CrossMark

The Opportunity Cost of Negative Screening in Socially Responsible Investing

Pieter Jan Trinks[1] · Bert Scholtens[1,2]

Abstract This paper investigates the impact of negative screening on the investment universe as well as on financial performance. We come up with a novel identification process and as such depart from mainstream socially responsible investing literature by concentrating on individual firms' conduct and by studying a much wider range of issues. Firstly, we study the size and financial performance of fourteen potentially controversial issues: abortion, adult entertainment, alcohol, animal testing, contraceptives, controversial weapons, fur, gambling, genetic engineering, meat, nuclear power, pork, (embryonic) stem cells, and tobacco. We investigate an international sample of more than 1,600 stocks for more than twenty years. We then analyze the impact of applying negative screens to a market portfolio. Our findings suggest that the choice for negative screening strategies does matter for the size of the investment universe as well as for risk-adjusted return performance. Investing in controversial stocks in many cases results in additional risk-adjusted returns, whereas excluding them may reduce financial performance. These findings suggest that there are opportunity costs to negative screening.

Keywords Controversial investments · Market capitalization · Responsible investing · Risk-adjusted returns · Screening

✉ Bert Scholtens
   l.j.r.scholtens@rug.nl

[1] Department of Economics, Econometrics & Finance, University of Groningen, PO Box 800, 9700, AV, Groningen, The Netherlands

[2] School of Management, University of Saint Andrews, Saint Andrews, Fife KY16 9RJ, Scotland, UK

Introduction

Sustainability, ethics, and social responsibility are notions that increasingly concern both individual and institutional investors. The domain of socially responsible investing (SRI) persuades investors to align ethical and financial concerns, as well as to impact on firms' environmental, social, and governance (ESG) performance (Renneboog et al. 2008; de Colle and York 2009). To achieve this, responsible investors have developed a variety of innovative strategies, including 'best-in-class' investing, active ownership, and ESG integration (Eurosif 2014). SRI has gradually matured when it became adopted by more and more investors (Sparkes and Cowton 2004). Interestingly though, the original SRI practice of excluding stocks of companies involved in harmful or controversial activities (so-called *sin stocks*) remains the most common SRI strategy today (GSIA 2012; Eurosif 2014).

What does it mean for an investor to employ negative screens on her universe of potential investments? And does it matter for financial performance which particular screens are being employed? In the past decade, numerous empirical studies have been conducted in the field of SRI. However, these studies offer inconclusive results and in several respects are still at an early stage. For instance, current research in SRI lacks sound definitions and metrics for responsibility, which probably forms an impediment for an adequate assessment of the size of SRI and its value to investors, companies, and society as a whole (Scholtens 2014). With respect to the impact on returns, studies on responsible and 'sin' investing have provided conflicting results. While some find positive abnormal returns for sin stocks (e.g., Hong and Kacperczyk 2009), others do not find them at all Lobe and Walkshäusl (2011) and conclude that shunning these sin stocks does not significantly impact

Springer

financial performance (Salaber 2009; Humphrey and Tan 2014). As to the number of stocks (and their combined market capitalization) to be invested in, various studies have suggested that screening has a fairly small impact (Fabozzi et al. 2008; Hong and Kacperczyk 2009; Salaber 2009; Durand et al. 2013a, b; Salaber 2013; Humphrey and Tan 2014). Lastly, most previous research has left unexplored a wide range of issues other than the widely studied combination of tobacco, alcohol, and gambling stocks (known as the *Triumvirate of Sin*).

We employ a comparative analysis on fourteen potentially controversial issues for the period 01/1991–12/2012. In contrast to most other studies, we do not rely on broad industry classification and discarding complete industries, but rather check the fourteen issues at the level of the individual firm. As such, we investigate more than 1,600 controversial stocks. The issues we analyze are abortion/abortifacients, adult entertainment, alcohol, animal testing, contraceptives, controversial weapons, fur, gambling, genetic engineering, meat, nuclear power, pork, (embryonic) stem cells, and tobacco. Some of these issues are highly prevalent reasons for exclusion in responsible investing (e.g., alcohol, weapons, gambling, tobacco, adult entertainment). Other issues are less well-established and institutionalized. Still, these are being used in private mandates of investors, and the number of controversies seems to increase (Eurosif 2014). To clarify, part of our study is closely related to current investment practices as it investigates widely adopted screens. But part of the research might be regarded as somewhat more hypothetical as it investigates the exclusion of particular potentially controversial activities that are not yet widely recognized as controversial in the responsible investment industry, although for several individual investors excluding firms engaging with these activities would well align with their personal values. For such controversial activities, our study highlights the potential financial effect on investment performance.

We find that controversial investments generally yield positive abnormal returns, and that screening produces suboptimal financial performance. Furthermore, in contrast to previous research, we observe screening to be applicable to a large number of stocks, representing substantial market capitalization. Lastly, we show that controversial issues other than the usually studied ones also are material and therefore relevant to the study of responsible and 'sin' investing.

Our paper contributes to the literature on SRI in several ways. First, it opens up for philosophical analyses of the concept of negative screening in SRI. Second, we provide detailed new insights into the characteristics of controversial issues, and the effect negative screens may have on the universe of common stocks available to investors through a unique directly constructed sample of controversial stocks

using the ORBIS database. Third, we include methods that are novel to the literature by presenting least absolute deviation (LAD) estimations.

The remainder of the paper is structured as follows. We first provide a background to responsible investing and negative screening. Then we introduce the data and methods used. Next, we present and discuss our results. We end with a brief conclusion.

Background

In this section, we briefly review the empirical literature and highlight important (but often overlooked) philosophical concerns with negative screening policies.

While there is a large body of literature suggesting a positive relationship between measures of CSR or SRI and stock performance (Sparkes and Cowton 2004; Landier and Nair 2009; Margolis et al. 2009), there is growing evidence to the effect that investing in controversial stocks results in superior performance. Hong and Kacperczyk (2009) find that a classical *Triumvirate of Sin* portfolio (comprising alcohol, gambling, and tobacco stocks) of 156 US stocks outperform industry-comparable stocks during 1965–2006. These controversial stocks might be undervalued as a result of being neglected by norm-constrained investors (e.g., pension funds). The above findings have been confirmed by Durand et al. (2013b) for the US, Salaber (2013) for Europe, Visaltanachoti et al. (2009) for China and Hong Kong, and Capelle-Blancard and Monjon (2014) for France. A global study by Fabozzi et al. (2008) finds similar results for the classical controversial sectors combined with defense, biotech (comprising animal testing, genetic engineering, and ordinary stem cells), and adult entertainment.

However, several studies are unable to confirm the above findings. For instance, Salaber (2009) finds no outperformance for 183 US *Triumvirate of Sin* stocks relative to industry-comparable stocks. Similar results are obtained by Lobe and Walkshäusl (2011). Durand et al. (2013a) even observe negative risk-adjusted returns for controversial stocks in the Pacific-Basin markets. Lastly, outperformance of sin stocks may be contingent on religious and legal factors (Salaber 2013).

With respect to negative screening, the academic literature is inconclusive as well. Some authors have suggested that funds consisting of sin stocks outperform the market as well as screened SRI portfolios (Chong et al. 2006; Jo et al. 2010; Liston and Soydemir 2010; Durand et al. 2013b). A recent study by Humphrey and Tan (2014) directly applies *Triumvirate of Sin* plus Defense screens to benchmark indexes (e.g., the S&P500) and finds that excluding controversial stocks from these indexes does not damage performance. A study by Capelle-Blancard and Monjon (2014) for SRI funds

in the period 2004–2007 finds a curvilinear relationship between sectoral screening intensity and financial performance.

In all, the existing literature needs to be interpreted with caution, and it is important to realize that findings might be specific to country, culture, investor characteristics, time period, and methodological choices (cf. Kumar et al. 2011; Kiymaz 2012; Durand et al. 2013a; Salaber 2013; Scholtens and Sievänen 2013; Hood et al. 2014).

Prior to the empirical issues discussed above at least one philosophical issue deserves our attention. The concern is that, ironically, the practice of screening out controversial sectors itself is controversial. Firstly, the avoidance of controversial stocks by institutional investors might be incongruent with fiduciary duty law when controversial investments outperform responsible ones (Richardson 2013). Secondly, de Colle and York (2009) argue that sector-based negative screening practices of SRI fund managers are not justified as these fail to accurately reflect investors' values and moral orientations. Ethical screening may thereby not be as 'ethical' as its name suggests (cf. Schwarz 2003; de Bruin 2013). For example, de Bruin (2013) observes hardly any solid normative foundations for screens on the alcohol industry. He opposes the common public goods argument for such negative screens by raising concerns with democratic legitimacy and effectiveness. While some investors might rightly refuse to invest in particular industries or firms, this would be based on their personal beliefs and values. However, such a basis applies to a limited number of investors and is problematic when used by involuntary chosen funds (as is the case with many pension schemes and healthcare plans). We think it is likely that sectoral screens other than those on the alcohol industry are built on unsound normative arguments as well.

Considering the instrumental ineffectiveness of negative screening to advance responsible corporate behavior (see e.g., Heinkel et al. 2001), some authors have argued in favor of positive screening practices (e.g., Landier and Nair 2009). In addition, norm-based negative screening could be used to improve ethics and governance issues in global financial markets (Richardson 2013). However, these instances of screening face severe problems as well, including informational complexity, large differences in measuring methods, and a lack of clarity and transparency on specific methodologies (Margolis et al. 2009; Sandberg et al. 2009). And, of course, any form of screening (positive or negative) is vulnerable to the above-mentioned problems as the practice of screening by definition implies the selection of some companies and the exclusion of others. Therefore, other strategies such as engagement might be preferable over negative screening for pragmatic reasons (cf. Woodbridge 2011). In all, while negative screens have been important to the development and some success stories of SRI (Landier and Nair 2009; Richardson 2013), they often seem to lack normative substance. Investors, both individual and institutional, may reconsider their motivations for sector-based screens and critically evaluate whether shunning these stocks would be the best strategy to achieve their stated ends.

In the remainder of this study, we will investigate what could be the financial consequences of investment decisions to shun particular groups of stocks. We do realize that for some investors, financial consequences are not very important or might not play a role at all. Several of the controversial issues we investigate already are widely accepted as screens within responsible investing, such as alcohol, tobacco, and gambling. For others, the responsible investment industry does not yet offer investment vehicles, and investors will have to manage their own investment portfolio in line with their views and values.

Data and Method

We employ a comparative mean-variance analysis on fourteen potentially controversial issues. We investigate the prevalence of 'sin' in the investment universe, and whether negative screening consequently represents a loss in size (i.e., market capitalization). Second, we investigate risk-adjusted returns of controversial stocks and test for differences between controversial stocks and the market, as well as between negatively screened portfolios and the market. In this section, we describe our sample construction process and the methods of investigating size and return performance.

Previous studies have used broad industry classifications to construct 'sin portfolios.' We, however, contend that this approach is questionable for several reasons. First of all, broad industry codes do not capture all potential involvement in particular controversial sectors as 'sin' is not the basis for industry classification, leading to an incomplete representation of the actual universe of controversial stocks. Second, a substantial number of (potential) controversial issues have no industry classifications or are difficult to capture using broad classifications (e.g., abortion, adult entertainment, genetic engineering). Third, contrary to the statements of Salaber (2013), we think that issues other than tobacco, alcohol, and gambling are also of interest to responsible investors and therefore material to the study of controversial stocks.

The fourteen controversial issues are selected and defined analogous to commonly employed definitions of 'sin' by ESG raters and SRI funds (see Renneboog et al. 2008; Fabozzi et al. 2008; MSCI 2013; Capelle-Blancard and Monjon 2014). We want to emphasize that we by no means contend that one should regard these issues as being sinful or immoral, but rather that investors in principle can and in practice frequently seem to do so (see Lobe and Walkshäusl 2011; Capelle-Blancard and Monjon 2014). Appendix 1 provides a brief description of the fourteen issues.

Apart from analyzing individual controversial portfolios we also consider various combinations or clusters, including the classical *Triumvirate of Sin* (alcohol, tobacco, and gambling), the so-called *4Bs* portfolio of 'booze, bets, bombs, and butts' (alcohol, gambling, controversial weapons, and adult entertainment), as proposed by Ahrens (2004), the *Sextet of Sin* portfolio (alcohol, tobacco, gambling, controversial weapons, adult entertainment, and nuclear power), as studied by Lobe and Walkshäusl (2011), and potentially controversial medical stocks, comprising firms engaging with abortion, animal testing, contraceptives, genetic engineering, and embryonic stem cells (labeled as *Medical Sin*).

We construct our sample of controversial stocks as follows. Using the ORBIS company database, we employ a combination of industry classifications (using NAICS and NACE codes) and elaborate keyword search techniques to retrieve potentially controversial stocks for each of the fourteen controversial issues.[1] We refine the retrieved controversial stock lists on a manual basis by checking general, history, and activity descriptions of each company on whether that company indeed satisfies our definitions of 'sinfulness.' In this way, we end up with a preliminary sample of controversial stocks for each of the fourteen issues (see row 1 in Table 1). To circumvent survivorship bias, we include in our sample all dead (delisted) stocks.

The preliminary sample of controversial stocks based on ORBIS data is matched with Thomson Reuters' DataStream end-of-the-month Euro data on returns, market capitalization (price times the number of outstanding common shares), and USdollar–Euro exchange rates, by removing the parts for which either data on returns or market value are not available. We measure returns as the natural logarithm of a stock's *Total Return index* on $t = 0$ divided by the index on $t = -1$. The Total Return index is created by DataStream to depict a stock's theoretical growth in value, assuming that dividends are reinvested. This is in line with Lobe and Walkshäusl (2011) and Salaber (2013). Additionally, although not very common in the literature, we carefully and systematically address zero-returns in stocks' return series. We require a stock's return series to cover at least eleven months of continuous data (which is in line with Salaber 2009, 2013). Furthermore, the zero-returns period must not persist longer than three months or occur more than three times in a possible eleven months series. If these conditions are satisfied, we

---

[1] A potential shortcoming of using the ORBIS dataset would be our inability to directly control for changed company activities, since searches had to be made within product descriptions which are updated to the most recent activities. However, potential biases will be marginal as company involvement in particular controversial issues generally does not change radically over time, and our searches within company history descriptions would capture these changes as well.

Table 1 Number of stocks in various categories after clearing for data availability

	Abortion	Adult entertainment	Alcohol	Animal testing	Contraceptives	Controversial weapons	Fur industry	Gambling	Genetic engineering	Meat	Nuclear power	Pork	Stem cells	Tobacco	Total
Initial list (ORBIS)	6	21	494	118	110	22	46	144	76	708	372	179	20	128	2444
Availability of return data	6	21	399	113	104	21	25	133	75	502	344	119	20	90	1972
Availability of sufficient return data	4	20	377	112	96	20	25	129	74	433	328	113	18	80	1829
Availability of sufficient MV data	4	20	361	112	93	20	22	126	74	403	327	107	18	76	1763
Final sample Non-duplicate: 1634	4	20	361	112	93	20	22	126	74	403	327	107	18	76	1763

replace these 'incidental' zero-returns with the market return. When the zero-returns are non-incidental, the series are deleted up to the point for which conditions are satisfied, keeping the longest available valid series, in some cases implying deleting the stock altogether.

After clearing the sample for available return data (as provided by DataStream), sufficient available return data (considering zero-returns), and available data on market value (DataStream), we are left with a total of 1,763 controversial stocks across fourteen issues, which results in 1,634 stocks after removing duplicates. The table in Appendix 2 shows how these stocks are internationally distributed over the 94 countries in the sample. Our population as well as our final sample consists of companies in developed and emerging markets across the world (30.0 % is from North-America; 28.0 % from Asia; 26.9 % from Europe; 7.2 % from Australasia; 4.8 % from South-America; 3.1 % from Africa). It appears that most controversial stocks are from the US (23.4 %), Australia (6.7 %), Japan (6.5 %), Canada (5.4 %), India and China (5.0 % each).

The total number of controversial stocks in our study is considerably larger than that of previous studies, which typically have around 200 controversial stocks using industry classification-based data on the *Triumvirate of Sin* issues. Lobe and Walkshäusl (2011)'s study is a notable exception, having a sample of 755 sin stocks across the *Sextet of Sin* issues. The large sample size of our study is in part the result of having a broader range of controversial issues. It also is a direct result of our more detailed sample selection process. As such, we are able to look into controversial issues with firms outside the 'usual suspect' industries. For instance, our *Sextet of Sin* sample consists of 930 firms, whereas with Lobe and Walkshäusl (2011), it is 755. Moreover, our sample of *Triumvirate of Sin* stocks is on average more than twice as large as that in existing studies. Table 1 shows the sample clearing process with per issue information on the number of stocks in the remaining investable sample.

In line with the conventional literature, we estimate the risk-adjusted performance of our controversial portfolios using the Carhart (1997) regression model. We refrain from utilizing national market indexes as this would imply adherence to the somewhat ambiguous assumption of completely home-biased investors, i.e., investors do not invest a single euro in stocks which are not listed on their domestic stock exchange. We will follow mainstream finance literature in assuming (semi-strong) globally efficient and well-diversified markets. We use the Fama and French global factors and the US one-month Treasury-bill rate risk-free rate provided by Kenneth French's website. Fama and French (2012) recommend using the global factors in applications to explain the returns on global portfolios, e.g., to evaluate the performance of mutual funds that hold a global portfolio of stocks, as long as the portfolio does not have a strong tilt toward microcaps or toward the stocks of a particular region, in which our sample does not (we took the factors from: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The Fama–French factor data are transformed into Euros using DataStream's Dollar–Euro exchange rates.

We investigate risk-adjusted returns for controversial stocks using a Carhart (1997) regression model. To test the return differences between controversial stocks and the market portfolio, we estimate zero-investment Carhart regressions (see also Hong and Kacperczyk 2009; Salaber 2009). In line with Humphrey and Tan (2014), we analyze the impact of avoiding controversial stocks on market capitalization and risk-adjusted return performance by negatively screening the S&P500. We retrieve the beginning and end-of-the-year constituents list, returns, and market value data for the S&P500 from DataStream. In order to construct the negatively screened portfolio, we delete the stocks belonging to the *TotalSin* portfolio (all fourteen issues) from the end-of-the-year S&P500 constituents. We redo the analysis by only excluding the *Triumvirate of Sin* (alcohol, tobacco, gambling) stocks. The size and significance of the alphas obtained from the zero-investment regressions will provide the abnormal performance of controversial stocks when benchmarked against the market (as in Fabozzi et al. 2008) and the negatively screened market portfolios relative to the unscreened portfolio (see also Humphrey and Tan 2014).

We use value-weighted returns since these are in accord with typical practice of SRI or sin investors and funds (Humphrey and Tan 2014) and are most feasible for institutional investors, which make up for the largest part of the SRI market. Moreover, value-weighting is common in the related literature (e.g., Lobe and Walkshäusl 2011; Salaber 2013). Furthermore, the Fama–French market index is weighted at market value too.

Since our return data are non-normally distributed (as Jarque–Bera and Anderson–Darling normality tests convincingly establish), we use median regression which employs a LAD estimator to fit the conditional median of the response variable. Median regression is a special case of quantile regression (QR).[2] In accord with most of the

---

[2] The value-added of QR is its ability to provide a more complete and more robust description of the conditional distribution of returns than ordinary mean regression analysis does. With fat tailed and asymmetric conditional distributions of the dependent variable, OLS would, moreover, not give robust and unbiased results. LAD estimation, in contrast, is robust to fat tails and skewed data (Koenker and Bassett 1978; Mata and Machado 1996), as it investigates how returns are affected at the median, which is a better measure of location and estimator of central tendency. QR techniques are increasingly popular and have lots of interesting applications in finance (Koenker and Hallock 2001). See Buhai (2005), Koenker (2005), McMillen (2013) for in-depth overviews.

literature, we will carry out the median regressions using the Design Bootstrap procedure (Basset and Chen 2001; Baur et al. 2012) with 10,000 replications.

Results

This section presents and discusses the results. We first go into the size and returns of the various controversial issues and provide a simple comparison with conventional market benchmarks. Then we present the estimation results of the multifactor model.

Table 2 shows the number of controversial stocks in the various categories as well as their combined market capitalization for the overall sample, the US, and stocks in the S&P500. Of the 1,634 controversial stocks, about one-fifth is related to alcohol, meat, and nuclear power each. As to market capitalization, firms engaged with contraceptives and nuclear power have the largest share. Combined these two categories make up more than half of overall market capitalization of the 1634 controversial stocks. Interestingly, the market capitalization of firms in adult entertainment and embryonic stem cells is very small. Firms engaging with meat make up a relatively large number of the total stocks, but their market capitalization is only modest. By and large, the same distribution pattern holds for the US. A notable exception is firms engaging with alcohol and contraceptives. The former are underrepresented in the US and the latter are overrepresented from an international perspective. As a result, the *Triumvirate of Sin, 4Bs,* and *Sextet of Sin* have much less market capitalization in the US sample than in the overall sample. To put figures in Table 2 in perspective, the total number of listed companies in the world at the beginning of 2011 is estimated by the World Federation of Exchanges at 45,595. Hence, the 1,634 controversial stocks make up about 3 % of the overall number of stocks. However, in relation to market capitalization, they make up almost 7 % of the investment universe. In the US, the S&P500 total market capitalization at year-end 2012 was € 9,979 bn. Firms engaging with the fourteen sin issues account for about 12 % of this market capitalization. This means that the employment of at least some negative screens results in a substantial loss in terms of size and potential diversification opportunities. For example, with the *Sextet of Sin*, US investors would forego more than 6 % of the investment universe in terms of market capitalization. Hence, in contrast to Humphrey and Tan (2014), we conclude that there are a large number of controversial stocks to be shunned from the investment universe when applying negative screens.

The market-weighted returns for several controversial clusters and negatively screened portfolios as well as those of conventional benchmarks are closely aligned. All correlation coefficients are high and highly significant (results are available upon request). The lowest correlation with the general market is that of *Medical Sin*, the *Triumvirate of Sin*, and the *4Bs* (all less than 60 %).

Table 3 provides the estimation results from regressing the excess returns of eight value-weighted zero-investment portfolios on the four Carhart (1997) factors.[3] We observe practically all controversial cluster portfolios to significantly outperform the market. In the case of controversial medical stocks, the outperformance is only marginally significant. The last two comparisons in Table 3 show that negative screening results in statistically significant underperformance. These findings suggest that negative screening can have significant financial costs, which are to be regarded as an opportunity cost. The results from Table 3 are well in line with most of the academic literature (cf. Chong et al. 2006; Jo et al. 2010; Liston and Soydemir, 2010; Durand et al. 2013b). The results furthermore support the finding of Capelle-Blanchard and Monjon (2014) that the intensity of sector-based screening significantly impacts risk-adjusted returns. Our findings also contradict Humphrey and Tan (2014) who suggest that SRI funds will neither gain nor lose from employing 'sin' screens. As to controversial investing, we find that not only the classical *Triumvirate of Sin* might offer positive abnormal returns, as was established in Hong and Kacperczyk (2009), but that also investing other controversial clusters can be financially attractive. We next turn to the attractiveness of the individual controversial issues.

Table 4 reports the return performance estimations on the basis of the Carhart model of the fourteen controversial issues as well as that of a number of controversial clusters. It shows that some controversial issues are financially more attractive than others. We find that alcohol, animal testing, contraceptives, fur, genetic engineering, and tobacco display statistically as well as economically significant positive abnormal returns. This underlines as well as specifies the results reported in Table 3. The findings are well in line with the predictions by Heinkel et al. (2001), and Hong and Kacperczyk (2009) that firms with controversial business activities have to come up with extra-financial performance in order to keep attracting investors. In this respect, adult entertainment and stem cells are exceptions as these issues exhibit mildly significant underperformance.

---

[3] Note that the coefficient of interest, the intercept of our model (alpha), provides us with an estimate of the median risk-adjusted returns, not with the average (or expected) risk-adjusted returns. Furthermore, the interpretation of the Adjusted R-squared is not analogous to the standard OLS case (see Koenker and Machado 1999 for details).

Table 2 Number of stocks and total market capitalization in absolute figures and as percentage of the total sample (€bn, 31/12/2012)

	Panel A: Total sample of controversial stocks				Panel B: US controversial stocks			
	#Stocks	%	MC (€bn)	%	#Stocks	%	MC (€bn)	%
Abortion	4	0.2	234.71	6.3	3	0.8	140.16	9.0
Adult entertainment	20	1.2	0.31	0.0	12	3.1	0.16	0.0
Alcohol	361	22.1	651.69	17.6	30	7.8	69.60	4.5
Animal testing	112	6.9	459.39	12.4	73	19.1	315.42	20.3
Contraceptives	93	5.7	1056.03	28.6	43	11.2	662.69	42.7
Controversial weapons	20	1.2	82.29	2.2	10	2.6	50.85	3.3
Fur	22	1.3	28.31	0.8	3	0.8	0.52	0.0
Gambling	126	7.7	88.85	2.4	48	12.5	15.07	1.0
Genetic engineering	74	4.5	194.64	5.3	33	8.6	58.18	3.7
Meat	403	24.7	192.64	5.2	50	13.1	59.24	3.8
Nuclear power	327	20.0	864.29	23.4	69	18.0	296.27	19.1
Pork	107	6.5	68.66	1.9	16	4.2	18.65	1.2
Stem cells	18	1.1	0.79	0.0	12	3.1	0.49	0.0
Tobacco	76	4.7	492.87	13.3	11	2.9	193.53	12.5
TotalSin	1634	100.0	3698.32	100.0	383	100.0	1553.53	100.0
TotalSin ex. Meat/Pork	1233	75.5	3312.18	89.6	335	87.5	1494.29	96.2
Triumvirate of Sin	563	34.5	1232.78	33.3	89	23.2	278.20	17.9
4Bs	527	32.3	822.51	22.2	100	26.1	135.52	8.7
Sextet of Sin	930	56.9	2179.59	58.9	180	47.0	625.47	40.3
Medical Sin	283	17.3	1749.55	47.3	153	39.9	885.72	57.0

	Panel C: Controversial stocks in S&P500			
	#Stocks	%	MC (€bn)	%
Abortion	1	0.2	140.05	1.4
Adult entertainment	–	0.0	–	0.0
Alcohol	7	1.4	29.84	0.3
Animal testing	5	1.0	271.15	2.7
Contraceptives	8	1.6	477.06	4.8
Controversial weapons	3	0.6	46.53	0.5
Fur	–	0.0	–	0.0
Gambling	–	0.0	–	0.0
Genetic engineering	1	0.2	38.40	0.4
Meat	4	0.8	38.57	0.4
Nuclear power	19	3.8	239.71	2.4
Pork	2	0.4	10.43	0.1
Stem cells	–	0.0	–	0.0
Tobacco	4	0.8	183.29	1.8
All controversial stocks	50	10.0	1206.70	12.1
Triumvirate of Sin stocks	10	2.0	213.12	2.1
Complete S&P500	500	100.0	9979.29	100.0

The TotalSin sample includes all the stocks involved with the fourteen controversial issues described in the data section and defined in Appendix 1. The TotalSin ex. Meat/Pork sample is the TotalSin sample minus companies that engage with meat including pork (see definition in Appendix 1). The Triumvirate of Sin refers to the companies involved in alcohol, gambling, and tobacco. The 4Bs portfolio comprises alcohol, gambling, controversial weapons, and adult entertainment. Sextet of Sin refers to alcohol, tobacco, gambling, controversial weapons, adult entertainment, and nuclear power. Medical Sin comprises companies engaging with abortion, animal testing, contraceptives, genetic engineering, and embryonic stem cells

* Numbers are based on respectively the TotalSin sample excluding duplicates (Panel A) and the US stocks within this sample (Panel B). Panel C exhibits the TotalSin and Triumvirate of Sin stocks as benchmarked against the complete S&P500 sample. Note that the numbers for individual sin portfolios and controversial clusters in some cases differ from the numbers of the TotalSin (overall) sample due to the exclusion of duplicates in each benchmark sample. E.g., the Medical Sin portfolio consists of companies involved in multiple issues, and consequently its reported number of stocks and market capitalization are lower than the summation of each individual issue's portfolio

Table 3 Return performance of zero-investment portfolios, 01/1991–12/2012

	α^a	β_{MKT}	β_{SMB}	β_{HML}	β_{WML}	Adj.R^2
TotalSin versus Market						
TotalSin–Market	0.0091**	−0.7080***	−0.2713*	−0.0846	0.0523	0.1747
	(0.0037)	(0.1140)	(0.1620)	(0.1503)	(0.0815)	
TotalSin less Meat/Pork–Market	0.0096***	−0.7504***	−0.3865***	−0.1540	0.0513	0.1842
	(0.0033)	(0.1129)	(0.1349)	(0.1348)	(0.0852)	
Triumvirate of Sin–Market	0.0105***	−0.7632***	−0.2121	−0.0088	0.0686	0.1707
	(0.0032)	(0.1110)	(0.1824)	(0.1861)	(0.1015)	
4Bs–Market	0.0076***	−0.5896***	−0.0101	0.0657	−0.0534	0.1400
	(0.0028)	(0.1069)	(0.174)	(0.1745)	(0.1257)	
Sextet of Sin–Market	0.0054*	−0.5923***	−0.0450	0.2270	0.0158	0.1496
	(0.0030)	(0.1089)	(0.1742)	(0.1589)	(0.0925)	
Medical Sin–Market	0.0091***	−0.8759***	−0.7109***	−0.2718*	−0.0384	0.1982
	(0.0030)	(0.0979)	(0.1574)	(0.1630)	(0.0779)	
Neg.screened S&P versus Unscreened S&P						
S&P less TotalSin–S&P complete	−0.0025***	0.0619***	0.0692***	−0.0216	−0.0104	0.1032
	(−0.0025)	(0.0619)	(0.0692)	(−0.0216)	(−0.0104)	
S&P less Triumvirate of Sin–S&P complete	−0.0021***	0.0112**	0.0102	−0.0058	−0.0101	0.0436
	(−0.0021)	(0.0112)	(0.0102)	(−0.0058)	(−0.0101)	

The Market refers to the return on the Fama–French market index. The TotalSin sample includes all the stocks involved with the fourteen controversial issues described in the data section and defined in Appendix 1. The TotalSin ex. Meat/Pork sample is the TotalSin sample minus companies that engage with meat including pork (see definition in Appendix 1). The Triumvirate of Sin refers to the companies involved in alcohol, gambling, and tobacco. The 4Bs portfolio comprises alcohol, gambling, controversial weapons, and adult entertainment. The Sextet of Sin refers to alcohol, tobacco, gambling, controversial weapons, adult entertainment, and nuclear power. Medical Sin comprises companies engaging with abortion, animal testing, contraceptives, genetic engineering, and embryonic stem cells. S&P refers to the S&P500 index. This table shows the measurement results from regressing the excess returns of various value-weighted zero-investment portfolios on the four Carhart (1997) factors using LAD estimation. Alpha is the intercept, indicating out- or underperformance relative to the Fama–French market portfolio or unscreened S&P500. β_{MKT}, β_{SMB}, β_{HML}, and β_{WML} are the coefficients on the market, Size, Book-to-Market, and Momentum factors, respectively, as described by Fama and French (2012). In brackets are the standard errors obtained using the Design Bootstrap procedure (10,000 replications)

* Statistical significance at the 10 % level. ** Statistical significance at the 5 % level. *** Statistical significance at the 1 % level

[a] We also perform the nonparametric Mann–Whitney U test (Mann and Whitney 1947) to make comparisons between two independent samples, e.g., the excess return on the TotalSin portfolio vis-à-vis the excess return on the market portfolio (in line with Lam et al. 2012; Durand 2013b; Gangi and Trotta 2013). The Mann–Whitney test detects significant outperformance for the Triumvirate of Sin portfolio and mildly significant outperformance for the 4Bs portfolio

Sensitivity Analysis

We employ five additional analyses to check the sensitivity and robustness of our results. We consecutively focus on a different market factor, OLS estimation, beginning- versus end-of-year S&P constituents, pre-crisis versus crisis period, and equally weighted return calculations. We briefly discuss our findings below (the results tables for these analyses are available upon request with the corresponding author).

First, we redid our analyses using a different proxy for the market portfolio, namely the MSCI All Country World Index (ACWI). This index very well reflects our sample distribution. Lam et al. (2012) also use the MSCI ACWI. The index furthermore shows returns very similar to the commonly employed MSCI World index (a correlation of 99.8 % during our study period). We find sin portfolios to exhibit lower outperformance using the MSCI ACWI.

Now, only the Triumvirate of Sin and the 4Bs are significant. Results for individual sin portfolios and for the negative screening analysis remain the same.

Second, we ran our main regressions again using OLS regression analysis, with Newey-West HAC standard errors and automatic lag selection (Newey and West 1987, 1994), in line with other studies (e.g., Hong and Kacperczyk 2009; Liston and Soydemir 2010; Lam et al. 2012; Nofsinger and Varma 2013). We find that compared to the LAD estimations, the use of OLS generally results in a somewhat lower outperformance (of roughly 2 % on an annual basis), yet outperformance remains significant. The results remain similar for the negatively screened (S&P) portfolios. In general, this seems to suggest that our results in Table 3 and 4 give a realistic view of the opportunity costs of negative screening as the results are generally unaffected by the estimation methodology used.

Table 4 Return performance of controversial portfolios, 01/1991–12/2012

	α^a	β_{MKT}	β_{SMB}	β_{HML}	β_{WML}	Adj.R^2
Abortion	0.0068*	0.3592***	−1.0886***	−0.3166	−0.0494	0.0842
	(0.0039)	(0.1329)	(0.1859)	(0.248)	(0.1092)	
Adult entertainment	−0.0136*	0.5653***	0.2314	0.0117	−0.2035	0.0065
	(0.0070)	(0.2043)	(0.381)	(0.2624)	(0.2798)	
Alcohol	0.0093***	0.2629**	−0.1614	0.0707	0.0347	0.0260
	(0.0030)	(0.1100)	(0.1637)	(0.1799)	(0.1255)	
Animal testing	0.0103***	0.0913	−0.8368***	−0.3809**	−0.0116	0.0521
	(0.0039)	(0.1107)	(0.1831)	(0.1783)	(0.0976)	
Contraceptives	0.0118***	0.0930	−0.8306***	−0.3762**	−0.0193	0.0641
	(0.0037)	(0.1151)	(0.1842)	(0.1772)	(0.0962)	
Controversial weapons	0.0084*	0.4646***	−0.2358	0.2124	0.0339	0.0273
	(0.0048)	(0.1502)	(0.2549)	(0.2143)	(0.1411)	
Fur	0.0119**	1.0470***	0.0168	−0.1314	−0.2698	0.1245
	(0.0055)	(0.2189)	(0.3339)	(0.2799)	(0.1655)	
Gambling	0.0056	0.6456***	0.6832***	0.0302	−0.0463	0.1314
	(0.0038)	(0.1521)	(0.1921)	(0.1729)	(0.122)	
Genetic engineering	0.0100***	0.7195***	0.0253	−0.1314	−0.0004	0.0982
	(0.0034)	(0.1206)	(0.2984)	(0.1901)	(0.139)	
Meat	0.0050	0.4583***	−0.2999	0.2604*	−0.0539	0.0566
	(0.0038)	(0.1496)	(0.2001)	(0.1483)	(0.1001)	
Nuclear power	0.0026	0.5789***	−0.0047	0.3247***	−0.0014	0.1257
	(0.0025)	(0.0936)	(0.1144)	(0.1091)	(0.0673)	
Pork	0.0023	0.4804***	−0.4399	0.3107	−0.005	0.0228
	(0.0048)	(0.1823)	(0.3085)	(0.2749)	(0.1565)	
Stem cells	−0.0173*	0.9776***	3.2035***	−1.4949***	0.0791	0.1594
	(0.0102)	(0.2436)	(0.5132)	(0.3335)	(0.2132)	
Tobacco	0.0166***	0.1724	−0.4197**	0.0644	0.0554	0.0076
	(0.0046)	(0.1166)	(0.176)	(0.1582)	(0.0998)	
TotalSin	0.0104***	0.2957***	−0.2368	−0.0600	0.0419	0.0550
	(0.0035)	(0.1115)	(0.1615)	(0.1552)	(0.0835)	
TotalSin excl. Meat/Pork	0.0112***	0.2633**	−0.4374***	−0.1725	0.0655	0.0659
	(0.0032)	(0.1118)	(0.1408)	(0.1398)	(0.0880)	
Triumvirate of Sin	0.0124***	0.2410**	−0.3110*	0.0461	0.0613	0.0209
	(0.0032)	(0.1099)	(0.1827)	(0.1917)	(0.1015)	
4Bs	0.0090***	0.4150***	−0.0102	0.1110	−0.0421	0.0385
	(0.0030)	(0.1085)	(0.1796)	(0.1785)	(0.1267)	
Sextet of Sin	0.0069**	0.3984***	−0.0418	0.1925	0.0298	0.0683
	(0.0030)	(0.1096)	(0.1715)	(0.1595)	(0.0900)	
Medical Sin	0.0105***	0.1387	−0.7400***	−0.2863*	−0.0301	0.0689
	(0.0032)	(0.0990)	(0.1603)	(0.1593)	(0.0777)	

TotalSin sample includes all the stocks involved with the fourteen controversial issues described in the data section and defined in Appendix 1. The TotalSin ex. Meat/Pork sample is the TotalSin sample minus companies that engage with meat including pork (see definition in Appendix 1). The Triumvirate of Sin refers to the companies involved in alcohol, gambling, and tobacco. The 4Bs portfolio comprises alcohol, gambling, controversial weapons, and adult entertainment. The Sextet of Sin refers to alcohol, tobacco, gambling, controversial weapons, adult entertainment, and nuclear power. Medical Sin comprises companies engaging with abortion, animal testing, contraceptives, genetic engineering, and embryonic stem cells. This table shows the measurement results from regressing the excess returns of the value-weighted individual and combined controversial portfolios on the four Carhart (1997) factors using LAD estimation. Alpha is the intercept, indicating the size and significance of abnormal returns. β_{MKT}, β_{SMB}, β_{HML}, and β_{WML} are the coefficients on the market, Size, Book-to-Market, and Momentum factors, respectively, as described by Fama and French (2012). In brackets are the standard errors obtained using the Design Bootstrap procedure (10,000 replications)

* Statistical significance at the 10 % level. ** Statistical significance at the 5 % level. *** Statistical significance at the 1 % level

[a] We furthermore perform a nonparametric one-sample Wilcoxon Signed Rank test (Wilcoxon 1945). Results are roughly similar to the LAD estimations, yet the Wilcoxon test labels all sectors significant except adult entertainment and stem cells

Third, we redid the screening analysis using beginning instead of end-of-year index constituents for screening out controversial stocks. Here, the results remain similar and hence appear to be robust in this respect.

Fourth, we considered the 'Great Recession,' which indicates the financial crises period 12/2007–12/2012 (in line with Salaber 2009; Nofsinger and Varma 2013). We compare this with the entire (pre-crisis) period until 11/2007.[4] We establish that our results are in line with Salaber (2009) and Campus (2013). Nearly, all combined controversial portfolios beat the market during the recessionary period in an economically significant way (on average they reach an annualized outperformance of 14.1 %). However, only with the Triumvirate of Sin portfolio, these results are statistically significant. With respect to the individual controversial portfolios, we find alcohol, contraceptives, fur, meat, and tobacco stocks to perform particularly well (but statistical significance can only be established for alcohol stocks). In addition, controversial stocks on average show increased ordinary and downside risk, but this increase is about twice as small compared to the market index. This is in line with the notion of sin stocks being recession-proof investments (e.g., Ahrens 2004). However, we find no evidence of reduced downside risk for either responsible or sin investments during the crisis. For the pre-crisis period, the outperformance of controversial investments is lower. Still, most controversial clusters show statistically significant outperformance. Individual issues that drive these results are alcohol, contraceptives, controversial weapons, genetic engineering, and tobacco. In addition, both the Triumvirate of Sin and Medical Sin portfolios beat the market in a statistically significant way. As to responsible investments (negatively screened S&P500), we find no significant abnormal returns during the crisis period (contrary to Gangi and Trotta 2013; Nofsinger and Varma 2013). For the pre-crisis period, negative screening nevertheless leads to significant underperformance (in line with Nofsinger and Varma 2013). Please keep in mind the fact that in the crisis period the market as a whole showed a lot of turbulence (by definition). As a result, the relative performance of the controversial stocks seems to have improved in a period when the market as a whole performed very badly.

In our last sensitivity analysis, we calculated returns using equally weighted portfolios instead of value-weighted ones.

Here, we arrive at different results. With equal weights, we find no out or underperformance of controversial clusters. On the part of the individual controversial portfolios, only tobacco stocks keep statistically significant outperformance, whereas gambling, pork, and stem cells underperform. Controversial clusters now are found to underperform the market but not in a statistically significant way. Results for negative screening remain very similar relative to our main analysis. The differences in measurement results on the basis of portfolio composition do not have to be problematic. As argued for above, equal-weighting might not be feasible or desirable, particularly for institutional investors. In all, our findings are in part contingent on return averaging methods and estimation techniques. Yet, we have justified the preference of our methods over alternatives by referring to the use and purpose of our study, which renders our main analysis empirically adequate (van Fraassen 2008).

Conclusion

Portfolio management relies on diversification. The default investment is the market portfolio, which is a value-weighted portfolio of all investable securities. But investors are a heterogeneous group, which increasingly wish their investments to reflect their values and beliefs. This can result in excluding particular firms and/or industries. We investigate the impact of such negative screening on their investment universe as well as on financial performance. We investigate fourteen potentially controversial issues: abortion, adult entertainment, alcohol, animal testing, contraceptives, controversial weapons, fur, gambling, genetic engineering, meat, nuclear power, pork, (embryonic) stem cells, and tobacco. Some of these issues are well-established reasons for exclusionary screens in responsible investment portfolios. Others are less prevalent but are already used in private mandates. More and more controversial issues are likely to come play a role in responsible investing. We employ four-factor time-series mean and median regressions using a uniquely constructed sample of 1763 stocks across the fourteen potentially controversial issues over the period 01/1991–12/2012.

In case an investor employs negative screens, we find that the universe of investment objects can become substantially smaller. The extent to which this is the case highly depends on the screen being applied. For example, screening for adult entertainment, fur, and stem cells does have a very limited impact on the investment universe, both in terms of number of stocks and in terms of market capitalization excluded. However, with screens for alcohol and nuclear power, investors forego both a large number of investment objects as well as substantial market capitalization. Most studies so far either argued or assumed screening had only limited impact (e.g., Salaber 2009; Lobe and Walkhäusl 2011; Humphrey

---

[4] We conduct Quant-Andrews tests on our regression outputs to detect possible structural breakpoints in our data (see Andrews 1993). We were unable to establish potential breakpoints in our time series, implying that stock performance has been relatively even over the entire study period. Consequently, we use an ex-ante starting and ending period of the recent global financial crisis. We ignore intermediate periods of crises, the most prominent example of which is the Dot-Com crisis, as it is hard to find an objective demarcated time period of those crises for our global sample.

and Tan 2014). We, however, contend that it does matter, depending on the market and the screen.

Furthermore, we establish that there seems to be a price to screening, namely the opportunity cost of refraining from investing in controversial firms. Again, this cost is dependent upon the type of screen. For example, while we find most controversial clusters to significantly outperform the market, this is not the case for the so-called *Sextet of Sin* (alcohol, tobacco, gambling, controversial weapons, adult entertainment, and nuclear power). These findings are in line with, for example, Hong and Kacperczyk (2009) and Lobe and Walkshäusl (2011). In addition, we find a screened market portfolio (S&P500) to significantly underperform the unscreened market portfolio when accounting for conventional risk factors. In the end, it is for the investor to decide whether this 'price' is worth it. The impact of screening on performance shows that there indeed can be a trade-off between values and beliefs on the one hand and financial returns on the other. If the investor really wants to refrain from particular activities, screening is a practical tool to do so.

A limitation of our study is that it involves an enormous amount of work to arrive at a list of firms that engage with controversial issues which in the end is sensitive to subjective assessment. This contrasts with simply selecting particular industries and shunning all the firms in such an industry. However, our approach is much more detailed and exact. Since the future for SRI and negative screening will probably lie in putting more and more issues under public scrutiny (e.g., factory farming, coal, palm oil), we need insights into the relative performance of these issues as well. This might involve exploring the impact on private equity and debt markets, as some industries engaging with particular controversial issues primarily rely on these forms of financing. Additionally, an important next step is the study of the relative pay-offs and effectiveness of various other SRI strategies, even within 'sin' investing (cf. Cai et al. 2012). Our future research will be directed at the potential effects of new screens as well as on the effects of responsible and controversial investing on CSR and company behavior.

Appendix 1

Definitions of controversial issues

The demarcation of specific controversial issues or 'sins' is an inevitably subjective process. Differences exist firstly in the specific definitions of the sin screens employed by different ESG raters and SRI funds. For instance, there may be differing opinions on whether normal and emergency contraceptive pills should be regarded as abortifacients. Secondly, the sin categories may be treated separately or combined. We combine, e.g., abortion providers with abortifacients providers and embryonic stem cells with human cloning, as their topic of concern is identical.

Our attempt is to employ screens that are as 'objective' as possible and share universal consensus to the highest possible extent. We base the analysis on strict and narrow definitions of 'sin,' which are in line with commonly employed definitions of 'sin' by ESG raters and SRI funds (see Renneboog et al. 2008; Fabozzi et al. 2008; MSCI 2013). No revenue or ownership threshold items are included however (which is common among practitioners). This means that, for instance, Wal-Mart, which might not derive more than 10 % of its revenue from selling alcohol or tobacco, is not selected in our exclusion list of these issues. We see no fundamental reasons why thresholds *should be* employed by sin-shunning SRI investors. There is no company "only 10 % involved'' or "10 % sinful'', which would then mysteriously imply a complete acquittal of sin. Nonetheless, it may in some cases be hard to draw a line at where involvement exists and where it does not. For instance, producers of mobile phones, computers, and games may all be (indirectly) related to the production or distribution of adult entertainment (e.g., Woodbridge 2011). We therefore only select companies that are targeted at, i.e., 'directly' and 'obviously' (from the perspective of the average investor) involved in, particular issues and elaborately motivate each demarcation. Investors that consciously do not want to invest in—say—abortifacients or contraceptives generally will not hold stocks of Pfizer or Sanofi, irrespective of whether the companies in question derive more or less part of their revenue of those particular products. This enterprise has as its main advantages that it approaches objectivity and universality and is directly relevant to and applicable by the SRI investor, and all measurement results remain comparable and insightful. This means that relative to the purpose and use of our study, it adequately represents the real world (see van Fraassen 2008).

Below we briefly list the definitions of the various screens. Retrieved exclusion lists from ORBIS are carefully checked, particularly on whether company activity descriptions include the desired results, whether the company is really ''sinful'' according to our definitions.

Abortion/Abortifacients

Companies owning or operating facilities where abortions are performed, abortion providers, abortifacient manufacturers.

This does not include contraceptives, insurance companies that pay for elective abortions, and companies that provide financial support to Planned Parenthood. *Note* There are differences in what individual investors might view as abortion/abortifacient. Birth control pills/contraceptives are not selected here but may be excluded by some investors who view ('normal' and emergency) contraceptive pills as abortifacient.

Adult Entertainment

Companies targeted at the production or distribution of sexually explicit products and services, i.e., X-rated films, online products, production studios, printed materials, TV or radio programs, and adult clubs or bars. *Note* A narrow definition of adult entertainment is employed which does not include broad entertainment or on-demand video provider companies with marginal links to the adult entertainment industry.

Alcohol

Companies that have as its business the production and/or distribution of alcoholic products, including breweries, wineries, alcoholic beverage stores, wholesalers, and drinking places and excluding supermarkets, restaurants, etc.

Animal Testing

Companies do research or perform tests on animals for medical and cosmetic reasons (to determine safety and efficacy of particular products).

Contraceptives

Companies involved in the manufacturing of contraceptives, e.g., birth control pills, IUDs, sterilization procedures providers, condom manufacturers, and so on.

Controversial Weapons

Companies involved in nuclear, biological, chemical weapons, cluster munitions, and antipersonnel mines. This does not include companies that target at detection, safety, and other products or services. *Note* Controversial weapons can include different things in different countries. The case is further complicated as—according to international humanitarian and criminal law—also the trade in conventional arms can be regarded controversial if these weapons are destined for countries where human rights are violated or where genocide, war crimes, and crimes against humanity are committed. Nonetheless, for the purpose of most SRI investors, there appears to be a great degree of consensus about the above definition. See, e.g., "*Delta Lloyd AM excluded companies*" and MSCI (2013).

Fur industry

Companies that manufacture, sell, or distribute fur products.

Gambling

Companies that manufacture, own, or operate gambling machines or equipment, casinos, lotteries, betting activities, and so on. This excludes operators or owners of restaurants, hotels/motels, and broad (or non-gambling) entertainment activities.

Genetic Engineering

Companies perform genetic engineering or modification techniques for medical or agricultural or other purposes.

Meat

Companies involved in slaughtering, fishing, and processing of meat products. This does not include companies that are only involved in hogs, breading, etc.

Nuclear Power

Companies operating, constructing, or owning nuclear power plants or utilities, as well as companies involved in uranium mining.

Pork

Companies involved in the production, procession or wholesale distribution of pork products, in line with the 'Meat' sector's definition.

Stem Cells

Companies involved in (research in) *embryonic* stem cells, as well as human cloning. *Note* We account for human embryonic stem cell research, as this is the form of stem cell research controversy is generally about. Similarly, cloning research is limited to human cloning. SRI funds (e.g., in Germany) may screen on this particular instance of cloning.

Tobacco

Companies involved in the production, processing and wholesale distribution of tobacco products. This does not include broad stores, supermarkets, and other threshold items.

Appendix 2

Countries included in the sample of controversial stocks

	#Stocks in sample	% of total
UNITED STATES	383	23.4
AUSTRALIA	110	6.7
JAPAN	106	6.5
CANADA	89	5.4
INDIA	82	5.0
CHINA	81	5.0
UNITED KINGDOM	72	4.4
FRANCE	52	3.2
KOREA. REPUBLIC OF	45	2.8
GERMANY	40	2.4
POLAND	24	1.5
MALAYSIA	23	1.4
RUSSIAN FEDERATION	23	1.4
BULGARIA	20	1.2
GREECE	18	1.1
CROATIA	17	1.0
BRAZIL	16	1.0
SOUTH AFRICA	16	1.0
UKRAINE	15	0.9
CHILE	14	0.9
SPAIN	13	0.8
NORWAY	13	0.8
SWEDEN	13	0.8
BERMUDA	12	0.7
SRI LANKA	12	0.7
PERU	11	0.7
THAILAND	11	0.7
TAIWAN. PROVINCE OF CHINA	11	0.7
EGYPT	10	0.6
ISRAEL	10	0.6
CAYMAN ISLANDS	10	0.6
PHILIPPINES	10	0.6
SWITZERLAND	9	0.6
SINGAPORE	9	0.6
ARGENTINA	8	0.5
ICELAND	8	0.5
NETHERLANDS	8	0.5
ROMANIA	8	0.5
SERBIA	8	0.5
AUSTRIA	7	0.4
BANGLADESH	7	0.4
DENMARK	7	0.4
FINLAND	7	0.4
JORDAN	7	0.4
MEXICO	7	0.4

continued

	#Stocks in sample	% of total
NEW ZEALAND	7	0.4
OMAN	7	0.4
TURKEY	7	0.4
BELGIUM	6	0.4
HONG KONG	6	0.4
HUNGARY	6	0.4
ITALY	6	0.4
LITHUANIA	6	0.4
PORTUGAL	6	0.4
INDONESIA	5	0.3
IRELAND	5	0.3
NIGERIA	5	0.3
CYPRUS	4	0.2
CZECH REPUBLIC	4	0.2
KUWAIT	4	0.2
PAKISTAN	4	0.2
SAUDI ARABIA	4	0.2
SLOVENIA	4	0.2
BOSNIA AND HERZEGOVINA	3	0.2
COSTA RICA	3	0.2
GIBRALTAR	3	0.2
LUXEMBOURG	3	0.2
MOROCCO	3	0.2
MAURITIUS	3	0.2
VIRGIN ISLANDS, BRITISH	3	0.2
ZAMBIA	3	0.2
UNITED ARAB EMIRATES	2	0.1
ECUADOR	2	0.1
ESTONIA	2	0.1
GHANA	2	0.1
KENYA	2	0.1
MACEDONIA, FORMER YUGOSLAV REPUBLIC OF	2	0.1
PALESTINIAN TERRITORY, OCCUPIED	2	0.1
VENEZUELA, BOLIVARIAN REPUBLIC OF	2	0.1
ZIMBABWE	2	0.1
BAHRAIN	1	0.1
BOTSWANA	1	0.1
CÔTE D'IVOIRE	1	0.1
FAROE ISLANDS	1	0.1
JERSEY	1	0.1
LATVIA	1	0.1
MONACO	1	0.1
MONTENEGRO	1	0.1
MALTA	1	0.1
NAMIBIA	1	0.1
QATAR	1	0.1
RWANDA	1	0.1

continued

	#Stocks in sample	% of total
TUNISIA	1	0.1
UGANDA	1	0.1

References

Ahrens, D. (2004). *Investing in vice: The recession-proof portfolio of booze, bets, bombs, and butts*. New York: St. Martin's Press.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica, 61*(4), 821–856.

Bassett, G. W, Jr, & Chen, H. (2001). Portfolio style: Return-based attribution using quantile regression. *Empirical Economics, 26*(1), 293–305.

Baur, D. G., Dimpfl, T., & Jung, R. C. (2012). Stock return autocorrelations revisited: A quantile regression approach. *Journal of Empirical Finance, 19*(2), 254–265.

Buhai, I. S. (2005). Quantile regression: Overview and selected applications. *Ad-Astra-The Young Romanian Scientists Journal, 4*(1), 1–17.

Cai, Y., Jo, H., & Pan, C. (2012). Doing well while doing bad? CSR in controversial industry sectors. *Journal of Business Ethics, 108*(4), 467–480.

Campus, R. P. (2013). *Do Sin Stocks outperform the market during recessionary and expansionary periods?* Unpublished working paper, University of Puerto Rico. http://cbt2.nsuok.edu/kwok/conference/submissions/swfa2013_submission_112.pdf. Accessed 2 Jan 2014.

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, C. (1997). On persistence in mutual fund performance. *Journal of Finance, 45*(5), 57–82.

Chong, J., Her, M., & Phillips, G. M. (2006). To sin or not to sin? Now that's the question. *Journal of Asset Management, 6*(6), 406–417.

de Bruin, B. (2013). Socially responsible investment in the alcohol industry: An assessment of investor attitudes and ethical arguments. *Contemporary Social Science, 8*(1), 58–70.

de Colle, S., & York, J. G. (2009). Why wine is not glue? The unresolved problem of negative screening in socially responsible investing. *Journal of Business Ethics, 85*(1), 83–95.

Durand, R. B., Koh, S., & Limkriangkrai, M. (2013a). Saints versus Sinners. Does morality matter? *Journal of International Financial Markets, Institutions and Money, 24*(4), 166–183.

Durand, R. B., Koh, S., & Tan, P. L. (2013b). The price of sin in the pacific-basin. *Pacific-Basin Finance Journal, 21*(1), 899–913.

Eurosif (2014). *European SRI study 2014*. http://www.eurosif.org/semantics/uploads/2014/09/Eurosif-SRI-Study-2014.pdf. Accessed 10 Oct 2014.

Fabozzi, F. J., Ma, K. C., & Oliphant, B. J. (2008). Sin stock returns. *The Journal of Portfolio Management, 35*(1), 82–94.

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

Gangi, F., & Trotta, C. (2013). The ethical finance as a response to the financial crises: An empirical survey of European SRFs performance. *Journal of Management and Governance,*. doi:10.1007/s10997-013-9264-7.

Global Sustainable Investment Alliance (2012). *Global sustainable investment review*. http://gsiareview2012.gsi-alliance.org/. Accessed 8 Oct 2014.

Heinkel, R., Kraus, A., & Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis, 36*(4), 431–449.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93*(1), 15–36.

Hood, M., Nofsinger, J. R., & Varma, A. (2014). Conservation, discrimination, and salvation: Investors' social concerns in the stock market. *Journal of Financial Services Research, 45*(1), 5–37.

Humphrey, J. E., & Tan, D. T. (2014). Does it really hurt to be responsible? *Journal of Business Ethics, 122*(3), 375–386.

Jo, H., Saha, T., Sharma, R., & Wright, S. (2010). *Socially responsible investing versus vice investing*. Unpublished working paper, Santa Clara University. http://www.aabri.com/LV2010Manuscripts/LV10107.pdf. Accessed 10 June 2014.

Kiymaz, H. (2012). SRI mutual fund and index performance. In H. K. Baker & J. R. Nofsinger (Eds.), *Socially responsible finance and investing: financial institutions, corporations, investors, and activists* (pp. 425–442). Hoboken: Wiley.

Koenker, R. (2005). *Quantile regression*. Cambridge: Cambridge University Press.

Koenker, R., & Basset, G. (1978). Regression quantiles. *Econometrica, 46*(1), 33–50.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives, 15*(4), 143–156.

Koenker, R., & Machado, J. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association, 94*(448), 1296–1310.

Kumar, A., Page, J. K., & Spalt, O. G. (2011). Religious beliefs, gambling attitudes, and financial market outcomes. *Journal of Financial Economics, 102*(3), 671–708.

Lam, S., Jacob, G. H., Yee, A. T. S. (2012). *Socially responsible investment styles: equity risk, return and valuation*. Unpublished working paper, National University of Singapore. http://www.cbern.ca/UserFiles/Servers/Server_625664/Image/CBERN%20Events/PRI2012/Papers/Jacob.pdf. Accessed 01 Jan 2014.

Landier, A., & Nair, V. B. (2009). *Investing for change: Profit from responsible investment*. Oxford: Oxford University Press.

Liston, D. P., & Soydemir, G. (2010). Faith-based and sin portfolios: An empirical inquiry into norm-neglect vs norm-conforming investor behavior. *Managerial Finance, 36*(10), 876–885.

Lobe, S., & Walkshäusl, C. (2011). *Vice versus virtue investing around the world*. Unpublished working paper, University of Regensburg. http://dx.doi.org/10.2139/ssrn.1089827. Accessed 02 Jan 2014.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18*(1), 50–60.

Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). *Does it pay to be good. And does it matter? A meta-analysis of the relationship between corporate social and financial performance*. Unpublished working paper, Harvard University. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1866371. Accessed 21 Oct 2014.

Mata, J., & Machado, J. A. (1996). Firm start-up size: A conditional quantile approach. *European Economic Review, 40*(6), 1305–1323.

McMillen, D. P. (2013). Quantile regression: An overview' and 'linear and nonparametric quantile regression. In J. A. F.

Machado & J. Mata (Eds.), *Quantile regression for spatial data*. New York: Springer.

MSCI (2013). *Business involvement screening research. definitions and scope*. http://help.riskmetrics.com/Screener/Content/Metho dology/Business_Involv_Meth.pdf. Accessed 09 Jan 2014.

Newey, W. K., & West, K. D. (1987). A simple positive-definite heteroskedasticity and autocorrelation-consistent covariance matrix. *Econometrica, 55*(3), 703–708.

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies, 61*(4), 631–653.

Nofsinger, J. R., & Varma, A. (2013). Socially responsible funds and market crises. *Journal of Banking & Finance,*. doi:10.1016/j. jbankfin.2013.12.016.

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

Richardson, B. J. (2013). Socially responsible investing for sustainability: Overcoming Its incomplete and conflicting rationales. *Transnational Environmental Law, 2*(2), 311–338.

Salaber, J. (2009). *Sin stock returns over the business cycle*. Unpublished working paper, University of Bath. http://papers. ssrn.com/sol3/papers.cfm?abstract_id=1443188. Accessed 02 Jan 2014.

Salaber, J. (2013). Religion and returns in Europe. *European Journal of Political Economy, 32*(1), 149–160.

Sandberg, J., Juravle, C., Hedesström, T. M., & Hamilton, I. (2009). The Heterogeneity of socially responsible investment. *Journal of Business Ethics, 87*(4), 519–533.

Scholtens, B. (2014). Indicators of responsible investing. *Ecological Indicators, 36*(1), 382–385.

Scholtens, B., & Sievänen, R. (2013). Drivers of socially responsible investing: A case study of four nordic countries. *Journal of Business Ethics, 115*(3), 605–616.

Schwarz, M. S. (2003). The "Ethics'' of ethical investing. *Journal of Business Ethics, 43*(3), 195–213.

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.

van Fraassen, B. C. (2008). *Scientific representation: Paradoxes of perspective*. Oxford: Oxford University Press.

Visaltanachoti, N., Zou, L., Zheng, Q. (2009). *The performances of sin stocks in China*. Unpublished working paper Massey University. http://www.fma.org/Reno/Papers/sinstock.pdf. Accessed 02 Jan 2014.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83.

Woodbridge, V. (2011). Pornography in your investment portfolio? Ecumenical council for corporate responsibility (ECCR). *Bulletin, 82*(1), 5–11.