the S&P Global 100 Index. This index represents the financial performance of the 100 most important stocks in equity markets at a global level. Specifically, they are the firms with the highest capitalization in the S&P Global 1200, and are considered global businesses as they earn a large portion of their income doing business in different countries. This index was chosen for several reasons. Lydenberg and White (2015) point out that benchmarks are defined by region, size and sector, and consequently, to make a suitable comparison, both the benchmark and the firm's sample must have similar features. On that basis, the scope of firms making up the S&P Global 100 Index is global, with firms from around the world forming part of the index, as is the scope of the Global-100. The number of firms in the S&P Global 100 Index is the same as the Global-100. Their fundamental difference is precisely what we are looking for: i.e. the appeal of following SRI criteria versus capitalization criteria can be evaluated using the S&P Global 100 Index. While the Global-100 firms are rated for specific SRI requirements, the S&P Global 100 Index is for capitalization.

Descriptive statistics on the average monthly returns, standard deviation and risk/reward ratio for the Global-100 portfolio and S&P Global 100 Index are presented in Table 2. Although the Global-100 portfolio yields higher returns than the S&P Global 100 Index in more years, as well as in full sample period, these differences are not statistically significant. As to standard deviations, the Global-100 portfolio presents higher levels of risk than the S&P Global 100 Index in the frequent majority of cases. However, the risk/reward ratio shows that the relation between return and risk (standard deviation in this case) is somewhat better for the Global-100 portfolio than the S&P Global 100 Index.

Table 2. Descriptive statistics

The full sample period is form January 2005 to December 2014. Mean is the monthly arithmetic mean return, SD is the standard deviation. Mean diff (SD diff) is the mean return (standard deviation) of Global-100 portfolio (Global) minus S&P Global 100 Index (S&P) with *p*-values on *t*-tests (*F*-test) of equality of means (standard deviations). Risk/Reward ratio is the total return divided by standard deviation. Portfolios are rebalanced annually at the end of the month in which a new list is announced.

	Mean				SD				Reward/Risk	
	Global	S&P	Mean diff	*t-test*	Global	S&P	SD diff	*F-test*	Global	S&P
2005	0.0077	0.0024	0.0054	*0.5201*	0.0279	0.0223	0.0056	*1.5614*	0.2778	0.1060
2006	0.0222	0.0133	0.0089	*0.9592*	0.0241	0.0213	0.0027	*1.2723*	0.9230	0.6236
2007	0.0052	0.0067	-0.0015	*-0.1287*	0.0290	0.0278	0.0012	*1.0875*	0.1795	0.2416
2008	-0.0404	-0.0407	0.0003	*0.0125*	0.0697	0.0609	0.0087	*1.3064*	-0.5798	-0.6683
2009	0.0282	0.0167	0.0115	*0.3917*	0.0710	0.0727	-0.0017	*1.0486*	0.3973	0.2301

2010	0.0119	0.0022	0.0097	*0.3583*	0.0685	0.0641	0.0043	*1.1389*	0.1741	0.0343
2011	-0.0129	-0.0054	-0.0075	*-0.3386*	0.0557	0.0527	0.0030	*1.1162*	-0.2319	-0.1026
2012	0.0169	0.0078	0.0091	*0.4988*	0.0494	0.0399	0.0094	*1.5296*	0.3422	0.1945
2013	0.0125	0.0162	-0.0037	*-0.2796*	0.0361	0.0291	0.0069	*1.5345*	0.3470	0.5582
2014	-0.0013	0.0001	-0.0014	*-0.1196*	0.0296	0.0253	0.0043	*1.3734*	-0.0431	0.0031
Full period	0.0050	0.0019	0.0031	*0.4890*	0.0461	0.0416	0.0045	*1.2096*	0.1088	0.0463

Transaction costs are not considered in this study for several reasons: (1) the ability of retail investors to seek and negotiate the most favourable and advantageous investment alternatives will determine the final outcome of each investor; (2) transaction costs depends on aspects such as the amount of funds available for investing or the broker that retail investors use; (3) transaction costs affect the returns for retail investors investing in Global-100 stocks and in the S&P Global 100 Index; and (4) recent studies (e.g., Auer and Schumacher, 2016) consider transactions costs and find that this does not alter their main conclusions. Brzeszczyński and McIntosh (2014) point out that transaction costs would have to be disproportionately high to explain performance differences between SRI and conventional investment. Explanations for this can be found by taking a closer look at the changes of the ESG ratings over time, changes do not occur very often (Auer and Sshumacher, 2016), and because trading occurs only once a year and transaction costs are likely relatively trivial (Brammer et al., 2009).

4. Methods

We examine the financial performance with stock market-based measurements in line with Scholtens (2008), Carvalho and Areal (2016) and Leite and Cortez (2016), among others. Several researchers (Barber and Lyon, 1997; Fama, 1998; Loughran and Ritter, 2000) have shown that the magnitude, and sometimes even the sign, of the long-run abnormal returns are sensitive to alternative measurement methodologies. To determine the sensitivity of our results, we examine the financial performance using several approaches.

Sharpe ratio and significance tests

The Sharpe ratio (1966) - the ratio of excess return to standard deviation - is undoubtedly one of the most commonly used financial performance measure in the financial literature. Thus, as a general measure of financial performance and given the well-known interpretation of its results, retail investors may be interested in comparing the performance of alternative investment strategies according to this measure. From two investment portfolios *i* and *j* whose excess returns over the risk-free rate at time *t*

are r_{ti} and r_{tj} respectively, a total of *T* return pairs (r_{1i}, r_{1j}), …, (r_{Ti}, r_{Tj}) are observed. The difference between two Sharpe ratios is given by $\Delta = \mathrm{Sh}_i - \mathrm{Sh}_j = \mu_i/\sigma_i^2 - \mu_j/\sigma_j^2$, where μ and σ^2 are the sample mean and standard deviation respectively. As the value of the Sharpe ratio is really an estimate from historical return data, statistical inference has to be applied in order to compare the two indicators. To this end, previous studies (e.g. DeMiguel and Nogales, 2007; Gasbarro et al., 2007) used the test of Jobson and Korkie (1981) and the corrected version of Memmel (2003). However, this test is not valid if the returns distribution is non-normal or if the observations are correlated over time, both phenomena quite common on financial returns time series data. Recently, Ledoit and Wolf (2008), hereafter LW, propose a studentized time series bootstrap approach that works asymptotically and has satisfactory properties in finite samples. Previous literature (e.g., Hall, 1992; Lahiri, 2003) shows the enhanced inference accuracy of the studentized bootstrap over standard inference based on asymptotic normality. LW propose to test $H_0: \Delta = \mathrm{Sh}_i - \mathrm{Sh}_j = 0$ by inverting a bootstrap confidence interval. A two-sided bootstrap confidence interval with nominal level 1-α for Δ (true difference between the Sharpe ratios) is constructed and if zero is not contained in the interval, then H_0 is rejected at nominal level α. Specifically, LW propose to construct a symmetric studentized time series bootstrap confidence interval. To this end, the two-sided distribution function of the studentized statistic is approximated through the bootstrap by $\mathrm{F}\left(\left|\hat{\Delta} - \Delta\right|/s(\hat{\Delta})\right) \approx \mathrm{F}\left(\left|\hat{\Delta}^* - \Delta\right|/s(\hat{\Delta}^*)\right)$, where Δ is the true difference between the Sharpe ratios, $\hat{\Delta}$ is the estimated difference computed from the original data, $s(\hat{\Delta})$ is a standard error for $\hat{\Delta}$ (also calculated from the original data), $\hat{\Delta}^*$ is the estimated difference computed from bootstrap data, and $s(\hat{\Delta}^*)$ is a standard error for $\hat{\Delta}^*$ (also calculated from bootstrap data). Letting $z^*_{|\cdot|,\lambda}$ be a λ quantile of $\mathrm{F}\left(\left|\hat{\Delta}^* - \Delta\right|/s(\hat{\Delta}^*)\right)$, a bootstrap 1-α confident interval for Δ is given by $\hat{\Delta} \pm z^*_{|\cdot|,1-\alpha}\, s(\hat{\Delta})$. LW note that with heavy-tailed data or data of time series nature, this quantile will typically be somewhat larger than its standard normal counterpart (used in the traditional tests) in small to moderate samples, resulting in more conservative inferences. To generate the bootstrap data, we use the circular block bootstrap of Politis and Romano (1992), resampling blocks of pairs from the observed pairs (r_{ti}, r_{tj}), *t*=1,…, *T*, with replacement. Applying the studentized circular block bootstrap requires a choice of the block size *b* and LW propose to use the calibration procedure of Loh

(1987). LW suggest that M = 5000 bootstrap sequences is sufficient for reliable inference. The standard error $s(\hat{\Delta})$ is calculated through kernel estimation. Specifically, the prewhitened quadratic spectral kernel of Andrews and Monahan (1992). The standard error $s(\hat{\Delta}^*)$ is the natural standard error calculated from the bootstrap data, making use of special block dependence structure. The bootstrap *p*-values are computed as $PV = \{\tilde{d}^{*,m} \geq d\} + 1/M + 1$, where $d = |\hat{\Delta}|/s(\hat{\Delta})$, the original studentized test statistic, $\tilde{d}^{*,m} = |\hat{\Delta}^{*,m} + \hat{\Delta}|/s(\hat{\Delta}^{*,m})$, denote the centered studentized statistic computed form the *m*th bootstrap sample by $d^{*,m}$, *m*=1,…, *M*, and *M* is the number of bootstrap resamples.

Firm features and systematic risk

Another approach to evaluate portfolio performance involves computing alphas from multi-factor models, as in Galema et al. (2008), Brammer et al. (2006), Edmans (2011), and Humphrey et al. (2012). .We examine performance using the four-factor Carhart (1997) model that captures the risk premiums associated with size and value versus growth (as in Fama and French, 1993) as well as a momentum factor motivated by Jegadeesh and Titman (1993). The Carhart four-factor model is expressed by

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_{RMRF} RMRF_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{MOM} MOM_t + \varepsilon_{p,t} \quad (1)$$

where $R_{p,t}$ is the return of the portfolio p on time t, $R_{f,t}$ is the risk-free rate and α_p is the estimated performance measure of the portfolio. In relation to the risk factors, $RMRF_t$ represents market excess returns (relative to the risk-free rate) on time t; SMB_t is the difference between the returns on diversified portfolios of small stocks and large stocks; HML_t is the difference between the returns on diversified portfolios of high book-to-market (value) stocks and low book-to-market (growth) stocks; and MOM_t is the difference between the returns on diversified portfolios of winning and losing stocks in the past year. The betas in the model represent the estimated risk measures associated to the different risk factors: market, size, value and momentum. Finally $\varepsilon_{p,t}$ are the regression's residuals. To construct SMB and HML portfolios, we follow the recent Ferruz and Badía (2017) procedure, hereafter FB. They note that Fama and French (1993) construct portfolios once a year and maintain them invariable during full year; however, variations in the features of firms can occur during any given 12-month period, which is not accounted by the Fama and French procedure. Taking month-to-month data and rebuilding the value and size portfolios at the end of each month, FB yield a more dynamic procedure that enhances the ability of the risk-factors and the model. To construct the MOM portfolio, we use six value-weight portfolios formed on size and prior (2-12) returns. The portfolios are the intersections of two portfolios formed on size and three portfolios formed on prior (2-12) return. The MOM factor is also rebuilt at the end of each month. The monthly size breakpoint is the median market equity and the monthly prior return breakpoints are the 30th and 70th percentiles. Thus, MOM is the average return on the two high prior return portfolios (winners) minus the average return on the two low prior return portfolios (losers).

Geographical analysis

As outlined above, besides analysing performance at the global level, SRI financial performance is analysed at the regional level. Our international sample collected firms from 26 countries and the short country-specific sample in some cases could reduce the power of our tests. In order to mitigate it, we combine our 26 countries into diversified portfolios. Following the MSCI market allocation, we analyse five regions (portfolios): (I) America, that includes the United States and Canada; (II) Europe (except UK), that includes Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden and Switzerland; (III) United Kingdom; (IV) Pacific region, that includes Australia, Hong Kong, Japan, New Zealand and Singapore; and (V) Emerging markets, that includes Brazil, India, South Africa, South Korea and Taiwan. This allocation is akin to the one of Fama and French (1998, 2012) who group countries in regions mainly by geographic location and market integration. We exclude UK firms from the Europe portfolio due to the weight of UK in the full sample and the specificities of the UK market relative to continental Europe. Including UK firms in the Europe portfolio may bias the conclusions on Europe. Furthermore, it allows us to observe the SRI phenomenon on UK market and to compare our results with previous studies.

Market states

Additionally, we analyse the financial performance of SRI portfolios in different market states. Recent research shows that different market states (e.g., expansion and recession periods), affect the performance of SR equity funds, SR fixed-income funds, and SR stocks. We start by identifying the different market states across our sample period using the Pagan and Sossounov (2003), hereafter PS, approach. PS develop a statistical approach to determine peaks and troughs of a stock market index. According to PS, the peaks and troughs represent relatively high and low points of a stock index series during a period of time. A peak is identified at *t* time if the event $PK = [\ln P_{t-8}, \dots, \ln P_{t-1} < \ln P_t > \ln P_{t+1}, \dots, \ln P_{t+8}]$ occurs, where P_t represents the quotation of the relevant stock index, and a trough at time *t* if the event $TH = [\ln P_{t-8}, \dots, \ln P_{t-1} > \ln P_t < \ln P_{t+1}, \dots, \ln P_{t+8}]$ occurs. Consistent with literature, we qualify bear periods as those with a downtrend in the relevant stock market index of at least 20% from peak to trough. This process is recently used in financial studies such as Lee et al. (2013), Leite

and Cortez (2016) and Carvalho and Areal (2016) among others. The MSCI ACWI[1] is used as the relevant stock market index since is a coherent and complete representation that captures the full spectrum of the global equity opportunity set without home bias. The index collects stocks across 23 developed markets and 23 emerging markets. With 2,480 constituents, the index covers approximately 85% of the global investable equity opportunity set. Table 3 shows the global bear market period identified (Global-ACWI). The remaining periods are considered bull market periods. However, since this paper examines international SR stocks returns, we have to be cautious establishing unique global market states.

Considering the different geographic areas of analysis, we thus proceed to identify different market states at the regional level. The relevant stock market indexes used are: the MSCI North America Index (portfolio I: Americas); the MSCI Europe ex UK Index (portfolio II: Europe except UK); the MSCI United Kingdom Index (portfolio III: United Kingdom); the MSCI Pacific Index (portfolio IV: Pacific); and the MSCI Emerging Markets ex China Index[2] (portfolio V: Emerging markets).[3] The regional bear periods are showed in Table 3, and the remaining periods are considered as bull periods.

As expected, the downtrend associated to the international financial crisis (from 2007 to 2009) is observed both at the global market and regional level. Furthermore, we observe another bear market period in Europe ex-UK from May 2011 to May 2012 which can be associated to the euro sovereign debt crisis, as well as another bear market period in emerging markets from May 2011 to September 2011 as a possible financial contagion of fiscal risks in the US and sovereign debt sustainability in Europe.

[1] Index prices are in USD. Data information is obtained from www.msci.com.

[2] The MSCI Emerging Markets Index includes China as the most weighted country. We use the MSCI Emerging Markets ex China Index since China is not included our sample. Furthermore, the most weighted countries in this index are those included in our sample: South Korea 20.62%, Taiwan 16.79%, India 12.11%, Brazil 10.43%, and South Africa 9.09%. Anyhow, we computed the analysis with both indices and obtained exactly the same results.

[3] Prices for all indices are in USD. Data information is obtained from www.msci.com. Indices used for the remaining regions cover the same countries as our regional portfolios. The MSCI North America Index covers US and Canadian firms; the MSCI Europe ex UK Index covers firms from Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden and Switzerland; the MSCI United Kingdom Index covers stocks from UK; and the MSCI Pacific Index covers firms from Australia, Hong Kong, Japan, New Zealand and Singapore.

Table 3. Bear market states

This table identifies periods bear market according to the Pagan and Sossounov (2003) procedure. The sample period studied is from January 2005 to December 2014. The indices used are the MSCI ACWI Index (Global); the MSCI North America Index (portfolio I: Americas); the MSCI Europe ex UK Index (portfolio II: Europe except UK); the MSCI United Kingdom Index (portfolio III: United Kingdom); the MSCI Pacific Index (portfolio IV: Pacific); and the MSCI Emerging Markets ex China Index (portfolio V: Emerging markets). Consistent with literature, we require the rise (fall) of the market being greater (less) than either 20. We test the window breadth for eight, nine and ten months and obtain the same results.

Portfolio	Start date	Index value (Points)	End date	Index value (Points)	Change in market index	Length of bear period (months)
Global-ACWI	2007/11	408.105	2009/02	187.168	-0.5414	16
Americas	2007/11	1 558.805	2009/02	776.949	-0.5016	16
Europe except UK	2007/11	2 452.294	2009/02	985.823	-0.5980	16
	2011/05	1 794.745	2012/05	1 231.996	-0.3472	13
UK	2007/11	1 638.644	2009/02	672.550	-0.5896	16
Pacific	2007/11	2 763.476	2009/02	1 369.571	-0.5044	16
Emerging Markets	2007/11	4 030.146	2009/02	1 610.415	-0.6004	16
	2011/05	3 945.570	2011/09	3 011.914	-0.2366	5

Performance in different market states

To analyse the market state effect on financial performance we use a conditional model in line with Nofsinger and Varma (2014), Carvalho and Areal (2016) and Leite and Cortez (2016). This model allows risk and performance to vary over different market states. The conditional model incorporates two dummy variables in order to obtain different estimated coefficients on different market states, as follows:

$$R_{p,t} - R_{f,t} = \alpha_{Bull} D_{Bull,t} + \alpha_{Bear} D_{Bear,t} + \beta_{1Bull} RMRF_t D_{Bull,t} + \beta_{1Bear} RMRF_t D_{Bear,t} + \beta_{2Bull} SMB_t D_{Bull,t} + \beta_{2Bear} SMB_t D_{Bear,t} + \beta_{3Bull} HML_t D_{Bull,t} + \beta_{3Bear} HML_t D_{Bear,t} + \beta_{4Bull} MOM_t D_{Bull,t} + \beta_{4Bear} MOM_t D_{Bear,t} + \varepsilon_{p,t} \quad (2)$$

where $D_{Bull,t}$ is a dummy variable that takes value of one for bull market periods and zero otherwise and $D_{Bear,t}$ is a dummy variable that takes value of one for bear market periods and zero otherwise. α_{Bull} corresponds to the financial performance on bull markets and α_{Bear} on bear markets. β_{1Bull}, β_{2Bull}, β_{3Bull} and β_{4Bull} correspond to the factor loadings on bull periods, and β_{1Bear}, β_{2Bear}, β_{3Bear} and β_{4Bear} on bear periods. As Leite and Cortez (2016) point out, this procedure extends the model of Nofsinger and Varma (2014) by incorporating the dummy variables both for the alphas and for the risk factors, thereby enabling the analysis of financial performance and risk exposures on different market states.

5. Empirical Results

This section presents the empirical results. Table 4 shows the results of applying the Sharpe ratio and the LW procedure to estimate the statistical significance of the difference between the Sharpe ratio of the SRI portfolio (Global-100 stocks) and of the four-factor Carhart (1997) model to both portfolios. Furthermore, in line with previous studies (e.g. Nofsinger and Varma, 2014; Leite and Cortez, 2016), in order to investigate differences in financial performance between both portfolios, we also estimate the alphas of a portfolio constructed by subtracting the returns of the S&P Global 100 Index from the returns of the Global-100 portfolio

Table 4. Portfolio Financial Performance and Risk estimates

This table shows estimates of performance and risk for the Global 100 portfolio (Global) and the S&P Global 100 Index (S&P). Diff is the portfolio constructed by subtracting the returns of the S&P Global 100 Index from the returns of the Global-100 portfolio. The full sample period is form January 2005 to December 2014. Portfolio performance is evaluated by means of the Sharpe ratio and the alpha from the four-factor Carhart (1997) model. The LW procedure is used to identify statistical significant differences between the Sharpe ratio of both portfolios, and values in brackets represent the *p*-value for equal Sharpe ratios. The Carhart (1997) model is estimated by OLS based on the heterokedasticity and autocorrelation adjusted errors of Newey and West (1987). Portfolios SMB and HML are constructed following FB and MOM following the Carhart approach. The MSCI ACWI Index is the market proxy in the Carhart (1997) model. One-month US T-bills proxy for the risk-free rate. R2 Adj. is the adjusted coefficient of determination. Values in parenthesis are the *t*-statistics. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

	Sharpe	Alpha	RMRF	SMB	HML	MOM	R2 Adj.
Global	0.0751	0.0025**	0.9843***	0.1315***	0.2351***	-0.0194	0.9614
		(2.4432)	(37.0308)	(4.1246)	(4.8310)	(-0.7825)	
S&P	0.0162	-0.0010	0.9432***	0.1153***	-0.1250**	0.0169	0.9530
		(-0.9180)	(31.6051)	(4.2158)	(-2.1312)	(0.5718)	
Diff	0.0589*	0.0034**	0.0411	0.0162	0.3600***	-0.0363	0.4408
	[0.0569]	(2.2868)	(0.9930)	(0.3771)	(4.8316)	(-0.8797)	

Considering our full sample period, the Sharpe estimate for the Global-100 portfolio is 0.0751 and for the S&P Global 100 Index 0.0162, resulting in a difference of 0.0589. The LW test produces a *p*-value of 0.0570, meaning that the difference between the Sharpe ratio of both portfolios is statistically significant. These results are supported by the alpha estimates. The Global-100 portfolio shows a positive and significant alpha and the S&P Global 100 Index yields a negative although not statistically significant alpha. The difference in performance between both portfolios, measured by the alpha of the Diff portfolio, is statistically significant, indicating that the Global 100 portfolio outperforms the S&P Global 100 Index. Thus, both financial performance measures indicate statistically significant differences between SRI and conventional investments, and show that the Global 100 portfolio yields better financial performance than S&P

Global 100 Index. As to market sensitivities, both portfolios have positive and statistically significant exposure to the size factor, showing a tendency for the portfolios to be exposed to smaller firms. Furthermore, the Global 100 portfolio presents significant positive loading on the value factor, whereas the S&P Global 100 Index has significant and negative exposure. Considering the results of the Diff portfolio, we can conclude that, the SRI portfolio is significantly more exposed to value stocks. Regarding momentum factor, we do not find statistically significant coefficients. Our results are in line with previous studies such as Filbeck (2009), Edmans (2011) and Filbeck (2013) and suggest that SR retail investors, and equally institutional investors, are able to benefit from the outperformance of a SRI strategy relative to conventional investments.

The results on the portfolio performance of the SRI portfolios at the regional level are presented in Table 5. Estimates of the Sharpe ratio and four-factor model for each region are reported. With respect to the Sharpe ratios, three portfolios show positives values and two other show negative values for this measure. If the portfolios are ranked by the Sharpe values, portfolio P1 (Americas) yields the highest financial performance, followed by portfolio P2 (Europe ex-UK). Portfolio P5 (Emerging markets) obtains the lowest financial performance, followed by the P4 portfolio (Pacific). The alpha estimations further explore the performance analysis, controlling for the four risk factors. Portfolios P1 and P2 yield a positive and statistically significant alpha (at the 1% level); portfolios P3 and P4 obtain insignificant alphas, and portfolio P5 shows a marginal (at the 10% level) statistically significant negative alpha. These results suggest that the significant differences observed at the global level between the Global-100 portfolio and the S&P Global 100 Index are driven by portfolios P1 and P2. On the other hand, it is possible to appreciate how market sensitivities oscillate notably among regions. The size factor loses relevance in Pacific and Emerging markets; the value factor is only significant in the America and UK area; and momentum effect is documented solely in the American portfolio. Thus, the typical risk factors seem to present a limited capacity to explain some specific regional portfolio returns. Brzeszczyński and McIntosh (2014) document that the returns of the UK-SRI portfolios cannot be consistently explained by conventional factors other than the market factor. However, in contrast, our size and value risk factors constructed via FB are significant for this region. When analyzing American SR stocks, Brammer et al. (2006) find

negative loadings on the market, size, value, and momentum factors, although only size and momentum are statistically significant. In contrast, our results for the P1 portfolio (Americas) point out positive loadings on the market, size and value factors, and negative exposure on momentum, all of them being statistically significant. Using the sample data for constructing the size, value and momentum portfolios following the FB procedure likely has a positive influence on the significance of risk factors. As to the financial performance, our results are in line with previous evidence for the UK market (e.g. Humphrey et al., 2012; Brzeszczyński and McIntosh, 2014) and for the US market (e.g. Edmans, 2011; Filbeck, 2013), and are in contrast with Brammer et al. (2006) and Mollet and Ziegles (2014). Since this study is, as far as we are aware of, the first analyzing SR stocks in pacific and emerging markets focused on retail investor possibilities, our results are novel for this geography.

Table 5. SRI Financial performance and risk at the regional level

This table shows estimates of performance and risk for each regional portfolio. Five regional portfolios are constructed: P1 corresponds to Americas; P2 is Europe ex-UK; P3 is UK; P4 is Pacific; and P5 correspond to Emerging markets. The full sample period is from January 2005 to December 2014. The estimates for the P5 portfolio start in January 2010, considering previously there are no stocks from this region in the sample. Portfolio performance is evaluated by means of the Sharpe ratio and the alpha from the four-factor Carhart (1997) model. The Carhart (1997) model is estimated by OLS based on the heterokedasticity and autocorrelation adjusted errors of Newey and West (1987). Portfolios SMB and HML are constructed for each region specifically following FB and MOM following the Carhart approach. Market proxies are the MSCI North America for P1; the MSCI Europe ex UK for P2; the MSCI United Kingdom for P3; the MSCI Pacific for P4, and Emerging markets ex china for P5. One-month US T-bills proxy for the risk-free rate. R2 Adj. is the adjusted coefficient of determination. Values in parenthesis are the *t*-statistics. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

	Sharpe	Alpha	RMRF	SMB	HML	MOM	R2. Adj.
P1	0.1560	0.0033***	0.9247***	0.2486***	0.0742***	-0.0662***	0.9710
		(2.7030)	(37.4582)	(6.4825)	(3.1980)	(-3.5201)	
P2	0.0738	0.0033***	0.9566***	0.2297***	0.0528	-0.0353	0.9700
		(3.2330)	(42.9523)	(4.7386)	(1.5436)	(-1.0756)	
P3	0.0132	0.0024	0.8650***	0.2585***	0.1164**	-0.0931	0.8988
		(1.4895)	(19.5529)	(4.2252)	(2.1344)	(-1.5885)	
P4	-0.0275	-0.0025	1.0646***	0.0043	0.0420	-0.0280	0.8967
		(-1.5309)	(34.0875)	(0.0874)	(1.2782)	(-1.0155)	
P5	-0.0713	-0.0182*	1.1944***	0.0238	0.1256	-0.0320	0.8722
		(-1.8718)	(24.7482)	(0.3319)	(1.4017)	(-0.4880)	

The statistical differences between Sharpe ratios via the LW procedure by pairs of regions, as well as the alpha of the differences portfolio, also between pairs of regions, are presented in table 6. In the up-right side of the table, we can observe that the differences between the Sharpe ratios of portfolio P1 are statistically significant from those of portfolios P4 and P5. Portfolio P2 also shows statistically significant differences in relation to the Sharpe ratios of portfolios P4 and P5, whereas portfolio P3

yields a Sharpe ratio that is significantly difference from that of portfolio P5. The difference between portfolios P4 and P5 is not significant. Furthermore, in the down-left side of the table, we present the alpha estimates of the difference portfolios between pairs of regions. We can observe how, after controlling for four risk-factors, the alpha of portfolio P1 is statistically different from all other portfolios; the alpha of portfolio P2 is statistically different in relation to portfolios P3 and P5; and the alphas of portfolios P3 and P4 are statistically significant different from that of portfolio P5. These results deepen in the results showed in table 5. It appears that portfolios P1 and P2 are the main drivers of SRI financial performance. Given the statistical differences in financial performances among regional portfolios, in line with previous studies (e.g. Nilsson, 2008; Heimann et al., 2011; Hörisch et al., 2015; Bauer and Smeets, 2015), these results suggest country-specific factors may affect the relationship between corporate social and financial performance.

Table 6. Differences in SRI financial performance and risk at the regional level

This table shows financial performance differences between regional portfolios. Up-right side of the table presents the Sharpe differences between pairs of regions. The LW procedure is used to identify statistical significant differences between the Sharpe ratio of pairs of regional portfolios. Down-left side of the table shows the alpha estimates of the difference portfolios between pairs of regions. Difference portfolios are constructed by subtracting the returns of a regional portfolio from the returns of another one. Alphas are estimated by the four-factor Carhart (1997) model. This model is regressed by OLS based on the heterokedasticity and autocorrelation adjusted errors of Newey and West (1987). P1 corresponds to America's portfolio; P2 to Europe ex-UK; P3 to the UK; P4 to Pacific; and P5 to Emerging markets. The full sample period is from January 2005 to December 2014. Differences with the P5 portfolio are estimated from January 2010, considering previously there are no stocks from this region in the sample. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

	P1	P2	P3	P4	P5
P1	--	0.0822	0.1428	0.1836**	0.2773***
P2	0.0148***	--	0.0605	0.1013**	0.1682*
P3	0.0216***	0.0068*	--	0.0408	0.2067**
P4	0.0191***	0.0044	-0.0024	--	0.1146
P5	0.0140***	0.0129*	0.0135**	0.0191***	--

Finally, in table 7 we present estimates of performance and risk of the Global-100 portfolio, the S&P Global 100 Index, as well as the regional portfolios, across different market states. In panel A, we observe that in bear markets, the alpha is negative, although not statistically significant for both portfolios, indicating a neutral performance. During bull market periods, the Global-100 portfolio (G) yields a positive and statistically significant alpha whereas the S&P Global 100 index (S&P) shows a negative and marginal (at the 10% level) statistically significant alpha. In bull markets, the alpha of the differences portfolio is statistically significant, showing that an

outperformance of the Global-100 portfolio relative to the S&P Global 100 index. In bear market periods, there are no statistical significant differences between the performance of both portfolios. Brzeszczyński and McIntosh (2014) show that SR stock portfolios yield higher mean returns than conventional benchmarks during bull and bear market periods in the UK market, nonetheless the differences were reduced and not statistically significant. By means of a more robust methodology, Carvalho and Areal (2016) find that the financial performance of SR stocks is positive in and not affected during bear market periods. Similarly to them, we document that SR stocks are not affected during bear market periods. However, we also show that SR stocks render better than conventional stocks in bull market periods. The conditional multi-factor model with not only separate alphas but also separate betas on different market states allow us to document that the outperformance of the Global-100 portfolio in relation to the S&P Global 100 Index during bull markets is related to the higher exposure to the size and value factors. Panel B shows the performance and risk of the regional portfolios on different market states. Portfolios P1, P2 and P3 exhibit positive and statistically significant alphas in up markets and present a similar loading exposure to risk factors during these periods. All portfolios have significant positive exposure to size and value risks, although momentum is only significant and negative for the portfolio P1. Negative momentum significant exposure relative to SRI may be related to their more narrowed investment universe (Leite and Cortez, 2015). In contrast, portfolio P4 shows a marginal negative and statistically significant alpha in bear markets, as a consequence of a positive exposure on size and value factor, and negative on momentum. Portfolio P5 is the only one showing a positive and statistically significant alpha in down markets periods, related to a significant positive exposure on the size factor and a significant negative exposure on the value and momentum factors. The regional analysis of performance on different market states allows us to document that the performance of the Global-100 portfolio is mostly influenced by regional portfolios P1, P2 and P3. Overall, the results reinforce the argument in favour of country-specific features on the relationship between corporate social and financial performance.

Table 7. Financial performance on different market states

This table presents estimates of performance and risk of the Global-100 portfolio, the S&P Global 100 Index, as well as the regional portfolios, on different market states, based on the conditional model (equation 2). The model is estimated by OLS based on the heterokedasticity and autocorrelation adjusted errors of Newey and West (1987). The Pagan and Sossounov (2003) procedure is used in order to identify the different market states (bear and bull). G (S&P) corresponds to the Global-100 portfolio (S&P Global 100 index); P1 corresponds to America's portfolio; P2 to Europe ex-UK; P3 to the UK; P4 to Pacific; and P5 to Emerging markets. Diff is the portfolio constructed by subtracting the returns of the S&P Global 100 Index from the returns of the Global-100 portfolio. The coefficients β_1, β_2, β_3 and β_4 represent the factor loadings on the market excess return, size, value and momentum factors, respectively. The full sample period is from January 2005 to December 2014. The estimates for the P5 portfolio start in January 2010, considering previously there are no stocks from this region in the sample, therefore, only the second bear market period is studied. R2 Adj. is the adjusted coefficient of determination. Values in parenthesis are the *t*-statistics. The asterisks are used to represent the statistically significant coefficients at the 1% (***), 5% (**) and 10% (*) significance levels.

Panel A: The Global-100 portfolio and the S&P Global 100 Index.

	α_{Bear}	α_{Bull}	β_{1Bear}	β_{1Bull}	β_{2Bear}	β_{2Bull}	β_{3Bear}	β_{3Bull}	β_{4Bear}	β_{4Bull}	R2 Adj.
G	-0.0049	0.0028**	0.8822***	1.0241***	0.3536***	0.2583***	-0.0683	0.1753***	-0.1506***	0.0183	0.9667
	(-1.1553)	(2.4929)	(12.5231)	(37.8883)	(3.5519)	(4.6964)	(-1.3747)	(4.5157)	(-3.8889)	(0.5639)	
SP	-0.0040	-0.0024*	0.8816***	0.9926***	-0.1916	-0.2117***	0.2111***	0.0471	0.0686	0.0166	0.9551
	(-1.3350)	(-1.9430)	(18.4066)	(31.9604)	(-1.3656)	(-3.4536)	(3.8732)	(1.1775)	(1.3341)	(0.5940)	
Diff	-0.0010	0.0051***	0.0006	0.0315	0.5453***	0.4700***	-0.2794***	0.1283**	-0.2191***	0.0017	0.4837
	(-0.1505)	(3.3375)	(0.0051)	(0.9319)	(3.4629)	(6.3578)	(-4.0557)	(2.0921)	(-3.1430)	(0.0379)	

Panel B: Regional SRI portfolios.

	α_{Bear}	α_{Bull}	β_{1Bear}	β_{1Bull}	β_{2Bear}	β_{2Bull}	β_{3Bear}	β_{3Bull}	β_{4Bear}	β_{4Bull}	R2 Adj.
P1	-0.0025	0.0030**	0.7918***	0.9693***	0.1832**	0.3112***	0.0572*	0.0610**	-0.1268***	-0.0632***	0.9726
	(-0.4074)	(2.3875)	(7.3994)	(33.1152)	(2.4863)	(6.3660)	(1.8707)	(2.1757)	(-4.2221)	(-3.2156)	
P2	0.0004	0.0025*	0.9765***	0.9827***	0.0804	0.3224***	-0.0354	0.0761*	-0.0993***	0.0043	0.9723
	(0.2306)	(1.9502)	(22.7625)	(33.8107)	(0.7803)	(7.4091)	(-1.0696)	(1.8160)	(-3.0631)	(0.1191)	
P3	-0.0100	0.0043***	0.7833***	0.8236***	0.5750***	0.2285***	0.0265	0.1407***	-0.1535**	0.0458	0.9188
	(-1.3206)	(2.8773)	(10.5601)	(18.7884)	(4.4992)	(4.4297)	(0.4112)	(2.8883)	(-2.5930)	(1.1700)	
P4	-0.0119*	-0.0020	1.0495***	1.0701***	0.1557**	0.0378	0.3365***	0.0172	-0.2432***	-0.0118	0.9034
	(-1.7764)	(-1.1310)	(28.3980)	(22.8302)	(2.5524)	(1.0395)	(2.8888)	(0.5827)	(-4.1649)	(-0.3685)	
P5	0.0059***	-0.0129	-0.0280***	1.2431***	0.1911***	0.0192	-0.5914***	0.0650	-0.3284***	-0.0348	0.8703
	(-3.7E+05)	(-1.1872)	(-7.9E+11)	(23.2819)	(-8.3E+13)	(0.2339)	(-2.0E+13)	(0.6633)	(-2.5E+13)	(-0.4922)	

Robustness checks

Finally, in this section we report a variety of supplementary checks in order to verify the robustness of our results. First, alternative risk-free rates are used for the calculation of excess returns. Specifically, we calculate the excess returns using the 1-month European Interbank Offered Rate (EURIBOR) and the UK 1-month T-bill as the risk-free rates. Statistical significant financial performance differences between the Global-100 portfolio and the S&P Global 100 Index are even higher (p-value < 0.01) using

alternative risk-free rates[4]. Second, other commonly used global indices are employed as alternative conventional investment indices. We consider the Russell Global Index, the Thomson Reuters Global Index, the S&P Global 1200 Index, the STOXX Global 1800 Index, the World DataStream Market Index, and the FTSE Global Index, and assess the financial performance differences between the Global-100 portfolio and these additional indices. Statistically significant differences between the 1% and the 10% level are found for all differences portfolios, after controlling for the risk factors. Hence, again, there is strong evidence on the outperformance of SR portfolios relative to conventional investments. Finally, other financial performance evaluation measures are considered. We employ the modification proposed by Ferruz and Sarto (2004) regarding the Sharpe ratio (1966), used in different studies such as Scholz (2007) and Luo et al. (2015). Ferruz and Sarto (2004) note that the Sharpe ratio assumes positive portfolio excess returns. However, this is not always the case. Consequently, when this happens, the Sharpe ratio can present anomalous results. In this context, Ferruz and Sarto (2004) propose a correction to the Sharpe ratio, as follows: $FS_{p,t} = \left(R_{p,t}/R_{f,t}\right)/\sigma_{p,t}$, where $R_{p,t}$ is the portfolio p return on time t, $R_{f,t}$ is the risk-free return on time t, and $\sigma_{p,t}$ is standard deviation of the portfolio p on time t. We also employ the Sortino ratio (Sortino and van der Meer, 1991, Sortino and Price, 1994), used by authors such as Leggio and Lienv (2003), Meligkotsidou et al. (2009) and Auer (2016) to measure performance on the basis of the lower partial moments (LPM). According to the Sortino ratio, risk is measured by the negative deviations of returns in relation to a minimum acceptable return (e.g. zero, the risk-free rate or the average return). In our case, we use a rolling interest rate based on the evolution of the risk-free monthly interest rate. The Sortino specification is $S_{p,t} = R_{p,t} - \varphi/\left(\frac{1}{\mathrm{T}}\sum_{t=1}^{T}\max[\varphi - R_{p,t}\,, 0]^2\right)^{1/2}$, where $R_{p,t}$ is the portfolio p return on time t, and φ is the target return or minimum acceptable return. Using these financial performance measures, we find consistent results. As to the Ferruz and Sarto (2004) proposal, the Global-100 portfolio obtains a value greater than twice that of the S&P Global 100 Index. When we analyse the performance using the LPM, the difference is even greater. These results are limited to a descriptive comment because of the fact that procedures to test statistical significance of differences between measures are not available for these performances measures.

[4] The specific results of this section are not presented for the sake of brevity and because our main results and conclusions are not altered. Nonetheless, detailed results are available upon request.

6. Conclusions

In recent periods there has been a considerable increase in the popularity of SRI among retail investors and, moreover, the technological developments in trading systems, reducing transaction costs and commissions, have encouraged retail investors' trading. Previous evidence on the relation between SRI and financial performance is extensive. Yet most studies are conducted from the perspective of institutional investors and not from the perspective of retail investors who wish to construct SR portfolios. Research on the performance of SR portfolios constructed on the basis of free and available information to investors, which may be useful to retail investors, is somewhat scarce, and focuses mainly the US and the UK markets.

This paper highlights this issue and analyses the performance of SR portfolios constructed on the basis of the Global 100 list over the period 2005 to 2014. Since previous evidence is focused on specific countries we provide evidence of SRI financial performance at the worldwide level as well as at the regional level, for 5 regions (Americas, Europe except UK, United Kingdom, Pacific and Emerging markets). Additionally, since recent research shows that SRI performance can differ across market states, we analyse SRI portfolio performance in periods of bull and bear markets.

Our results show that the Global-100 portfolio outperformance the S&P Global 100 Index. In terms of investment styles, both SRI and conventional investments are more exposed to small firms, whereas SRI is more associated to value firms and conventional investments to a growth stocks.

The results on SRI financial performance and risk at the regional level show statistical differences in the financial performances among regional portfolios. The regional analysis allows us to conclude that the performance of the Global-100 portfolio is mostly influenced by three specific regional portfolios: Americas and Europe ex-UK (positively) and emerging markets (negatively). Thus, our results point out country-specific factors may affect the relationship between corporate social and financial performance. Nevertheless, as a limitation of our study, we do not study the influence of concrete social factors in investment decisions. Market sensitivities oscillate notably among regions and we find that the typical risk factors present a limited capacity to explain some specific regional portfolio returns. The analysis on the differences by pairs of regions highlights statistically significant differences among regional portfolios and further motivates the issue on the effect of country-specific factors.

As to the differences in performance between SRI and conventional investments across different market states, the results show that the financial performance in bear market periods neutral for both portfolios. In bull market periods, the Global-100 portfolio shows a positive and statistically significant performance whereas the S&P Global 100 index yields negative and marginal statistically significant financial performance. As a consequence, the Global-100 statistically outperforms the S&P 100 Index in up markets. Furthermore, we document that this outperformance is related to a positive and statistically significant exposure to the size and value risk factors. The regional analysis in this context shows how the regions present miscellaneous exposures on different market states. Our results are robust to several test related to the use of alternatives risk-free rates, benchmarks indexes, and financial performance measures.

In sum, our empirical evidence indicates that SR retail investors are able to implement a SRI strategy that outperforms the S&P Global 100. In addition, the different results uncovered at the regional level suggest that country-specific factors may affect the relationship between corporate social and financial performance. Finally, we document SRI is not negatively affected in bad times, and that their good times their performance increases, outperforming comparable conventional investments. This study has been performed from a retail investor perspective, but, of course the results are also useful for institutional investors when constructing their SRI strategies.

References

Anginer, D., & Statman, M. (2010). Stocks of admired and spurned companies. *The Journal of Portfolio Management, 36*(3), 71-77.

Auer, B. R. (2016). Do socially responsible investment policies add or destroy European stock portfolio value? *Journal of Business Ethics, 135*(2), 381.

Auer, B. R., & Schuhmacher, F. (2016). Do socially (ir) responsible investments pay? new evidence from international ESG data. *The Quarterly Review of Economics and Finance, 59*, 51-62.

Barber, B. M., & Lyon, J. D. (1997). Firm size, book-to-market ratio, and security returns: A holdout sample of financial firms. *The Journal of Finance, 52*(2), 875-883.

Bauer, R., & Smeets, P. (2015). Social identification and investment decisions. *Journal of Economic Behavior & Organization, 117*, 121-134.

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

Becchetti, L., Ciciretti, R., Dalò, A., & Herzel, S. (2015). Socially responsible and conventional investment funds: Performance comparison and the global financial crisis. *Applied Economics, 47*(25), 2541-2562.

Belghitar, Y., Clark, E., & Deshmukh, N. (2017). Importance of the fund management company in the performance of socially responsible mutual funds. *Journal of Financial Research, 40*(3), 349-367.

Benijts, T. (2010). A framework for comparing socially responsible investment markets: An analysis of the Dutch and Belgian retail markets. *Business Ethics: A European Review, 19*(1), 50-63.

Benson, K. L., & Humphrey, J. E. (2008). Socially responsible investment funds: Investor reaction to current and past returns. *Journal of Banking and Finance, 32*(9), 1850-1859.

Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. *Financial Management, 35*(3), 97-116.

Brammer, S., Brooks, C., & Pavelin, S. (2009). The stock performance of America's 100 best corporate citizens. *The Quarterly Review of Economics and Finance, 49*(3), 1065-1080.

Brière, M., Peillex, J., & Ureche-Rangau, L. (2017). Do social responsibility screens matter when assessing mutual fund performance? *Financial Analysts Journal, 73*(3), 1-14.

Brzeszczyski, J., & McIntosh, G. (2014). Performance of portfolios composed of British SRI stocks. *Journal of Business Ethics, 120*(3), 335.

Butt, H. A., & Virk, N. S. (2017). Momentum profits and time varying illiquidity effect. *Finance Research Letters, 20*, 253-259.

Capelle-Blancard, G., & Monjon, S. (2014). The performance of socially responsible funds: Does the screening process matter? *European Financial Management, 20*(3), 494-520.

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

Carvalho, A., & Areal, N. (2016). Great places to work®: Resilience in times of crisis. *Human Resource Management, 55*(3), 479-498.

Cortez, M. C., Silva, F., & Areal, N. (2012). Socially responsible investing in the global market: The performance of US and European funds. *International Journal of Finance & Economics, 17*(3), 254-271.

DeMiguel, V., & Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research, 57*(3), 560-577.

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The eco-efficiency premium puzzle. *Financial Analysts Journal, 61*(2), 51-63.

Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial Economics, 101*(3), 621-640.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.

Fama, E. F., & French, K. R. (1998). Value versus growth: The international evidence. *The Journal of Finance, 53*(6), 1975-1999.

Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance, 65*(5), 1915-1947.

Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics, 105*(3), 457-472.

Fatemi, A., Fooladi, I., & Tehranian, H. (2015). Valuation effects of corporate social responsibility. *Journal of Banking & Finance, 59*, 182-192.

Ferruz, L., & Badía, G. (2017). Adapting and testing the Fama and French model, with some variations of company characteristics. *Applied Economics Letters, 24*(5), 342-345.

Ferruz, L., & Sarto, J. L. (2004). An analysis of Spanish investment fund performance: Some considerations concerning Sharpe's ratio. *Omega, 32*(4), 273-284.

Ferruz, L., Muñoz, F., & Vargas, M. (2012). Managerial abilities: Evidence from religious mutual fund managers. *Journal of Business Ethics, 105*(4), 503-517.

Filbeck, G., & Preece, D. (2003). Fortune's best 100 companies to work for in america: Do they work for shareholders? *Journal of Business Finance & Accounting, 30*(5-6), 771-797.

Filbeck, G., Gorman, R., & Zhao, X. (2009). The "Best corporate citizens": Are they good for their shareholders? *Financial Review, 44*(2), 239-262.

Filbeck, G., Gorman, R., & Zhao, X. (2013). Are the best of the best better than the rest? the effect of multiple rankings on company value. *Review of Quantitative Finance and Accounting, 41*(4), 695.

Fulmer, I. S., Gerhart, B., & Scott, K. S. (2003). Are the 100 best better? an empirical investigation of the relationship between being a "great place to work" and firm performance. *Personnel Psychology, 56*(4), 965-993.

Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance, 32*(12), 2646-2654.

Gasbarro, D., Wong, W., & Kenton Zumwalt, J. (2007). Stochastic dominance analysis of iShares. *The European Journal of Finance, 13*(1), 89-101.

Geczy, C., Stambaugh, R. F., & Levin, D. (2005). Investing in socially responsible mutual funds.

Global sustainable investment review. (2016). Reviewed July 2017.

Graafland, J. J., & van de Ven, Bert W. (2011). The credit crisis and the moral responsibility of professionals in finance. *Journal of Business Ethics, 103*(4), 605-619.

Hall, P. (1992). *The bootstrap and edgeworth expansion*, Springer, New York.

Heimann, M., Pouget, S., Mullet, É., & Bonnefon, J. (2011). The experimental approach to trust in socially responsible investment funds. *Finance and sustainability:*

Towards a new paradigm? A post-crisis agenda, 169-183. Emerald Group Publishing Limited.

Hörisch, J., Ortas, E., Schaltegger, S., & Álvarez, I. (2015). Environmental effects of sustainability management tools: An empirical analysis of large companies. *Ecological Economics, 120*, 241-249.

Humphrey, J. E., Lee, D. D., & Shen, Y. (2012). Does it cost to be sustainable? *Journal of Corporate Finance, 18*(3), 626-639.

Humphrey, J. E., Warren, G. J., & Boon, J. (2016). What is different about socially responsible funds? A holdings-based analysis. *Journal of Business Ethics, 138*(2), 263-277.

Javed, M., Rashid, M. A., & Hussain, G. (2016). When does it pay to be good–A contingency perspective on corporate social and financial performance: Would it work? *Journal of Cleaner Production, 133*, 1062-1073.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91.

Jobson, J. D., & Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *The Journal of Finance, 36*(4), 889-908.

Lahiri, S. (2003). *Resampling methods for dependent data*, Springer, New York.

Lean, H. H., & Nguyen, D. K. (2014). Policy uncertainty and performance characteristics of sustainable investments across regions around the global financial crisis. *Applied Financial Economics, 24*(21), 1367-1373.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance, 15*(5), 850-859.

Lee, J., Yen, P., & Chan, K. C. (2013). Market states and disposition effect: Evidence from Taiwan mutual fund investors. *Applied Economics, 45*(10), 1331-1342.

Leggio, K. B., & Lien, D. (2003). An empirical examination of the effectiveness of dollar-cost averaging using downside risk performance measures. *Journal of Economics and Finance, 27*(2), 211-223.

Leite, P., & Cortez, M. (2016). The performance of European socially responsible fixed-income funds. *Working paper available at SSRN: abstract 2726094.*

Loh, W. (1987). Calibrating confidence coefficients. *Journal of the American Statistical Association, 82*(397), 155-162.

Loughran, T., & Ritter, J. R. (2000). Uniformly least powerful tests of market efficiency. *Journal of Financial Economics, 55*(3), 361-389.

Lu, W., Chau, K., Wang, H., & Pan, W. (2014). A decade's debate on the nexus between corporate social and corporate financial performance: A critical review of empirical studies 2002–2011. *Journal of Cleaner Production, 79*, 195-206.

Luo, C., Seco, L., & Wu, L. B. (2015). Portfolio optimization in hedge funds by OGARCH and Markov switching model. *Omega, 57*, 34-39.

Lydenberg, S., & White, A. (2015). Responsible investment indexes: Origins, nature and purpose. In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The routledge handbook of responsible investment*, 1st ed., 527-535. Routledge, New York.

Managi, S., Okimoto, T., & Matsuda, A. (2012). Do socially responsible investment indexes outperform conventional indexes? *Applied Financial Economics, 22*(18), 1511-1527.

Meligkotsidou, L., Vrontos, I. D., & Vrontos, S. D. (2009). Quantile regression analysis of hedge fund strategies. *Journal of Empirical Finance, 16*(2), 264-279.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters, 1*, 21-23.

Mollet, J. C., & Ziegler, A. (2014). Socially responsible investing and stock performance: New empirical evidence for the US and European stock markets. *Review of Financial Economics, 23*(4), 208-216.

Neher, A. L., & Hebb, T. (2015). The responsible investment atlas–an introduction. *The Routledge Handbook of Responsible Investment*, 53-57. Routledge, New York.

Nicolosi, M., Grassi, S., & Stanghellini, E. (2014). Item response models to measure corporate social responsibility. *Applied Financial Economics, 24*(22-24), 1449-1464.

Nilsson, J. (2008). Investment with a conscience: Examining the impact of pro-social attitudes and perceived financial performance on socially responsible investment behavior. *Journal of Business Ethics, 83*(2), 307-325.

Nilsson, J. (2015). Stakeholders of responsible investment: Retail investors. In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The routledge handbook of responsible investment*, 1st ed., 485-493. Routledge, New York.

Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking & Finance, 48*, 180-193.

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

Osthoff, P. (2015). What matters to SRI investors? In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The routledge handbook of responsible investment*, 1st ed., 705-724. Routledge, New York.

Pagan, A. R., & Sossounov, K. A. (2003). A simple framework for analysing bull and bear markets. *Journal of Applied Econometrics, 18*(1), 23-46.

Politis, D. N., & Romano, J. P. (1992). A circular block-resampling procedure for stationary data. *Exploring the Limits of Bootstrap*, 263-270.

Ramanathan, R. (2016). Understanding complexity: The curvilinear relationship between environmental performance and firm performance. *Journal of Business Ethics*, 1-11.

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*(3), 302-322.

Revelli, C., & Viviani, J. (2015). Financial performance of socially responsible investing (SRI): What have we learned? A meta-analysis. *Business Ethics: A European Review, 24*(2), 158-185.

Scholtens, B. (2008). A note on the interaction between corporate social responsibility and financial performance. *Ecological Economics, 68*(1), 46-55.

Scholtens, B. (2015). Contemporary issues in responsible finance and investment. In T. Hebb, J. P. Hawley, A. G. Hoepner, A. L. Neher & D. Wood (Eds.), *The routledge handbook of responsible investment*, 1st ed., 575-592. Routledge, New York.

Scholz, H. (2007). Refinements to the Sharpe ratio: Comparing alternatives for bear markets. *Journal of Asset Management, 7*(5), 347-357.

Schröder, M. (2007). Is there a difference? the performance characteristics of SRI equity indices. *Journal of Business Finance & Accounting, 34*(1- 2), 331-348.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business, 39*(1), 119-138.

Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing, 3*(3), 59-64.

Sortino, F. A., & Van Der Meer, R. (1991). Downside risk. *The Journal of Portfolio Management, 17*(4), 27-31.

Statman, M. (2006). Socially responsible indexes. *The Journal of Portfolio Management, 32*(3), 100-109.

Statman, M., & Glushkov, D. (2016). Classifying and measuring the performance of socially responsible mutual funds. *The Journal of Portfolio Management, 42*(2), 140-151.

Ur Rehman, R., Zhang, J., Uppal, J., Cullinan, C., & Akram Naseem, M. (2016). Are environmental social governance equity indices a better choice for investors? an Asian perspective. *Business Ethics: A European Review, 25*(4), 440-459.

Utz, S., & Wimmer, M. (2014). Are they any good at all? A financial and ethical analysis of socially responsible mutual funds. *Journal of Asset Management, 15*(1), 72-82.

Van Duuren, E., Plantinga, A., & Scholtens, B. (2016). ESG integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics, 138*(3), 525-533.

Van Hoorn, A. (2015). The global financial crisis and the values of professionals in finance: An empirical analysis. *Journal of Business Ethics, 130*(2), 253.

Wimmer, M. (2013). ESG-persistence in socially responsible mutual funds. *Journal of Management and Sustainability, 3*(1), 9-15.