ICMA Centre

Henley Business School
UNIVERSITY OF READING

Discussion Paper

Creating More Stable and Diversified Socially Responsible Investment Portfolios

January 2015

Ioannis Oikonomou
ICMA Centre, Henley Business School, University of Reading

Emmanouil Platanakis
ICMA Centre, Henley Business School, University of Reading

Charles Sutcliffe
ICMA Centre, Henley Business School, University of Reading

Discussion Paper Number:
ICM-2015-03

The aim of this discussion paper series is to disseminate new research of academic distinction. Papers are preliminary drafts, circulated to stimulate discussion and critical comment. Henley Business School is triple accredited and home to over 100 academic faculty, who undertake research in a wide range of fields from ethics and finance to international business and marketing.

admin@icmacentre.ac.uk

www.icmacentre.ac.uk

Creating More Stable and Diversified Socially Responsible Investment Portfolios

Abstract

This study is the first to apply a robust estimation technique when constructing Socially Responsible Investing (SRI) portfolios and to highlight that the selection of the optimisation process in this industry matters. We go beyond the mean-variance Markowitz framework in order to bypass issues surrounding the significant estimation risk that causes unstable, poorly diversified and suboptimal portfolios. Using data from MSCI KLD on the social responsibility of US firms, we construct SRI portfolios which exhibit higher risk-adjusted performance, lower total risk, greater stability and diversification compared to conventional and SRI equity indexes, as well as more naïve forms of optimization. Our main conclusions are robust to a series of tests, including the use of different estimation windows, stricter screening criteria, and alternative ways of evaluating portfolio performance.

Keywords

corporate social responsibility, CSR, CSP, sustainability, robust estimation

JEL Classifications

C61, G11, M14

Acknowledgements

We would like to thank Professor Chris Brooks and all the participants of the ICMA Centre weekly academic seminar series for their comments on the paper. All remaining errors are the sole responsibility of the authors.

Contact

Dr Ioannis Oikonomou, ICMA Centre, Henley Business School, University of Reading, PO Box 242 Whiteknights, Reading, RG6 6BA, UK. E-mail: i.oikonomou@icmacentre.ac.uk.

1 Introduction

Corporate Social Responsibility (CSR) and Corporate Social Performance (CSP)[1] have become crucially important concepts in the modern business world. Broadly defined as "a management concept whereby companies integrate social and environmental concerns in their business operations and interactions with their stakeholders"[2], it has gained traction over the past 20 years. A growing number of stakeholders have increased societal demands that corporations perform well financially, while operating in a responsible and ethical manner.

This trend is noticeable in the latest surveys. Grant Thornton's International Business Report[3] in 2014 surveyed 2,500 firms in 34 countries and showed that more and more businesses are being driven to socially and environmentally sustainable practices and initiatives. These range from charitable donations and active participation in local community causes to improving energy efficiency and applying more effective waste management. The majority of these firms cite client/consumer demand as one of the dominant driving forces behind their decision to move towards more sustainable business formats. Similarly, the Nielsen Global Survey on Corporate Social Responsibility (2013) used a sample of 29,000 participants from 58 countries and indicated that at least half of global consumers are willing to "walk the talk" and pay a premium for goods and services produce by socially responsible firms.

In line with these developments, demand for CSP in financial markets, also known as Socially Responsible Investing (SRI)[4], has also been growing rapidly. According to the Global Sustainable Investment Review 2012, which is a product of the collaboration of a variety of organizations and sustainable investment forums across the world, approximately US$13.6 trillion worth of assets under professional management incorporate environmental, social or governance considerations into the investment selection process. This number represents more than 20% of the total assets under professional management in the areas covered in the report, and includes positive and negative screening, shareholder activism strategies, norms-based screening, best-in-class approaches and other forms of SRI. Even if one assumes that the criteria for an investment to be deemed socially responsible are not strict, it is still undeniable that SRI is nowadays a large and expanding segment of the financial markets.

---

[1] The two terms have been used interchangeably in relevant empirical research. In this paper, we use CSP.

[2] United Nations Industrial Development Organization, retrieved October 2014 from http://www.unido.org/en/what-we-do/trade/csr/what-is-csr.html

[3] For additional information, the interested reader is directed at http://www.grant-thornton.co.uk/en/Media-Centre/News/2014/Global-survey-finds-good-CSR-makes-good-business-sense-British-businesses-reacting-to-stakeholders-demands/, retrieved October 2014.

[4] Also referred to as Environmental, Social, and Governance Investing, Sustainable Investing and Impact Investing, though there are some conceptual differences between these terms.

As a result, a significant amount of scholarly research has been dedicated to the investigation of the nature of the relationship between CSP and firm financial performance. Meta-studies focusing on this area (Margolis et al., 2009; Orlitzky et al., 2003) are demonstrative of both its depth and breadth. Using data from hundreds of relevant papers going as far back as 1972, these studies provide evidence of an overall positive link between the two concepts. At the portfolio level of analysis, comparing SRI funds and indices with "conventional" funds and indices of otherwise similar characteristics commonly points to statistically indistinguishable performance (Renneboog et al., 2008; Schroder, 2007; Statman, 2000; Statman, 2006), although there are indications of SRI outperformance in certain contexts (Derwall and Koedijk, 2009; Kempf and Osthoff, 2007).

Despite the size of this literature, a very small number of studies has investigated optimal ways to construct SRI portfolios, either in the sense of the screening criteria used to narrow the investment universe or the optimization process employed. Barnett and Salomon (2006) is one of the few papers that focuses on the effects of screening intensity in SRI funds, and provides evidence of a U-shaped relationship between the number of social/environmental screens used and fund performance. Similarly, there are only a handful of papers (Ballestero et al. 2012; Drut, 2012; Utz et al. 2014) which explore the portfolio optimization frameworks used in SRI. Although such studies significantly contribute to this underdeveloped part of the relevant literature, they are limited in that they do not go far beyond the Markowitz (1952) mean-variance optimisation framework, in terms of allowing for estimation risk. They simply extend it by adding SRI preferences as an additional constraint, or incorporate them in the objective function. This is not ideal as Markowitz optimizations suffer from significant estimation risk (Green and Hollofield, 1992; DeMiguel et al., 2009a), and lead to solutions that are very sensitive to the parametric inputs and to the generation of portfolios that are unstable and poorly diversified. This is unfortunate, as SRI portfolios are characterised by a greater level of uncertainty in these inputs compared to conventional portfolios, due to the inherent complexity in measuring CSP and the largely discretionary nature of CSP reporting.

Our study makes a significant contribution to the literature by using a robust estimation approach which imposes portfolio norm constraints on SRI portfolios. This technique has been shown to have multiple advantages over Markowitz's framework and other similar approaches (DeMiguel et al. 2009a, Xing et al. 2014) including better out-of-sample performance in the presence of estimation risk, lower sensitivity to the input parameters leading to more stable and better diversified portfolios, and lower transaction costs. Using data from MSCI KLD on the social responsibility of US firms, we construct robust SRI portfolios which exhibit higher risk-adjusted performance, lower total risk, greater stability and diversification compared to conventional and SRI equity indexes, and to portfolios constructed using more naïve forms of optimization. Our main conclusions are robust to a series of tests, including the use of different estimation windows, stricter screening criteria and alternative metrics for various aspects of portfolio performance evaluation.

The remainder of the paper is structured as follows: Section 2 reviews previous studies of the application of optimization techniques to forming SRI portfolios and discusses the merits of the robust portfolio approach. Sections 3 contains the details of the CSP database we utilise and describes our optimization model. Section 4 presents our empirical results, and Section 5 concludes.

2 Related literature and motivation of the study

The vast majority of scholarly research dedicated to SRI portfolios focuses on identifying the ways in which they are different from (or similar to) conventional investments in terms of the performance they achieve, the risk they bear and their constituents. A surprisingly small number of academic papers investigate ways in which the portfolio construction process, be it through the use of alternative security selection criteria or different optimization techniques, can lead to the generation of better performing, more efficient and stable SRI portfolios.

Barnett and Salomon (2006) shed some light on the optimal number and type of sustainability related criteria used by SRI funds. Their findings depict a non-linear link between screening intensity and fund/portfolio performance: SRI portfolios where just a few or many social screens are employed outperform portfolios with an intermediate number of such screens. In addition, the authors also investigate the financial contribution of particular types of screening, and they find that community relations screening increases financial performance, whereas environmental and labour relations filters tend to decrease financial performance. Capelle-Blancardand and Monjon (2014) on the other hand, investigate French SRI funds and find that sectoral screens (i.e. avoiding investing in the so-called ‘sin’ stocks) decrease financial performance while other types of CSP screens do not have a noticeable financial impact on fund performance.

Along similar lines, and in an effort to investigate the common claim that SRI funds are in reality nothing more than conventional funds in disguise, Kempf and Osthoff (2008) compare the sustainability characteristics of the portfolio holdings of SRI funds to those of conventional funds. Their investigation focuses on US equity funds and demonstrates that the social and environmental ratings of their constituent stocks are indeed higher than those of otherwise similar conventional funds. Thus any outperformance of these funds can be attributed to the higher CSP levels of the securities they include.

Complimentary to this line of academic research is the small, and fairly new, part of the literature dedicated to the implementation of alternative optimisation approaches in order to construct well-diversified and efficient SRI portfolios. Hallerbach et al. (2004) were the first to point out that the SRI literature was desperately lacking suggestions for combining the social characteristics of risky assets with standard financial information in the portfolio optimization process. They presented an interactive multiple goal programming approach for managing an investment portfolio where the decision criteria include social effects.

An alternative approach was suggested by Drut (2012), who used the classical mean-variance model proposed by Markowitz (1952) and imposed an extra constraint for the CSP rating in order to derive optimal investment strategies, and to investigate whether adding restrictions regarding socially responsible performance leads to portfolios that underperform otherwise similar conventional investments. He concludes that this depends "on the link between the returns and the responsible ratings and on the strength of the constraint" (p. 28). Hence, including additional CSP considerations may not necessarily lead to suboptimal portfolio performance.

On the other hand, Ballestero et al. (2012) used goal programming within the framework of classical Markowitz mean-variance optimization to allow for investors who take into account ethical issues, in addition to the standard financial information. They considered an opportunity set including both "green" and "conventional" assets, and used a two-dimensional objective function (financial and environmental). Their numerical analysis revealed that substantial green investment is generally outperformed by modest green investment – a rare result within the core empirical literature – and hence discourage investors from investing a large part their portfolio in green assets.

The most recent relevant work in the area comes from Utz et al. (2014) who extended the Markowitz model by adding a social responsibility objective, in addition to the portfolio return and variance, causing the traditional efficient frontier to become a surface. When applying their framework to both conventional and SRI mutual funds they did not find any evidence that social responsibility, used as a third criterion and measured by CSP scores, plays an important role in the financial outcome of asset allocation. The authors did, however, find a modestly lower volatility associated with socially responsible compared to conventional funds.

In short, the studies of the optimisation techniques for SRI portfolios tend to focus, not on the effectiveness of the techniques themselves in creating well-performing, stable and diversified portfolios, but rather on providing generic frameworks that integrate financial with social and environmental considerations. They attempt to investigate whether there is a financial cost to including these additional CSP considerations, and whether SRI portfolios tend to outperform or underperform otherwise similar conventional portfolios. Contrary to the above, our work applies a different methodology with the aim of computing SRI portfolios with superior characteristics.

An additional common denominator of previous studies is the use of the Markowitz framework (or extensions of it) in the estimation of SRI portfolios, and this has several important drawbacks. The application of Markowitz mean-variance optimisation requires the estimation of the means, variances and covariances of the asset returns in the investment universe under consideration. The asset weights of the Markowitz portfolio are often very sensitive to the input parameters, i.e. the mean and covariance matrix (Green and Hollofield, 1992; DeMiguel et al., 2009a; and DeMiguel et al., 2009b). In practice this means that, if the sample mean and covariances are subject to estimation error, optimal portfolios constructed via Markowitz optimization are unstable, and characterised by poor diversification and out-of-sample performance. This phenomenon has been well-substantiated by the portfolio selection literature. For instance, Michaud (1999) states that although Markowitz theory provides a convenient framework for portfolio optimization, in practice it is an “error-prone process” that often leads to the construction of portfolios with problematic properties. Broadie (1993) also studied the effect of estimation risk on the construction of the Markowitz efficient frontier, while a more comprehensive review of the influence of estimation errors on portfolio selection can be found in Ziemba and Mulvey (1998).

Furthermore, a strong case can be made that estimation errors in the input parameters constitute a more important issue when constructing SRI portfolios, compared to conventional portfolios. There are several arguments supporting this rationale. First, CSP is a concept which has proved to be very hard to define. Many definitions have been vague or too inclusive. In the words of Votaw (1973): ‘The term is a brilliant one; it means something, but not always the same thing, to everybody’. The work of Carroll (1991) has been influential in defining CSP, and makes reference to a variety of tiers or levels of firm responsibilities (economic, legal, ethical and philanthropic) that taken together constitute CSP. The European Commission on the other hand simply refers to CSP as a concept whereby "companies are taking responsibility for their impact on society"[5].

Second, CSP is characterised by a large amount of variability and heterogeneity in its various dimensions. This makes the accurate measurement of CSP a problematic task (Abbott and Monsen, 1979; Griffin and Mahon, 1997). CSP may be related to, inter alia, issues involving a firm's treatment of the natural environment, employee welfare, philanthropic activity, engagement with local societies and interaction with controversial industries. Subjective judgements are involved, not only in assessing a company's performance in all of the above, but in measuring the relative importance of each CSP dimension for a firm belonging to a particular industry and operating within a specific socio-cultural environment. For example, it could be judged that oil and energy companies should put more emphasis on the environmental aspects of their CSP due to their significant footprint, whereas firms in the financial services sector should be more concerned about product quality and ethical business practices. The quantification of CSP is a complex task which requires the collection and assessment of information both internal and external to the firm by sophisticated, independent assessors such as MSCI, Sustainalytics, Oekom and other agencies producing social ratings for companies.

Third, CSP disclosures remain a discretionary part of corporate reporting in most countries (Orlitzky, 2013). Due to this, voluntary CSP reports are not subject to the same government oversight and regulatory scrutiny as compulsory company reporting. Hence, intentionally erroneous or misleading CSP reporting may not lead to legal and financial sanctions, making such disclosures more susceptible to manipulation brought about by opportunistic firm managers (Edwards, 2008). This further complicates the issue of the accurate measurement of CSP.

Overall, whether it is due to the inherent definitional complexity and heterogeneity of CSP or the strategic misinformation surrounding CSP issues, there is a greater degree of ambiguity when considering CSP as an additional criterion in portfolio creation. Orlitzky (2013) even goes as far as suggesting that CSP may increase the overall level of "noise trading". As we have noted, there is a plethora of studies showing that CSP influences both asset returns (Brammer et al, 2006; Galema et al., 2008; Edmans, 2011; Hillman and Keim, 2001; Von Arx and Ziegler, 2014), and financial risk (Bouslah et al., 2013; Lee and Faff, 2009; Oikonomou et al., 2012). Both qualitative literature reviews (Margolis and Walsh 2003) and statistical meta-analyses (Margolis et al., 2009; Orlitzky et al., 2003) broadly substantiate this conclusion. Thus, it can be surmised that the noise

[5] http://europa.eu/rapid/press-release_MEMO-11-730_en.htm

associated with CSP also leads to additional noise in the pricing of assets included in SRI portfolios (which are screened according to CSP). Therefore SRI portfolios are characterised by a greater degree of estimation errors in the input parameters, i.e. return and risk, and an optimisation method which is less sensitive to these values should be employed. However, the SRI literature is lacking in providing meaningful suggestions.

There are various alternative portfolio optimisation frameworks which could be used for the construction of SRI portfolios with good properties. One such technique requires shrinking the covariance estimator by imposing structural models (e.g. factor models or Bayesian models) on the covariance matrix (Ledoit and Wolf, 2003; Ledoit and Wolf, 2004; Candelon et al., 2012), or by directly shrinking the inverse covariance matrix (Kourtis et al., 2012). A different approach is to generate many datasets, usually through simulation methods, which are then used to compute optimal portfolios. The optimal solution is given by taking the average of these portfolios, Michaud (1999). A third alternative imposes linear constraints on the asset proportions to eliminate the extreme solutions produced by the presence of estimation risk (Board and Sutcliffe, 1994). Finally, a more sophisticated practice requires imposing norm constraints on the portfolio weights, see for instance Ledoit and Wolf (2003), Ledoit and Wolf (2004) and Fan et al. (2008).

We elect to use a technique that falls within this last category, and adopt a robust portfolio technique, inspired by Xing et al. (2014) among others, to construct superior portfolios in the presence of estimation risk – which, as we noted above, is higher when creating SRI portfolios. This approach encourages the creation of sparse portfolios with relatively few active positions and significantly reduced associated transaction costs. It is also particularly well suited to the preferences of the SRI investing community. It has been documented that long-term institutional investment is positively related to corporate social performance (Johnson and Greening, 1999; Cox et al., 2004). This demand arises principally from pension funds and life assurance companies who are investors characterized by high levels of risk aversion and wish to consider worst-case scenarios in order to ensure that their investment decisions are guided by prudence and safety. Insurance companies must comply with the prudential regulations in the UK Financial Conduct Authority handbook, while defined benefit pension schemes must meet their pensions promise, and so both these institutional investors tend to have a low tolerance for risk. In addition, pension schemes often have members who are keen on SRI.

Furthermore, this technique is computationally tractable and time efficient, while alternative portfolio optimization techniques based on generating future scenarios (e.g. stochastic programming or stochastic simulation) are too computationally demanding to be widely applied in practice. Finally, it is more widely comprehensible in comparison with other techniques with the same functionality, and can thus be more easily implemented by practitioners. In the next section we explain the details of our implementation of the robust portfolio estimation approach.

3 Model and dataset

3.1 Model

The mathematical optimization framework proposed by Markowitz (1952) assumes that the expected value $(\boldsymbol{\mu})$ and the covariance $(\boldsymbol{\Sigma})$ of asset returns are known with certainty. Specifically, if $\boldsymbol{\Phi}$ denotes the column vector of portfolio weights (decision variables) defined as $\boldsymbol{\Phi} = [\Phi_1, \Phi_2, ..., \Phi_N]^T$ with N assets in the portfolio, a sample variance-covariance matrix of asset returns ($\boldsymbol{\Sigma}$) and a column vector of mean asset returns ($\boldsymbol{\mu} = [\mu_1, \mu_2, ..., \mu_N]^T$), then the minimum variance portfolio selection problem is expressed as follows:

$$
\begin{aligned}
\underset{\boldsymbol{\Phi}}{\mathit{min}} \quad & \boldsymbol{\Phi}^T \boldsymbol{\Sigma} \boldsymbol{\Phi} \\
\text{s.t.} \quad & \boldsymbol{\mu}^T \boldsymbol{\Phi} \geq \alpha \\
& \mathbf{1}^T \boldsymbol{\Phi} = 1
\end{aligned}
$$

where the objective is the selection of a portfolio $\boldsymbol{\Phi}$ that minimizes the risk (variance) among all feasible portfolios. The constraint $\mathbf{1}^T \boldsymbol{\Phi} = 1$ requires that the portfolio weights sum to one while $\boldsymbol{\mu}^T \boldsymbol{\Phi} \geq \alpha$ imposes a minimum threshold for the portfolio's expected returns.

A major criticism of the Markowitz mean-variance framework is that optimal portfolios are extremely sensitive to the parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, since they are usually estimated from noisy data. To deal with the effects of parameter uncertainty we apply a robust portfolio estimation strategy which is inspired by previous studies of robust asset allocation with norm constraints on the portfolio weights, such as DeMiguel et al. (2009a). Specifically, we follow Xing et al. (2014) and impose a constraint of an l_1 norm, $\|\boldsymbol{\Phi}\|_1$ (taxicab or Manhattan norm) and an l_∞ norm, $\|\boldsymbol{\Phi}\|_\infty$ (maximum norm) on the portfolio weights.

The Markowitz model takes a long or short position in every available asset, some of which can be very large due to estimation errors, making the Markowitz optimal solution impractical in most cases. The taxicab norm (l_1) is the sum of the absolute values of a vector, and setting an upper bound on l_1 discourages short positions and tends to result in many zero positions, with long positions in only a few assets, i.e. sparse portfolios, see for instance Brodie et al. (2008). Having active positions in only a few assets leads to the significant practical benefit of lower transaction costs. However, after applying the l_1 norm constraint, some of these positions may be very large. The additional use of the l_∞ norm achieves this goal, see for instance Brondell and Reich (2008). The maximum norm (l_∞) of a vector is the largest absolute value of the elements in the vector, and an upper bound on l_∞ prevents large long or short positions in any asset. Therefore a combination of the l_1 and l_∞ upper bounds tends to produce sparse portfolios without any very large individual weights.

By setting a lower bound (α) on portfolio expected returns when we minimise the variance of expected returns, the optimization problem we solve can be written as follows:

$$\begin{aligned} \min_{\boldsymbol{\Phi}} \quad & \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Phi} \\ \text{s.t.} \quad & \|\boldsymbol{\Phi}\|_1 + \|\boldsymbol{\Phi}\|_\infty \leq \mathrm{c} \\ & \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Phi} \geq \alpha \\ & \mathbf{1}^{\mathrm{T}}\boldsymbol{\Phi} = 1 \end{aligned} \tag{1}$$

where $\|\boldsymbol{\Phi}\|_1 = \sum_{i=1}^{N} |\Phi_i|$ denotes the l_1 norm, $\|\boldsymbol{\Phi}\|_\infty = \max_{1\leq i\leq N}\{|\Phi_i|\}$ represents the l_∞ norm ($|\Phi_i|$ denotes the absolute value of Φ_i), while $\mathbf{1}$ is a column vector of ones[6]. Furthermore, $\alpha \geq 0$[7] denotes the lower bound[8] on the portfolio expected return, and $\mathrm{c} \geq 1 + \frac{1}{\mathrm{N}}$ is the upper bound of the constraint that involves the l_1 and l_∞ norms. In order to make the most of the l_1 and l_∞

[6] See Appendix A for more technical details of the algorithm used to solve problem (1)

[7] In our analysis, we set the lower bound of the mean portfolio return (parameter α) to 1% on an annual basis. We assume this is the minimum expected return an investor would be willing to accept in order to invest in a risky portfolio. Results are also available for different values of the parameter α. The selection of the exact value of α is not crucial in this framework and does not influence the conclusions drawn from our results.

[8] The lowest feasible value of *c* occurs when it is equal to 1+ 1/N and all the asset weights are equal to *1/N*. Simulations of robust portfolio models usually start with a value of *c* just above 1. Setting *c* some way above 1 permits the optimization process to make a trade-off between preventing short sales and allowing large asset weights.

norms and achieve portfolios that are both diversified and sparse with few massive positions we follow the previous literature (Fan et al., 2008; Fan et al., 2012; Xing et al., 2014) who solve norm-constrained portfolio optimization problems with values of c in the [1, 10] range.

3.2 Dataset

To create SRI portfolios, we use CSP metrics constructed using information in the MSCI ESG STATS database[9]. This dataset is the most frequently used in relevant research, to the extent that it has been characterised as *"the best-researched and most comprehensive"* (Wood and Jones, 1995) in this field, as well as *"the de facto research standard at the moment"* for measuring CSP (Waddock, 2003, p. 369). It is a multi-dimensional CSP database rich in both the cross section of firms analysed (currently about 3,000 US firms) and the timespan covered (23 years) and it has been shown to be characterised by reliability, consistency and construct validity (Sharfman, 1996).

The MSCI ESG STATS data contains annual assessments of the societal and environmental policies and practices of US corporations since 1991. Firms from every sector and industry are assessed on a plethora of indicators relevant to distinct aspects of CSP which are referred to as "qualitative issue areas". These are: community relations, diversity in the workplace, treatment of employees, environmental issues, product (or services) level of safety and quality, corporate governance framework, and respect for human rights. The relevant assessment is done separately on positive aspects ("strengths") and controversial aspects ("concerns") for each qualitative issue area. Sources both internal to the companies (e.g. proxy statements, quarterly reports and other firm documentation) and external to them (e.g. articles in the business and financial press, periodicals, and general media) are used to conduct the assessments of their social performance. In 1991 the dataset covered 650 firms, including all the firms listed in the S&P 500 Composite Index and the Domini 400 Social Index (now the MSCI KLD 400 Social Index). In 2001 this number grew as the relevant universe incorporated the largest 1,000 US companies in terms of market value, an expansion which continued in 2003 with the inclusion of the 3,000 largest US firms. Since 2003 the number of firms in the dataset has remained stable at approximately 3,000.

We follow the relevant empirical work which uses the MSCI ESG STATS database (Hillman and Keim 2001; Oikonomou et al., 2012), and focus solely on those qualitative business issues that can be directly connected with primary stakeholder groups. This is based on the stakeholder theory framework developed by Clarkson (1995) which broadly posits that building strong

[9] Known as KLD STATS before the acquisition of KLD (as part of RiskMetrics) by MSCI in 2010.

collaborative links with those stakeholder groups that are essential to the firm's viability and operational well-being (i.e. the primary stakeholders) are the only ones that will produce tangible financial benefits to the firm. Hence, the CSP measures used to create SRI portfolios are based on those qualitative issue areas which are considered important for applying effective stakeholder management with local communities, employees (including diversity issues), customers and environmental groups/activists (Hillman and Keim, 2001). An outline of the five indicators used in the assessment of each CSP issue area we are interested in can be found in Appendix B.

For the core part of our analysis we construct aggregate measures of CSP for each firm-year observation in the MSCI ESG STATS universe between 1991 and 2012. For each of the five issue areas of interest we sum all the indications for social strengths and deduct the sum of the respective indications for social concerns for a given firm in a given year. Then, we calculate the arithmetic average of all five of these scores in order to create a single, multidimensional CSP rating indicative of the firm's overall social and environmental profile[10]. Our approach follows previous scholarly work in the area of CSP and finance (Jo and Harjoto, 2012 and Deng, Kang and Low, 2013 being two notable examples). Lastly, based on these aggregated CSP scores, we estimate the ranking of each firm across the entire universe covered by MSCI (formerly, KLD) in a given year, and average this relative ranking across the years when the firm is included in the database. We exclude firms for which we cannot construct aggregate scores for at least 10 years out of the 22 in our sample, which helps to ensure the robustness and consistency of the CSP standing of each company. This process results in the estimation of average, aggregate, CSP rankings for 1,362 US firms. We identified the 100 firms with the highest scores as the sub-set of CSP screened firms. This ensures that we have a large enough number of stocks to benefit from the risk reducing effects of diversification when we form portfolios which consist entirely of top CSP performers. We match this dataset with total returns (i.e. returns that include dividends) for these firms and the three indices from Thomson Reuters DataStream.

[10] Creating such a multidimensional CSP measure automatically raises questions about the appropriate way to weight each dimension (i.e. the relative important of each dimension). The common practice in the literature has been to use equal weighting (Deng et al., 2013; Oikonomou et al, 2012) which is what we do. In addition, in subsection 4.3 we look at robust SRI portfolios based on individual CSP dimensions to investigate whether our results can be achieved using each of the five individual CSP measures.

4 Results

4.1 Main results

Due to the smaller coverage of firms by KLD during its earlier stages, as well as missing observations for quite a few firms over that period, it is not feasible to include years prior to 1993 in the data. Tables 1 and 2 depict the details of the estimation and investment periods (in months) we use to evaluate the robust SRI portfolios and the related benchmarks.

Table 1: Six-Year Estimation Periods

Periods(t)	Start	End	Length
Estimation Period 1	1994M1	1999M12	72
Estimation Period 2	1997M1	2002M12	72
Estimation Period 3	2000M1	2005M12	72
Estimation Period 4	2003M1	2008M12	72

Table 2: Non-Overlapping Three-Year Investment Periods

Periods(t)	Start	End	Length
Investment Period 1	2000M1	2002M12	36
Investment Period 2	2003M1	2005M12	36
Investment Period 3	2006M1	2008M12	36
Investment Period 4	2009M1	2011M12	36

In the literature, the length of the estimation period varies, but five to ten years is generally considered to be appropriate. For instance, Xing et al. (2014) use rolling windows of five years (60 months), ten years (120 months) and 15 years (180 months) to evaluate out-of-sample performance; DeMiguel et al. (2009a) use ten years (120 months); DeMiguel et al. (2009b) use ten years (120 months), 30 years (360 months) and 500 years (6000 months) to evaluate the out-of-sample performance of simulated data, while Platanakis and Sutcliffe (2014) use a six year (72 months) rolling window. The choice of a six year estimation window which starts in 1994M1 and ends in 1999M12 for the first estimation period seems to be sensible for the main part of our analysis since it lies within the range used by previous studies, while robustness checks are conducted below with a nine year estimation window. An additional reason for this choice is that the Calvert Social Index, which is used as a benchmark, starts in 2000M1. If an estimation window of less than six years had been used the first investment period would have started before 2000, and we would have been unable to use an important benchmark.

We use data for the first estimation period to compute the optimal portfolio (optimal asset allocation) for the following three years (first investment period). Then, we roll forward the data by 36 months, so that the second estimation period is now used to compute the optimal portfolio for the second investment period and so on, providing a total of four out-of-sample test periods of three years each, or 144 out-of-sample months (12 years) in total.

We run the robust SRI strategy for $c = 4$ (where c is the upper bound of the norm constraint on portfolio weights, see model 1 in section 3). In doing so we follow Fan et al. (2008) and Fan et al. (2012) who run norm-constrained portfolio optimization problems with discrete and continuous values of c in the range [1, 10], as well as Xing et al. (2014) who run norm-constrained optimization problems with values of c in the range [1, 8][11].

Table 3 contains the core of our empirical results and compares the performance of the robust SRI portfolios with five benchmarks. These are the naïve deterministic method (the minimum variance portfolio produced by applying the Markowitz optimization technique to the 100 firms in the SRI sub-sample), and is our predominant benchmark; the 1/N strategy (naive-diversification applied to the SRI sub-sample); the S&P500 Composite index; the MSCI KLD 400 Social Index and the Calvert Social Index. This comparison is made over the 144 out-of-sample months (12 years) which include four investment periods (4×36 months), with different optimal portfolios applying for each three year period (36 months). The results in Table 3 were adjusted to give annualized figures. Comparison of the robust SRI portfolios with the alternative methods allows us to quantify the benefits of the robust estimation process within the SRI sub-sample, while comparisons with a generic US equity index (S&P 500) and SRI equity indexes of the same market (MSCI KLD 400 and Calvert) constitute useful performance benchmarks in order to assess the practical impact of this approach[12].

[11] Robustness checks were carried out with different values for this parameter ($c = 2$, 3 and 5), and do not change the main conclusions of the study. These results are not reported for the sake of parsimony, but are available on request.

[12] Further techniques for deriving optimal portfolio strategies which might have been used as additional benchmarks include: stochastic programming, e.g. Geyer and Ziemba (2008), dynamic programming, e.g. Rudolf and Ziemba (2004), and stochastic simulation, e.g. Boender (1997). However, they are computationally challenging, making them inappropriate for use in practice for the sizeable portfolios we consider. For instance, Platanakis and Sutcliffe (2014) mention that the number of scenarios required by stochastic programming exceeds 24 billion for a portfolio with just 14 assets, four non-overlapping investment periods and five independent outcomes for each uncertain parameter per estimation period.

Table 3

Performance Measures	R-SRI	N-D	1/N	S&P	KLD	CALVERT
Annualized Mean Standard Deviation of Returns	0.1557	2.6026	0.2119	0.1772	0.1771	0.1755
Annualized Mean Risk-Adjusted Returns	0.2530	0.0400	0.0667	0.0491	0.0261	-0.0490
Annualized Sortino Ratio	0.3968	0.0611	0.0953	0.0731	0.0392	-0.0718
Mean Diversification	0.1740	60.2843	0.0100	-	-	-
Mean Stability	0.2300	128.4314	0.0000	-	-	-
VaR(5%)	0.0832	1.3661	0.1216	0.0898	0.0872	0.0888
VaR(10%)	0.0484	0.8988	0.0716	0.0735	0.0783	0.0729
Mean(DD) rate	0.0748	0.8538	0.1160	0.1893	0.2029	0.2113

Comparison of the robust SRI portfolio (R-SRI) comprised of 100 securities with the naïve deterministic portfolio estimation (N-D), the naive-diversification approach (1/N), the S&P 500 Composite Index (S&P), the MSCI KLD 400 Social Index (KLD) and the Calvert Social Index (CALVERT). VaR stands for Value at Risk and Mean(DD) stands for mean drawdown rate.

The portfolios formed in these six different ways are compared using eight criteria which examine risk (four measures), risk-adjusted returns (two measures), diversification, and portfolio stability over time. Natural log returns are used in the estimation of every criterion. The first row of Table 3 shows the annualized average out-of-sample standard deviation of returns. The robust SRI portfolios have the lowest standard deviation at 0.1557 and the naïve deterministic strategy the highest at 2.6026. This is particularly important for the risk-averse, long-term institutional investors who form a significant portion of the demand for SRI. The second row shows the average annualized risk-adjusted returns, which are the asset returns dividend by the standard deviations of returns. The robust SRI portfolios have by far the highest annualized out-of-sample performance at 0.2530, with the second highest being about a quarter of that (0.0667). The relative rankings are very similar when we adjust returns for downside risk (using the lower partial moment of the second order) and compute Sortino ratios.

We also measure the diversification of the portfolios by summing up the squared portfolio weights for each constituent and each estimation period following Blume and Friend (1975):

As a result these techniques are not used as benchmarks in our study due to the computational load they would entail.

$$Diversification_t = \sum_{i=1}^{\mathrm{N}} \Phi_{i,t}^2 \tag{2}$$

where $t = 1, 2, 3, 4$ denotes the estimation period 1, 2, 3 and 4. For naive-diversification (e.g. the 1/N strategy) the score is 1/N, while for zero diversification in a long-only portfolio it is equal to one. Row 4 of Table 3 shows the average results across the four estimation periods. Since we do not have historical data on the relative weights of their constituents, we cannot measure the diversification of the indexes (S&P 500, MSCI KLD 400 Index and Calvert Social Index). The robust SRI portfolios are much more diversified (0.1740) than the naïve deterministic portfolios (60.2843)[13] though, by definition, it cannot beat the naively diversified (1/N) portfolio in this respect.

The portfolio stability between two successive investment periods was measured by summing the squares of the differences between each asset's portfolio weights in adjacent investment periods, see for instance Goldfarb and Iyengar (2003). Mathematically:

$$Stability_{t \to t+1} = \sum_{i=1}^{\mathrm{N}} \left(\Phi_{i,t+1} - \Phi_{i,t} \right)^2 \tag{3}$$

where t denotes the corresponding estimation period. We can view the stability measure as a proxy for transaction costs if we assume that transaction costs are the same across assets and a linear function of the sum invested. The relevant results reveal that robust SRI generates more stable portfolios (0.2300) in comparison with the naïve-deterministic approach (128.4314). Consequently these portfolios are associated with lower transaction costs. Given our fixed sample of firms, the (1/N) strategy has complete stability. As with the measurement of diversification, we cannot make any comparisons with the stability of the benchmark indexes.

The sixth and seventh rows of Table 3 contain the value at risk at the 5% and 10% confidence levels. Irrespective of the confidence interval used, the robust SRI strategy easily outperforms the alternative approaches and indices. This corroborates the results for the standard deviations of the different portfolios, and indicates the lower risk of the robust SRI portfolios.

Lastly, row eight of Table 3 compares the average drawdown rate for the three strategies and the three indexes over the 144 out-of-sample months. Drawdown measures the declines from peaks in cumulative wealth over a specific time horizon. Hence, if a strategy has a lower mean

[13] The N-D diversification measure is greater than one because we have permitted short selling.

drawdown rate over a specific time horizon in comparison with others, it tends to have lower volatility and value-at-risk.

We define the drawdown rate at time t=1,2,…,144 (144 out-of-sample months) as follows (Grossman and Zhou, 1993):

$$DD_t = \frac{\max\limits_{0 \le \tau \le t} \left\{ R_\tau \right\} - R_t}{1 + \max\limits_{0 \le \tau \le t} \left\{ R_\tau \right\}} \tag{4}$$

where the cumulative logarithmic return is defined as:

$$R_t = \sum_{\tau=1}^{t} r_\tau ,\ R_0 = 0 \tag{5}$$

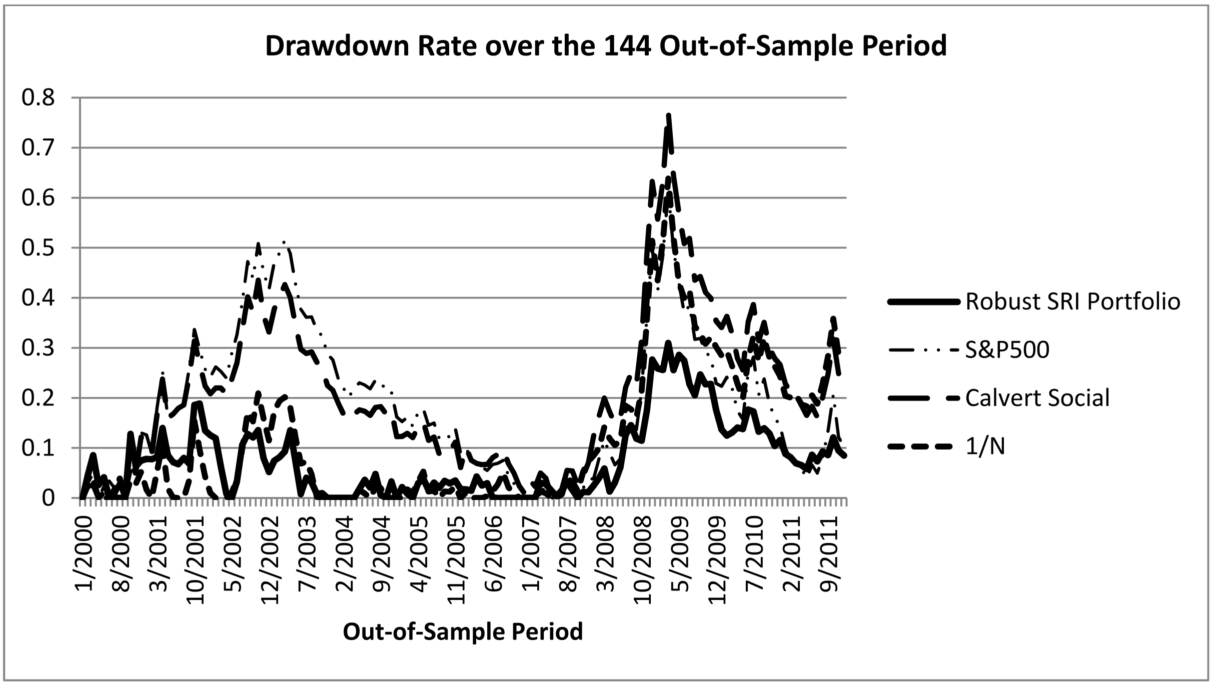
and the cumulative wealth (again, based on logarithmic returns) is:

$$W_t = 1 + \sum_{\tau=1}^{t} r_\tau ,\ \mathrm{W}_0 = 1 \tag{6}$$

Notice that because log returns are used, it is theoretically possible that drawdown rates (as well as value at risk) can be more than 100%. Robust SRI portfolios have the lowest mean drawdown rate (0.0748), while the naïve-deterministic strategy has the highest (0.8538), indicating the lower risk associated with the robust approach. Plotting the drawdown rate of the robust SRI strategy against that of the S&P500, Calvert and 1/N[14] over the entire investment period in Figure 1, the value of robust SRI portfolios is much more stable and less susceptible to market fluctuations.

[14] We do not plot the drawdown for the MSCI KLD 400 index as its performance is so similar to that of the S&P500 (as has also been observed by Statman, 2006) that the lines become nearly indistinguishable. In addition, we do not plot the drawdown rate for the naïve deterministic model as the relevant rate is so much larger compared to alternative strategies and indexes that the scale of the graph changes and it becomes challenging to notice the variability of drawdown rates between different approaches.

Figure 1: Drawdown rate of robust SRI portfolios versus alternative portfolios and indices over the 12 years of the out of sample period.



In Figure 2 we plot the asset allocation of the robust SRI strategy by comparing the average portfolio weights for each security across the four investment periods with the corresponding average for the naïve-deterministic strategy. The naïve-deterministic approach leads to consistently larger positive and negative holdings in numerous assets, and thus high transaction costs, while the robust SRI estimation leads to a much more sparse portfolio with the elimination of "spikes" indicating very large positions in single assets. To be more specific, the robust SRI portfolio has an average weight of 2% or greater in 26 assets, and requires short-selling another 5 assets at 2% or more. The maximum average weight is 9.17% and the minimum average weight is −5.61%. On the other hand, the naïve deterministic approach produces a portfolio which has an average weight of 2% or greater in 42 assets, and requires short-selling another 51 assets with absolute weights of 2% or more. This results in a long-short portfolio with more short positions than long positions, a maximum average weight of 96.69% and an even more unrealistic minimum average weight of −103.17%. There are two main things to note based on these statistics. First, the average robust SRI portfolio has 31 assets with an absolute weight of 2% or more, while the average Markowitz portfolio has 93 positions (out of a possible 100) above the same threshold. Hence, the robust SRI portfolio is more sparse, leading to much lower transaction costs. Second, the robust SRI portfolio has many fewer large negative holdings than the naïve deterministic portfolio, where more than half of the assets are short-sold[15]. Given that the aim is to select a SRI portfolio, it is highly unlikely that "social" investors would be willing to short more than half of this universe, as this would contradict their values.

[15] In allowing short-selling we follow all previous papers that focus on the portfolio optimisation process within the SRI framework (Ballestero et al. 2012; Drut, 2012; Utz et al. 2014).

Figure 2: Comparison of average asset weight between the robust and naïve deterministic approach for portfolios of 100 assets.

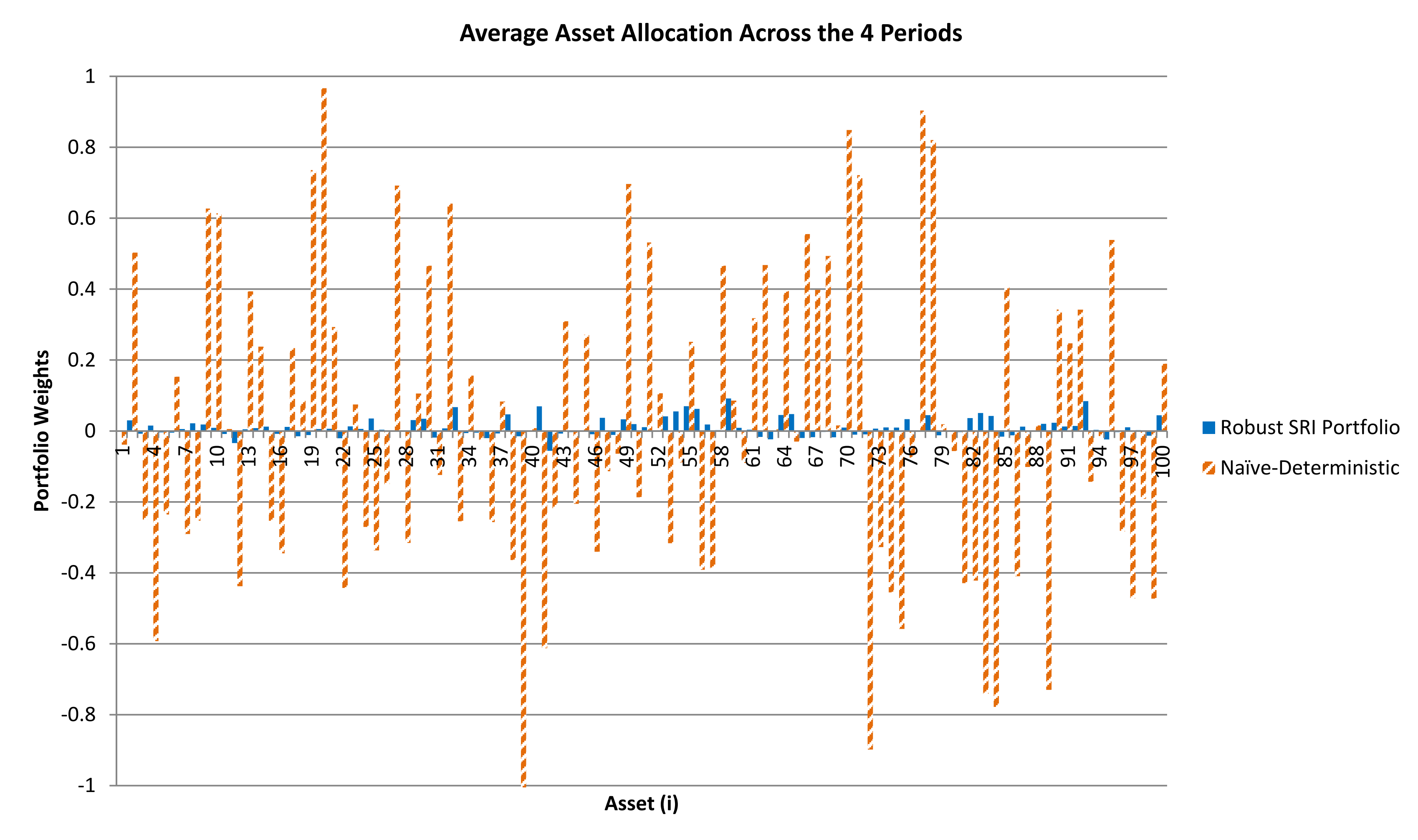
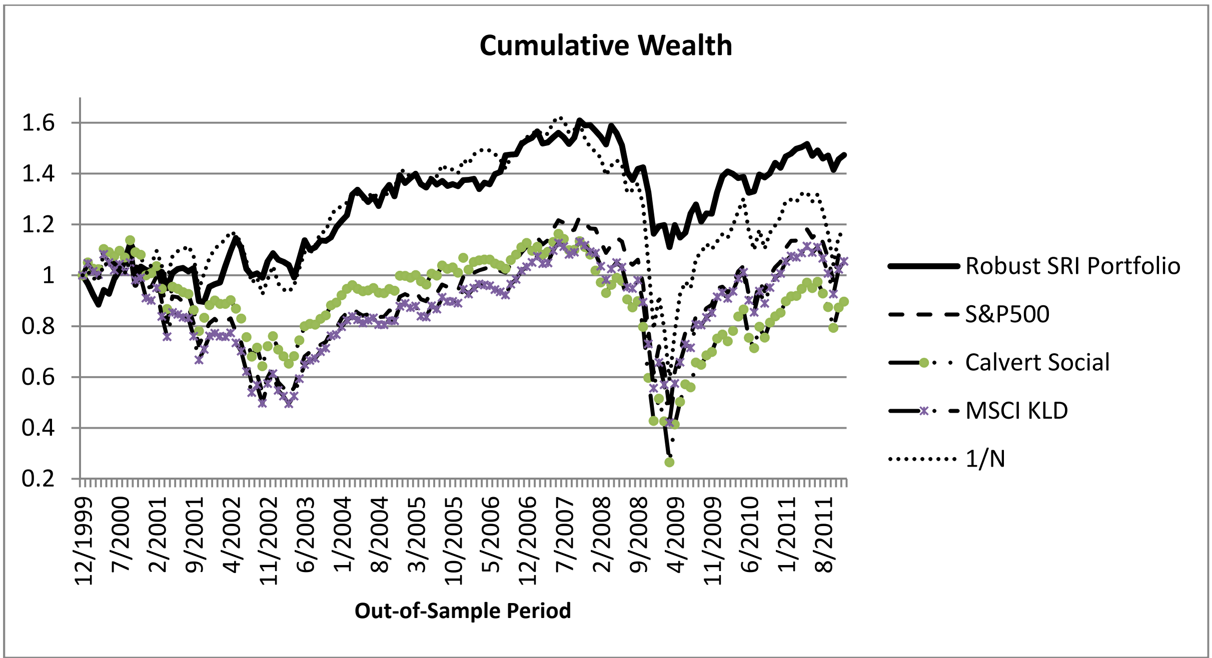


Figure 3 shows the cumulative wealth associated with the robust SRI strategy. We compare the cumulative performance of the robust SRI portfolio with that of the naively diversified portfolio as well as the S&P 500, MSCI KLD 400 and Calvert Social indices to get an indication of the dynamic evolution of the robust portfolio performance[16].

It is clear that the robust SRI investment dominates the naively diversified portfolio, and the S&P 500, MSCI KLD 400 and Calvert Social indices over the span of the four investment periods, which totals 12 years (2000-2011). This is impressive given that our objective function is to minimize portfolio risk, and not maximize portfolio returns. After the first 21 months, the robust SRI portfolio clearly distinguishes its performance vis-à-vis all the indices, and over time gains a significant advantage which it maintains irrespectively of the overall direction of the market. When compared to the naively diversified portfolio, robust SRI produces very similar returns for the majority of the period, but clearly outperforms from around the time of the beginning of the global credit crisis (2007-2008).

Figure 3: Cumulative wealth of the robust SRI portfolios compared to the S&P 500.



[16] We do not plot the dynamic performance of the cumulative wealth of the Markowitz portfolio as it is extremely unstable, with highs and lows of much greater magnitude compared to every other approach and index. Hence, its inclusion in the graph would make every other pattern nearly indistinguishable.

4.2 Robustness tests

To test the robustness of our results, we narrow our investment universe to the top 80[17] firms in terms of aggregate CSP. According to traditional finance theory, restricting the investment universe should lead to inferior portfolio characteristics and an SRI portfolio that may now underperform its benchmarks. On the other hand, given the strong empirical link between higher CSP and lower financial risk (Orlitzky and Benjamin, 2001; Godfrey et al., 2009; Oikonomou et al., 2012), applying more intense CSP screening criteria may improve the performance of robust SRI portfolios. Hence, we decrease the number of equities for the robust SRI portfolio to 80 and use the same estimation and investment periods as in our previous analysis. The results are summarized in Table 4.

Table 4

Performance Measures	R-SRI	N-D	1/N	S&P	KLD	CALVERT
Annualized Mean Standard Deviation of Returns	0.1613	0.9972	0.2135	0.1772	0.1771	0.1755
Annualized Mean Risk-Adjusted Returns	0.2388	0.1797	0.0890	0.0491	0.0261	-0.0490
Annualized Sortino Ratio	0.3711	0.2848	0.1277	0.0731	0.0392	-0.0718
Mean Diversification	0.2236	14.2906	0.0125	-	-	-
Mean Stability	0.2423	29.9809	0.0000	-	-	-
VaR(5%)	0.0765	0.4688	0.1100	0.0898	0.0872	0.0888
VaR(10%)	0.0574	0.3103	0.0729	0.0735	0.0783	0.0729
Mean(DD) rate	0.0841	0.2433	0.1106	0.1893	0.2029	0.2113

Comparison of robust SRI portfolios (R-SRI) comprised of 80 stocks with the naïve deterministic portfolio estimation (N-D), the naive-diversification approach (1/N), the S&P 500 Composite Index (S&P), the MSCI KLD 400 Social Index (KLD) and the Calvert Social Index (CALVERT). VaR stands for Value at Risk and Mean(DD) stands for mean drawdown rate.

The reduction in the number of assets included in the robust SRI portfolio does not change our previous conclusions. The robust SRI approach clearly produces the least risky portfolio as indicated by its lower annualized standard deviations, value at risk, and drawdown compared to the alternative methods and equity indices. In addition, it outperforms the Markowitz portfolio in every respect. When we move from a universe of 100 to 80 assets the risk-adjusted returns of

[17] We also replicated the process for portfolios comprising of the top 120 and 300 stocks in terms of CSP. The results are qualitatively very similar in terms of the outperformance of the robust SRI approach, but are not reported here for the sake of parsimony and are available on request.

the N-D portfolio increase by more than four times, yet this increase is still insufficient to equal the R-SRI risk-adjusted returns.

As a second robustness test, we keep the number of assets at 100 but change the length of the estimation periods to nine years (108 months) instead of six years (72 months). Tables 5 and 6 provide the relevant details of the new estimation and investment periods. We now have only three estimation periods and three investment periods.

Table 5: Nine-Year Estimation Periods

Periods(t)	Start	End	Length
Estimation Period 1	1994M1	2002M12	108
Estimation Period 2	1997M1	2005M12	108
Estimation Period 3	2000M1	2008M12	108

Table 6: Non-Overlapping Three-Year Investment Periods

Periods(t)	Start	End	Length
Investment Period 1	2003M1	2005M12	36
Investment Period 2	2006M1	2008M12	36
Investment Period 3	2009M1	2011M12	36

The results are summarized in Table 7. All our conclusions remain valid as the robust SRI portfolios outperform all the alternative approaches and indices in terms of riskiness, stability, diversification and risk-adjusted returns. Similarly, the Markowitz approach produces the worst results in terms of standard deviation of returns, diversification, stability and value at risk.

Table 7

Performance Measures	R-SRI	N-D	1/N	S&P	KLD	CALVERT
Annualized Mean Standard Deviation of Returns	0.1255	0.3644	0.2237	0.1756	0.1741	0.1751
Annualized Mean Risk-Adjusted Returns	0.6758	0.4894	0.0659	0.2955	0.2822	0.0862
Annualized Sortino Ratio	1.3136	1.1184	0.0923	0.4336	0.4232	0.1223
Mean Diversification	0.1533	1.9029	0.0100	-	-	-
Mean Stability	0.1185	4.1117	0.0000	-	-	-
VaR(5%)	0.0611	0.1413	0.1262	0.0977	0.0872	0.1013
VaR(10%)	0.0404	0.0851	0.0753	0.0497	0.0526	0.0524
Mean(DD) rate	0.0207	0.0631	0.1326	0.0778	0.0701	0.1285

Comparison of robust SRI portfolio (R-SRI) for 100 stocks with the naïve deterministic portfolio estimation (N-D), the naive-diversification approach (1/N), the S&P 500 Composite Index (S&P), the MSCI KLD 400 Social Index (KLD) and the Calvert Social Index (CALVERT). VaR stands for Value at Risk and Mean(DD) stands for mean drawdown rate. Nine year estimation periods are used.

4.3 Additional analyses

It has been documented that different measures of CSP based on different aspects or dimensions of corporate sustainability relate to distinct stakeholder groups (Griffin and Mahon, 1997; Mattingly and Berman, 2006) and this may have different impacts on financial performance. This is especially relevant when looking at samples of firms from different industries where the key performance indicators, as well as the social and environmental issues of note can be significantly different. So far in our analysis we circumnavigated this issue by using an aggregate, multidimensional measure of CSP in order to construct robust SRI portfolios. In this subsection, we create five different robust SRI investment data sets, each based on one of the CSP qualitative issue areas from which the aggregate CSP measure was constructed; i.e. relationships with local communities, diversity in the workplace, employee relations, environmental considerations, and product safety and quality.

To construct these robust SRI portfolios we follow the principles outlined in subsection 4.1. Thus, we use the top 100 firms for each of the qualitative issue areas, and the estimation and investment periods described in Tables 1 and 2. The performance metrics employed, as well as the approaches and indices with which the robust SRI portfolios are compared also remain the same. The results appear in Table 8 which contains five different panels, each of which focuses on one of the five CSP dimensions.

Table 8

Community relations	R-SRI	N-D	1/N	S&P	KLD	CALVERT
Annualized Mean Standard Deviation	0.1782	1.5158	0.2134	0.1772	0.1771	0.1755
Annualized Mean Risk-Adjusted Returns	0.0875	0.0294	0.1542	0.0491	0.0261	-0.0490
Annualized Sortino Ratio	0.1341	0.0468	0.2182	0.0731	0.0392	-0.0718
Mean Diversification	0.1797	44.0440	0.0100	-	-	-
Mean Stability	0.2229	85.3885	0.0000	-	-	-
VaR(5%)	0.0697	0.7465	0.1165	0.0898	0.0872	0.0888
VaR(10%)	0.0581	0.5243	0.0715	0.0735	0.0783	0.0729
Mean(DD) rate	0.1023	0.7990	0.0989	0.1893	0.2029	0.2113
Diversity	**R-SRI**	**N-D**	**1/N**	**S&P**	**KLD**	**CALVERT**
Annualized Mean Standard Deviation	0.1678	2.0359	0.2097	0.1772	0.1771	0.1755
Annualized Mean Risk-Adjusted Returns	-0.0463	-0.0972	-0.0074	0.0491	0.0261	-0.0490
Annualized Sortino Ratio	-0.0695	-0.1424	-0.0104	0.0731	0.0392	-0.0718
Mean Diversification	0.1752	41.1643	0.0100	-	-	-
Mean Stability	0.2250	84.0139	0.0000	-	-	-
VaR(5%)	0.0878	1.2635	0.1019	0.0898	0.0872	0.0888
VaR(10%)	0.0697	0.6433	0.0745	0.0735	0.0783	0.0729
Mean(DD) rate	0.1674	1.8963	0.1423	0.1893	0.2029	0.2113
Employee relations	**R-SRI**	**N-D**	**1/N**	**S&P**	**KLD**	**CALVERT**
Annualized Mean Standard Deviation	0.1654	2.4512	0.2031	0.1772	0.1771	0.1755
Annualized Mean Risk-Adjusted Returns	0.2030	-0.2104	0.1297	0.0491	0.0261	-0.0490
Annualized Sortino Ratio	0.2938	-0.2780	0.1812	0.0731	0.0392	-0.0718
Mean Diversification	0.1752	53.3399	0.0100	-	-	-
Mean Stability	0.2186	107.6395	0.0000	-	-	-
VaR(5%)	0.0686	1.0118	0.1220	0.0898	0.0872	0.0888
VaR(10%)	0.0554	0.7435	0.0656	0.0735	0.0783	0.0729
Mean(DD) rate	0.0897	4.2362	0.1070	0.1893	0.2029	0.2113

Environment	R-SRI	N-D	1/N	S&P	KLD	CALVERT
Annualized Mean Standard Deviation	0.1540	2.2737	0.2190	0.1772	0.1771	0.1755
Annualized Mean Risk-Adjusted Returns	-0.0440	-0.4374	0.0592	0.0491	0.0261	-0.0490
Annualized Sortino Ratio	-0.0647	-0.6043	0.0858	0.0731	0.0392	-0.0718
Mean Diversification	0.1982	57.3060	0.0100	-	-	-
Mean Stability	0.2559	120.9026	0.0000	-	-	-
VaR(5%)	0.0975	1.1227	0.1212	0.0898	0.0872	0.0888
VaR(10%)	0.0697	0.8543	0.0805	0.0735	0.0783	0.0729
Mean(DD) rate	0.1461	1.7096	0.1272	0.1893	0.2029	0.2113
Product safety and quality	**R-SRI**	**N-D**	**1/N**	**S&P**	**KLD**	**CALVERT**
Annualized Mean Standard Deviation	0.1635	2.8024	0.2449	0.1772	0.1771	0.1755
Annualized Mean Risk-Adjusted Returns	0.2735	0.4888	-0.0224	0.0491	0.0261	-0.0490
Annualized Sortino Ratio	0.4230	0.9408	-0.0316	0.0731	0.0392	-0.0718
Mean Diversification	0.1897	77.4157	0.0100	-	-	-
Mean Stability	0.2140	159.0189	0.0000	-	-	-
VaR(5%)	0.0775	0.9723	0.1311	0.0898	0.0872	0.0888
VaR(10%)	0.0534	0.8652	0.0731	0.0735	0.0783	0.0729
Mean(DD) rate	0.0986	0.1875	0.1620	0.1893	0.2029	0.2113

Comparison of the robust SRI portfolios (R-SRI) with 100 stocks constructed using the five different dimensions of CSP, with the naïve deterministic portfolio estimation (N-D), the naive diversification approach (1/N), the S&P 500 Composite Index (S&P), the MSCI KLD 400 Social Index (KLD) and the Calvert Social Index (CALVERT). VaR stands for Value at Risk and Mean(DD) stands for mean drawdown rate.

Table 8 reveals that, although there is variability, the robust SRI portfolios generally tend to outperform the alternatively constructed portfolios and indexes. Across all five CSP dimensions R-SRI has one of the lowest risk measures (standard deviation, value at risk and drawdown). R-SRI is superior to N-D on all the performance measures for the dimensions of community relations, diversity issues, employee relations, and environment; and superior on all measures except the risk-adjusted return and the Sortino ratio for the dimension of product safety and quality. For employee relations and product safety and quality, R-SRI is superior to the S&P500, KLD and Calvert indices on all the available performance measures, and for community relations it only underperforms the indices in terms of the standard deviation measure of risk. Relative to the 1/N approach, R-SRI is superior on every dimension for the performance measures of standard deviation and value at risk, and superior on risk-adjusted returns and the Sortino ratio for the dimensions of employee relations and product safety and quality. Overall, Table 8 shows that the results from the solo use of the other CSP measures produce slightly different conclusions. This suggests that, while the results are not highly sensitive to the weighting scheme involved in computing the aggregate CSP, it is beneficial to use the aggregate CSP measure when constructing robust SRI portfolios. The distinctiveness of CSP dimensions and the variability of the financial impacts of each has been well documented in the empirical CSP literature (Hillman and Keim, 2001; Mattingly and Berman, 2006; Oikonomou et al., 2012). Hence, these results are compatible with previous findings.

5 Conclusions

We expand the SRI literature by moving beyond the question of whether portfolios comprising "sustainable equities" outperform conventional investments, and focus on finding optimal ways to construct SRI portfolios. We show that the selection of the optimisation process for SRI portfolios matters and that, given the noise surrounding CSP measurement, it is essential to apply a method which is less sensitive to the input parameters. In particular, we apply a robust portfolio construction approach in order to alleviate concerns arising from the estimation risk commonly associated with the traditional Markowitz framework, as well as other optimization models.

Our out-of-sample results show that, when using a multidimensional CSP criterion to construct SRI portfolios, the robust approach outperforms the Markowitz (naïve deterministic) model and the naively diversified portfolio, as well as three equity indexes, both generic (S&P 500) and social (MSCI KLD 400 and Calvert). Specifically, the robust SRI portfolios are characterised by lower total risk, value at risk and drawdown, greater stability of asset weights between rebalancing periods, greater diversification and higher risk-adjusted returns. These findings are shown to be robust to different lengths of the estimation and investment periods, and to the use of more stringent CSP screening criteria. In addition, when implementing the robust estimation approach using just one CSP dimension, we find that, in most cases, they are still less risky than the five benchmarks, and generally support the superiority of robust SRI portfolios.

Overall, our findings demonstrate that the robust estimation approach is beneficial for the SRI industry. Fund managers, institutional and retail investors (especially those who are more risk-averse and have longer-term investment horizons) will significantly benefit from applying this technique, instead of a more naïve approach, after they have implemented the social and environmental screens of choice to the investment universe of interest. Although our study is innovative within the SRI field, it is limited by considering only one asset class (equities) and the geographic coverage of the markets considered (US). Future studies can extend our analysis in either of these directions. Furthermore, alternative estimation techniques with different advantages and drawbacks from the robust approach can also be considered for the purpose of building SRI portfolios. Lastly, the selection of the CSP criteria and dataset used is always an important issue within the literature (Griffin and Mahon, 1997). Using different social and environmental sources of data and alternative CSP metrics would provide a useful test for the reliability of different optimization methods for SRI.

References

Abbott, W. F. and Monsen R. J. (1979). On the measurement of corporate social responsibility: Self-reported disclosures as a method of measuring corporate social involvement. *Academy of Management Journal*: vol. 22, pp. 501-515.

Ballestero, E., Bravo, M., Pérez-Gladish, B., Arenas-Parra, M., and Plà-Santamaria, D. (2012) Socially Responsible Investment: A Multicriteria Approach to Portfolio Selection Combining Ethical and Financial Objectives. *European Journal of Operational Research*, vol. 216, pp. 487–494.

Barnett, M. L. and Salomon, R. M. (2006). Beyond dichotomy: The Curvilinear Relationship between Social Responsibility and Financial Performance, *Strategic Management Journal* vol. 27, pp. 1101-1122.

Blume, M.E. and Friend, I. (1975) The Asset Structure of Individual Portfolios and Some Implications for Utility Functions, *Journal of Finance*, vol. 30, pp. 585-603.

Board, J.L.G. and Sutcliffe, C.M.S. (1994) Estimation Methods in Portfolio Selection and the Effectiveness of Short Sales Restrictions: UK Evidence, *Management Science*, Vol. 40, No. 4, April, pp. 516-534.

Boender, C.G.E. (1997) A Hybrid Simulation-Optimization Scenario Model for Asset-Liability Management, *European Journal of Operations Research*, vol. 99, pp. 126-135.

Bondell, H.D. and Reich, B.J. (2008) Simultaneous regression shrinkage, variable selection, and supervised clustering of predictors with OSCAR, *Biometrics*, vol. 64, pp. 115-123.

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

Broadie, M. (1993) Computing Efficient Frontiers Using Estimated Parameters, *Annals of Operations Research*, vol. 45, pp. 21-58.

Brodie, J.I., Daubechies, C.D., Giannone, M.D. and Loris, I. (2008) Sparse and Stable Markowitz Portfolios, *European Central Bank Working Paper Series*, no. 936.

Candelon, B., Hurlin, C. and Tokpavi, S. (2012) Sampling Error and Double Shrinkage Estimation of Minimum Variance Portfolios, *Journal of Empirical Finance*, vol. 19, September, pp. 511–527.

Capelle-Blancard, G. and Monjon, S. (2014), The Performance of Socially Responsible Funds: Does the Screening Process Matter?. *European Financial Management*, vol. 20, pp. 494–520.

Carroll, A. B. (1991). The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders. *Business Horizons* vol. 34, pp. 39-48.

Clarkson, M. B. E. (1995). A stakeholder framework for analyzing and evaluating corporate social performance. *Academy of Management Review,* vol. 20, pp. 92-117.

Cox, P., Brammer, S. and Millington, A. (2004). An empirical examination of institutional investor preferences for corporate social performance. *Journal of Business Ethics*, vol. 52, pp. 27-43.

Deng, X., Kang, J. K. & Low, B. S. (2013). Corporate social responsibility and stakeholder value maximization: Evidence from mergers. *Journal of Financial Economics*, vol. 110, pp. 87-109.

DeMiguel, V., Garlappi, L., Nogales, F.J. and Uppal, R. (2009a) A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms, *Management Science*, vol. 22, no. 5, pp. 1915–1953.

DeMiguel, V., Garlappi, L. and Uppal, R. (2009b) Optimal Versus Naive diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953.

Derwall, J. and K. Koedijk (2009). Socially Responsible Fixed Income Funds. *Journal of Business Finance & Accounting*, vol. 36, pp. 210-229.

Drut, B. (2012) Social Responsibility and Mean-Variance Portfolio Selection, *Amundi Working Paper*.

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

Edwards, M. (2008). Small change: Why business won't save the world. San Francisco: Berrett-Koehler.

Fan, J., Zhang, J. and Yu, K. (2008) Asset allocation and risk assessment with gross exposure constraints for vast portfolios, *The Annals of Statistics*, vol. 25, pp. 1425–1432.

Fan, J., Zhang, J. and Yu, K. (2012) Vast Portfolio Selection with Gross-Exposure Constraints, *Journal of the American Statistical Association*, vol. 107, pp. 592-606.

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

Geyer, A. and Ziemba, W.T. (2008) The Innovest Austrian Pension Fund Financial Planning Model InnoALM, *Operations Research*, vol. 56, pp. 797-810.

Godfrey, P. C., Merrill, C. B. and Hansen, J.M. (2009). The relationship between corporate social responsibility and shareholder value: an empirical test of the risk management hypothesis. *Strategic Management Journal*, vol. 30, pp. 425-445.

Goldfarb, D. and Iyengar, G. (2003) Robust Portfolio Selection Problems, *Mathematics of Operations Research*, vol. 28, no. 1, pp. 1-38.

Green, R. and Hollifield, B. (1992) When will Mean–Variance Efficient Portfolios be well

Diversified?, *Journal of Finance*, vol. 47, no. 5, pp. 1785–1809.

Griffin, J. J. and. Mahon, J. F (1997). The corporate social performance and corporate financial performance debate. *Business & Society*, vol. 36, pp. 5-31.

Grossman, S. and Zhou, Z. (1993) Optimal investment strategies for controlling drawdowns, *Mathematical Finance*, vol. 3, pp. 241–276.

Hallerbach, W., Ning, H., Soppe, A., and Spronk, J. (2004) A framework for managing a portfolio of socially responsible investments. *European Journal of Operational Research*, vol. 153, pp. 517–529.

Jo, H., and M. Harjoto. (2012). The causal effect of corporate governance on corporate social responsibility. *Journal of Business Ethics*, vol. 106, pp. 53-72.

Hillman, A. J. and G. D. Keim (2001). Shareholder value, stakeholder management, and social issues: what's the bottom line? *Strategic Management Journal*, vol. 22, pp. 125-139.

Johnson, R. D. and D. W. Greening (1999), The Effects of Corporate Governance and Institutional Ownership Types on Corporate Social Performance, *Academy of Management Journal,* vol. 42, pp. 564-576.

Bouslah, K., Kryzanowski, L. amd M'Zali, B. (2013), The impact of the dimensions of social performance on firm risk, *Journal of Banking & Finance*, vol. 37, pp. 1258-1273.

Kempf, A. and P. Osthoff (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, vol. 13, pp. 908-922.

Kempf, A. and Osthoff, P. (2008), SRI Funds: Nomen est Omen. *Journal of Business Finance & Accounting*, vol. 35, pp. 1276-1294.

Kourtis, A., Dotsis, G. and Markellos, R.N. (2012) Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix, *Journal of Banking and Finance*, vol. 36, pp. 2522-2531.

Ledoit, O. and Wolf, M. (2003) Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection, *Journal of Empirical Finance*, vol. 10, December, pp. 603–621.

Ledoit, O., Wolf, M. (2004) A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis*, vol. 88, February, pp. 365–411.

Lee, D. D. and Faff, R. W. (2009), Corporate Sustainability Performance and Idiosyncratic Risk: A Global Perspective. *Financial Review*, vol. 44, pp. 213–237.

Margolis, J. D. and J. P. Walsh (2003). Misery loves companies: Rethinking social initiatives by business. *Administrative Science Quarterly,* vol. 48, pp. 268-305.

Margolis, J., Elfenbein, H., & 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. Harvard Business School Working Paper No. 229.

Markowitz, H. (1952) Portfolio Selection, *Journal of Finance*, vol. 7, no. 1, pp. 77-91.

Mattingly, J. E. and Berman, S. L. (2006). Measurement of corporate social action. *Business & Society*, vol. 45, pp. 20-46.

Michaud, R. (1999) *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*, Harvard Business School Press.

Oikonomou, I., Brooks, C., & Pavelin, S. (2012). The impact of corporate social performance on financial risk and utility: A longitudinal analysis. *Financial Management*, vol. 41, pp. 483–515.

Orlitzky, M. and J. D. Benjamin (2001). Corporate social performance and firm risk: A meta-analytic review. *Business & Society*, vol. 40, pp. 369-396.

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

Orlitzky, M. (2013). Corporate social responsibility, noise and stock market volatility, *The Academy of Management Perspectives*, vol. 27, pp. 238-254.

Platanakis, E and Sutcliffe, C. (2014) Asset Liability Modelling and Pension Schemes: The Application of Robust Optimization to USS, ICMA Centre Discussion Paper & Henley Business School Discussion Paper

Renneboog, L., J. Ter Horst, and C. Zhang (2008). The Price of Ethics and Stakeholder Governance: The Performance of Socially Responsible Mutual Funds, *Journal of Corporate Finance* 14, 302-322.

Rudolf, M. and Ziemba, W.T. (2004) Intertemporal Surplus Management, *Journal of Economic Dynamics and Control*, vol. 28, pp. 975 990.

Schroder, M. (2007). Is there a difference? The performance characteristics of SRI equity indices. *Journal of Business Finance and Accounting*, vol. 34, pp. 331–348.

Sharfman, M. (1996). The construct validity of the Kinder, Lydenberg & Domini social performance ratings data. *Journal of Business Ethics*, vol. 15, pp. 287-296.

Statman, M. (2000). Socially responsible mutual funds. *Financial Analysts Journal,* vol. 56, pp.: 30-39.

Statman, M. (2006). Socially responsible indexes. *Journal of portfolio Management,* vol. 32, pp.: 100-109.

Utz, S., Wimmer, M., Hirschberger, M. and Steuer, R.E. (2014) Tri-Criterion Inverse Portfolio Optimization with Application to Socially Responsible Mutual Funds, *European Journal of Operational Research*, vol. 234, pp. 491-498.

Von Arx, U., & Ziegler, A. (2014). The effect of corporate social responsibility on stock performance: New evidence for the USA and Europe. *Quantitative Finance*, vol. 14, pp. 977-991.

Votaw, D. (1973). Genius becomes rare. The Corporate Dilemma-Traditional Values versus Contemporary Problems. Prentice Hall, Englewood Clifs, NJ.

Waddock, S.A. (2003). Myths and Realities of Social Investing, *Organization and Environment,* vol. 16, pp. 369-380.

Wood, D. J. and Jones, R. E. (1995). Stakeholder mismatching: A theoretical problem in empirical research on corporate social performance. *International Journal of Organizational Analysis* vol. 3, pp. 229-267.

Xing, X., Hu, J. and Yang, Y. (2014) Robust Minimum Variance with L-Infinity Constraints, Journal of Banking and Finance, vol. 46, pp. 107-117.

Ziemba, W.T. and Mulvey, J.M. (1998) *Worldwide Asset and Liability Modeling*, Cambridge University Press.

Appendix A

Assuming that $\left|\Phi_N\right| = \max\limits_{1\le i\le N}\left\{\left|\Phi_i\right|\right\}$, and writing the portfolio weights and their absolute values as $\Phi_i = \Phi_i^+ - \Phi_i^-$ and $\left|\Phi_i\right| = \Phi_i^+ + \Phi_i^-$ (see also Appendix B in the paper of Xing et al. (2014) for details concerning this transformation), problem (1) in section 3 is written as follows:

$$
\begin{array}{lll}
\min\limits_{\boldsymbol{\Phi}^+,\boldsymbol{\Phi}^-} & \boldsymbol{\Phi}^{T}\boldsymbol{\Sigma}\boldsymbol{\Phi} & \\
\text{s.t.} & \left|\Phi_i\right| \le \left|\Phi_N\right|, & \forall i=1,\dots,N-1 \\
 & \sum\limits_{i=1}^{N}\left(\Phi_i^+ + \Phi_i^-\right) + \left(\Phi_N^+ + \Phi_N^+\right) \le c & \\
 & \sum\limits_{i=1}^{N}\left(\Phi_i^+ - \Phi_i^-\right) = 1 & \\
 & \sum\limits_{i=1}^{N}\left(\Phi_i^+ - \Phi_i^-\right)\mu_i \ge \alpha & \\
 & \Phi_i^+ \ge 0, & \forall i=1,\dots,N \\
 & \Phi_i^- \ge 0, & \forall i=1,\dots,N
\end{array}
$$

where there are N possible maxima of the portfolio weights. Hence, the above problem has $3N+2$ linear constraints.

Appendix B

MSCI KLD Qualitative issue areas of interest	Strengths	Concerns
Community	- Charitable Giving	- Investment Controversies
	- Innovative Giving	- Negative Economic Impact
	- Non-US Charitable Giving	- Indigenous Peoples Relations
	- Support for Housing	- Tax Disputes
	- Support for Education	- Other Concern
	- Indigenous Peoples Relations	
	- Volunteer Programs	
	- Other Strength	
Diversity	- CEO's identity	- Controversies
	- Promotion	- Non-Representation
	- Board of Directors	- Other Concern
	- Work/Life Benefits	
	- Women & Minority Contracting	
	- Employment of the Disabled	
	- Gay & Lesbian Policies	
	- Other Strength	
Employee Relations	- Union Relations	- Union Relations
	- No-Layoff Policy	- Health and Safety Concern
	- Cash Profit Sharing	- Workforce Reductions
	- Employee Involvement	- Retirement Benefits Concern
	- Retirement Benefits Strength	- Other Concern
	- Health and Safety Strength	
	- Other Strength	
Environment	- Beneficial Products and Services	- Hazardous waste
	- Pollution Prevention	- Regulatory Problems
	- Recycling	- Ozone Depleting Chemicals
	- Clean Energy	- Substantial Emissions
	- Communications	- Agricultural Chemicals
	- Property, Plant, and Equipment	- Climate Change
	- Management Systems	- Other Concern
	- Other Strength	
Product Safety and Quality	- Quality	- Product Safety
	- R&D/Innovation	- Marketing/Contracting Concern
	- Benefits to Economically Disadvantaged	- Antitrust
	- Other Strength	- Other Concern