



Performance Analysis of an SRI Screen Based on Employment
Quality in High and Low Human Capital Industries
- International Evidence

NOVA SBE / Maastricht University
School of Business and Economics
Lisbon, 30th December 2014
da Costa Avelar, M.
ID: 828 / i6000558
Supervisors: Prado, M. / Derwall, J.
Writing Assignment: Master Thesis

Table of Contents

1. Abstract	3
2. List of Figures.....	4
3. List of Tables	5
4. Introduction.....	6
5. Literature Review.....	8
The Basis for this Study	8
Employee Satisfaction and Firm Value	9
Traditional View on Motivation.....	14
Mispricing of Intangible Assets.....	15
Other SRI Screens and their Effect on Portfolio Performance	18
6. Data and Summary Statistics.....	23
Division of Companies.....	23
Scoring.....	27
Descriptive Statistics.....	29
7. Methodology	30
8. Results.....	33
9. Discussion and Implications of Results	38
10. Conclusion.....	42
Appendix.....	44
References.....	47

1. Abstract

This research focuses on the performance analysis of a positive socially responsible investment (SRI) screen based on employment quality. The screen is applied to two different industry portfolios, of which one is human capital intense and the other is less human capital intense. A scoring is created in order to assess the employment quality in each company of the sample and to form best-in-class portfolios. The database used to create the scoring is Thomson Reuters Asset4 ESG. The analysis is conducted for three distinct regions: the US, Europe and Japan. The sample period for the US is from January 2002 to December 2012, for Europe from January 2004 to December 2012 and for Japan from April 2003 to March 2013. Monthly risk-adjusted returns using the Carhart (1997) four-factor model result in mostly negative and insignificant alphas. The results show that this type of SRI screen does not seem to enhance portfolio performance, independent of certain weighting methods for the portfolios (equal- and value-weighted) or benchmarks (risk-free rate and industry returns) used. Overall, the findings are in line with a learning effect of investors and stock markets in incorporating employment related information in their stock evaluations. Abnormal returns from the usage of this kind of SRI screen seem to have vanished as a result of this learning in the recent past.

2. List of Figures

Figure 1: Impact of employee satisfaction on long-run stock returns within the company (upper scheme) and in the stock market (lower scheme).....	9
--	---

3. List of Tables

Table 1: Division of Industries	26
Table 2: Risk-adjusted returns of equal-weighted best-in-class portfolios (risk-free rate).....	33
Table 3: Risk-adjusted returns of value-weighted best-in-class portfolios (risk-free rate).....	35
Table 4: Risk-adjusted returns of equal-weighted best-in-class portfolios (industry return).....	36
Table 5: Risk-adjusted returns of value-weighted best-in-class portfolios (industry return).....	37
Table 6: Alpha coefficients.....	42
Table 7: ASSET4 ESG data structure	44
Table 8: Scoring components.....	44
Table 9: Descriptive statistics of industries per region.....	45
Table 10: Weighting of payroll data for average payroll calculation per company	46

4. Introduction

Socially responsible investing (SRI) is a promising idea, as it gives the opportunity to earn higher returns on investments while at the same time making ethically and socially acceptable investments. Apart from the financial gains an investor might also gain non-financial utility from these sorts of investments, by doing something “good”. It is therefore not surprising that it is given more attention in the last couple of years and is also applied by more funds in the market. From 1995 until 2012, the number of funds that incorporated environmental, social and corporate governance (ESG) factors in their investment selection process in the US increased from 55 to 720, while the total net assets increased from \$12 billion to \$1,013 billion (SIF, 2012, p.11). Other global pressures, like global warming and overall resource depletion give the topic further importance and SRI provides investors with an alternative to conventional investment strategies. Many studies have put emphasis on this subject, analyzing different sorts of screens in different geographical areas or analyzed, for instance, SRI fund performances. The empirical results seem to differ quite significantly and seem to depend on the applied screens, regions and sample periods on which they are applied to.

This study focuses on analyzing the performance of a single positive SRI screen in form of employment quality, applied to two different types of industry, namely human capital intense industries and less human capital intense industries. One of the main hypotheses of this study is that the screen should work better and result in statistically significant positive four-factor alphas when applied to human capital intense industries. Positive alphas are also expected for the application of the SRI screen on the less human capital intense industry. The results should be, however, economically smaller and statistically less significant compared to the high human capital application. As human capital might be valued higher in industries where it is most needed, competing for valuable employees and getting better employees could result in a competitive advantage for a firm in those industries. Getting “good” employees in an industry, where, for instance, innovation plays an important role, as for a company like Apple, certainly might be more important than for Walmart.

This analysis is done for three distinct regions: the US, Japan and continental Europe (consisting of the countries in the European monetary union of 1999). The portfolios are constructed in a best-in-class way. The Asset4 ESG database from Thomson Reuters is used to create a scoring for each company on a monthly basis. For this scoring, 10 factors are

considered, where each factor can be seen as a question similar to: “Does the company claim to provide a bonus plan to most employees?”. The 10 questions cover three different themes that should be important in the determination of employee satisfaction: employment quality, health & safety and training & development. Two of the 10 questions are of numerical nature, while the others are qualitative Yes/No questions. The 20 companies with the highest score form a portfolio. This portfolio construction is repeated for high and low human capital intensity industries. In order to determine whether a specific industry is of high human capital intensity, an income-/cost-based approach is used. Data from the U.S. census bureau on the payrolls per employee in each industry in the US from 2011 is used to identify the human capital intensity in each industry of the Thomson Reuters universe. The higher the payroll is, the more likely it is that human capital has a bigger importance, since it is valued higher. Following the logic from above, both equally- and value-weighted portfolios are constructed on a monthly basis and compared to each other. The commonly used Carhart (1997) four-factor model, consisting of the Fama and French three-factor model and a momentum factor, is used in order to assess the performance of each portfolio.

This study extends general studies on SRI screen performances. By adding an industry perspective, especially the prior work of Edmans (2011) is expanded. This study hypothesizes that certain SRI screens might work better in industries that are more related to the SRI screen than others. Furthermore, a different database is used, which might be considered less publicly available as the approach picked by Edmans. This will show whether similar results are obtained although a different database is used to assess the employment quality in the companies. Theoretically this might result in a better performance of the strategy, since investors might fail to incorporate the less publicly available information in their stock evaluations even more.

The results, however, show mostly negative and insignificant alphas suggesting that the SRI screen does not improve portfolio performance. Neither the weighting methods for the portfolios (equal- and value-weighted) nor the benchmarks (risk-free rate and industry returns) used, seem to affect these results. The findings are in line with a learning effect theory stating that investors might learn from their mispricing behavior in the past (Derwall et al., 2011). As a result of this learning, investors incorporate information about employment quality better in their stock evaluations. Ultimately this makes abnormal returns associated with this type of SRI screen disappear.

The paper is structured as follows. Section 5 gives a general overview of the existent literature on the topic of SRI and human capital. The section is followed by a description of the data, division into the industry types and the scoring used to assess the sample companies. Section 7 gives shortly the methodology, which is followed by the presentation of the results of this research. A discussion of the results and further implications of these follows in section 9. The last part concludes the paper.

5. Literature Review

The Basis for this Study

Edmans (2011) forms the foundation for the idea of this study. He studies the relationship between employee satisfaction and stock returns in the long run in the US. As a measure for employee satisfaction he uses a list of the “Best companies to work for in America”. A variation of portfolios is then created using all companies, which are included in that list and the portfolios are updated on a yearly basis. A commonly used Fama-French-Carhart (FFC) four risk factor model is then evaluated to assess the performance. The results are significant alphas that are robust with different benchmarks. Two further findings are presented by Edmans. The first one shows that the market apparently does not value intangibles fully. The list used in his study is publicly available and Edmans leaves the market with enough time to incorporate it in the stock prices. The outperformance is still existent and indicates that the market does not include the information in the pricing of the stocks. His study further shows that SRI screens may improve portfolio performance, although there is also partially some evidence from other authors against it, as will be shown below.

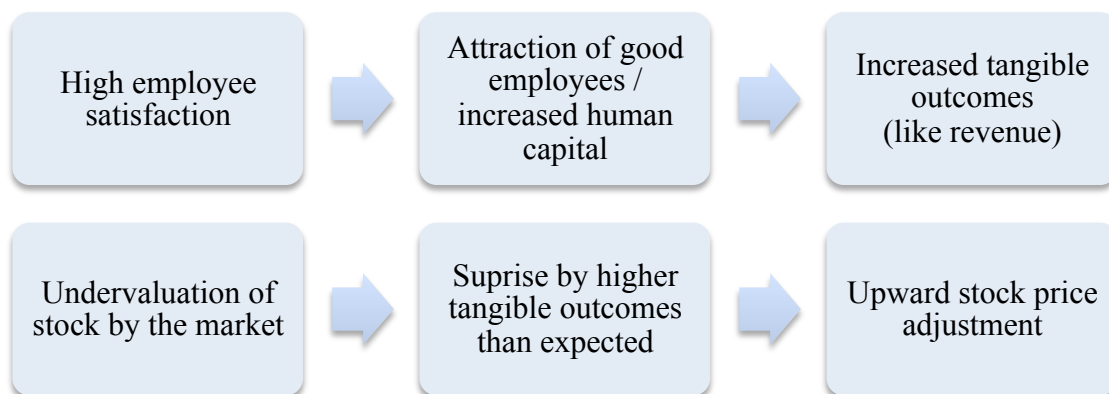
In a more recent study, Edmans, Li and Zhang (2014) extend this type of study to 14 other countries and analyze whether labor market flexibility in each country could explain outperformance of this type of SRI screen. They find that abnormal returns seem to exist when the labor market flexibility in a country is high. This is the case, for instance, in the US and the UK. Countries like Germany do have lower labor market flexibility and therefore do not show abnormal returns. These findings suggest that in countries with less restriction by legislation, in terms of firing and hiring, employee satisfaction could be value enhancing. The motivation, recruitment and retention of employees are affected more positive by employee satisfaction in flexible labor markets than elsewhere. Employee satisfaction is simply more valuable, if legislation does not provide a certain standard of employee welfare already.

According to the authors, diminishing returns might result from further increasing spending on employee welfare and can therefore be detrimental to stock performance.

Recognizing from both studies above, two points seem to be needed in order for employee satisfaction to have an impact on long run returns. First of all, employee satisfaction needs to be somehow beneficial to the firm value. It needs to affect, for instance, the motivation of workers for them to work better and therefore result in increased output or better products. Secondly, the stock market needs to misprice the effect of this employee satisfaction on firm value initially. This means that the potentially positive effect of employee satisfaction on the firm value is undervalued initially and correctly valued in the long run when it results in tangible outcomes, like increased earnings. The following two schemes (figure 1) show the above-mentioned logic. The upper scheme shows the process in the company and the lower one shows the process in the stock market.

In the following subsections, the possible effects of employee satisfaction on firm value and the mispricing of intangibles like employee satisfaction will be discussed. Additionally some evidence for the effects of other SRI screens on portfolio performance will be presented.

Figure 1: Impact of employee satisfaction on long-run stock returns within the company (upper scheme) and in the stock market (lower scheme)



Employee Satisfaction and Firm Value

The idea of employee satisfaction or human capital as an important part of a business is not new. Already in the early 20th century, Taylor (1911) writes about the importance of human capital in creating more efficiency in production. By now, the ideas he mentions are standard, but in 1911 these ideas were new and revolutionized production processes. Back then, no

training of the workforce was implemented and a general idea like “captains of industry are born, not made” was normal. An antagonistic relationship between the management of the company and normal workers was common and hindered efficient production. His ideas are related to ideas of motivation of the workforce to be beneficial for the company. Workers with satisfied needs and increased motivation perform better (Maslow, 1943). According to Taylor, prosperity for both employer and employee can be increased at the same time, without it being mutually exclusive. This increase in prosperity, however, can only be obtained through the highest efficiency possible. While in the old system workers were forced to produce more for the same pay and without knowing how to be more efficient, the newer system suggested making an analysis of the processes and introducing the best practices to the workers. Furthermore, workers should be paid accordingly higher, if their productivity increases, which would further increase the workers motivation. The higher pay would result in financial safety for the worker and thereby satisfy an essential need of him (Maslow, 1943). With this need satisfied the worker could be more focused on his work. The training of the workforce and overall increased cooperation between workforce and management were the essential new ideas of Taylor. Another factor that affected the productivity of workers back then was that workers were all paid the same amount, independent of their individual qualities. A worker that could perform the work better or faster therefore slowed down to the pace of the slowest worker to avoid unnecessary extra work. It would not make sense for him to work more than his colleague for the same pay. Taylor therefore argues that management and workers have to cooperate with each other and that workers have to be trained in order to achieve higher efficiency. Investing in human capital gives an advantage against competitors in his opinion. It is worth mentioning that his analysis is based on manufacturing companies and not e.g. the modern service companies of the 21st century. The main ideas, however, should be applicable also to those companies.

By now, the nature of the firms changed. Businesses from the early 20th century were very asset-intensive and were highly vertically integrated with very strong control over their employees. Nowadays, human capital is emerging as one of the most crucial asset that firms have, due to need of process innovation and quality improvement (Zingales, 2000). This changes how employees have to be considered within a firm and has implications on corporate governance and valuations of firms. Employees can also be seen to have more power than years ago, as they can leave a company and take with them their expertise and quality, which can leave a company in troubles. Zingales (2000) gives an example for the

importance of human capital. An advertising agency loses its key executive employees, who afterwards form their own company and attract key accounts from the old business. In this case, the executive employees made use of their acquired personal relationships with old clients to lure them into the new business. While the damage of something like this can be quite small in some industries, it certainly is quite essential for an advertising agency. This means that employee satisfaction might be more valuable for industries where human capital is essential compared to industries where it plays a minor role in the business. This study will try to shed some light into this idea.

Some older studies explain the relation between wages and productivity. Higher wages could be related to higher employee satisfaction or at least be one factor of many that make employees satisfied and therefore might show how employee satisfaction could affect productivity. In one of these studies, Shapiro and Stiglitz (1984) show that increased real wages result in involuntary unemployment in the economy, which in turn serves as a threat to workers. The involuntary unemployment results, because companies hire less people for the higher real wage paid. If workers are caught shirking they face the punishment of being fired, which can be seen as a cost for the employee. This punishment would not exist, if there would not be unemployment. Without unemployment, workers would theoretically be fired and rehired at the same time. Therefore, higher wages indirectly reduce the probability of shirking and ultimately increase the productivity of employees. This explanation of the relationship of real wages and productivity is more indirect than the above-mentioned approach of Taylor (1911). Instead of having a direct effect on motivation of the workers, the productivity in this case is increased via the threat of unemployment that results from the higher real wages. Although the end effect is achieved through different channels, it is ultimately the same, namely higher productivity.

Still in the context of wages, Akerlof (1982) uses a more sociological approach that is based on the reciprocal nature of gift giving. He argues that employees, which receive wages in excess of the minimum required wage, tend to work more and better in return. The employee perceives the excessive wage as a gift, while the firm sees the extra amount of effort from the employee as a gift. This idea only works, however, when workers acquire sentiment for the firm, which ensures that the workers obtain utility from “making a gift” to the employer. This sentiment is more likely to be found in workers that work in the same company for an extended time. Traditional ideas, like the old ideas of Taylor (1911), assume that labor is

hired like a factor of production and works like capital. The idea of gift giving, however, takes into account the willingness of labor to cooperate with the firm to get the most productive use of labor. An employee that receives excess wages therefore likely responds with increased effort.

In a study about the relationship of corporate social responsibility (CSR) performance and shareholder value, Hillman and Keim (2001) show some evidence for increased shareholder value through better relations to employees. They argue that there are two types of CSR: stakeholder management and social issue participation. After analyzing data from the S&P 500, the conclusion is that the management of primary stakeholders, like employees, customers, suppliers and communities increases shareholder value, while social issue participation has a negative effect. Primary stakeholders, especially employees, might create intangible and valuable assets that may lead to competitive advantages. Furthermore, good relations to customers and suppliers might increase loyalty of these stakeholders. In contrast to this, social issue participation, defined as elements of CSR without direct relationship to primary stakeholders, seem to be shareholder value destroying. These results could explain, why some SRI screens work better than others. SRI screens that target issues of primary stakeholders, like the one used in this study, might improve portfolio performance more than others.

The above-mentioned studies share the idea that employee satisfaction or higher motivated employees have an enhancing effect on firm value and mostly see human capital as one of the key assets of a modern firm. There are, however, also studies that find evidence for factors, which increase employee satisfaction, to have a negative effect on firm value. Abowd (1989) finds statistical evidence for an inverse relationship between labor compensation and shareholder value. The strength of this effect on each other does depend, nonetheless, on the sign of the effect. Unexpected union wealth decreases have a lower positive effect on shareholder gains than the negative effect of unexpected union wealth increases on shareholder wealth. In other words, labor cost reductions are not perceived as positive by the market as labor cost increases are perceived negatively by the market. The results give empirical support for the idea that investors think that managers make profit-maximizing decision with regard to employment decisions. Moreover, these findings contradict the idea that union activity enhances the productivity of employees. It seems that shareholders do not expect to get compensated for the increased unexpected labor cost through an increase of

productivity. The shareholder reaction instead shows that they expect to bear the full cost of unexpected labor cost increases. These findings stand somewhat in contrast to what Shapiro and Stiglitz (1984) argue, although they do not analyze a direct link of labor compensation and shareholder value. They analyze the relationship between labor compensation and productivity instead. Higher productivity, however, should be associated with higher shareholder value.

In another attempt to see how employees might affect the firm value, Gorton and Schmid (2004) analyze the implications of labor representation in German supervisory boards on the decision making process within a corporation. They compare companies with equal representation of employees and shareholders in the supervisory board to companies, in which the employees are represented only by one third. Companies with higher representation of employees tend to trade on the stock market on a 31% discount compared to their peers. The authors argue that this might be the case, because the supervisory board representatives of the employees tend to maximize a different objective function than the supervisory board members picked by the shareholders. They simply act more in the interest of employees and might be less likely to accept restructuring processes that could harm employees, but would otherwise be beneficial for the firm. This reflects the idea of Taylor (1911) that there exists an antagonistic relationship between the workers and management, instead of cooperating with each other for the greater good. The existence of labor representation in supervisory boards is also in line with the ideas of Zingales (2000), who explains that employees are more powerful in modern firms. The shareholders in turn respond to the higher allocation of supervisory board seats to employees by linking the supervisory board compensation to the firm performance and taking on more debt. With this action, the shareholders align the interests of the supervisory board members to their own. Nonetheless, it is clear that these actions come with increased costs, which cause the stock market to trade the company at the above-mentioned discount.

It has to be noted that employee satisfaction is a broad term. It may consist of several factors that might be either value enhancing or destroying, which makes it difficult to clearly state whether employee satisfaction is beneficial for a company. This ambiguity is also reflected in the available literature. This research will be related to the idea of Zingales (2000) that employee satisfaction might be more valuable in certain industries and less valuable in others. The sample companies will be divided into human capital intense and less intense

industries in order to check whether an SRI screen based on employee satisfaction has indeed different effects on different types of industries. This approach assumes that human capital is a key asset in one industry and not in the other. Furthermore, the wage level, as mentioned by Akerlof (1982) to have a positive effect on employee performance, will be part of the score to evaluate the companies.

Traditional View on Motivation

Motivation is what ultimately drives a human being. It is therefore worth having a look at it in more detail in order to understand it also in a more work related context. Motivated workers are likely to perform better at their jobs than their unmotivated peers.

Human beings have five basic needs that need to be satisfied to some extent according to Maslow (1943). These basic needs are divided into the categories: physiological, safety, love, esteem and self-actualization. People that are satisfied in these needs are the ones that are the happiest and can focus on the things they can do best. Their motivation will be higher. If a person is of the creative type, the person will be able to come up with creative ideas or innovations. Employment satisfaction could partially satisfy three of the above-mentioned needs. The employer can provide safety to the employee, if the company provides a pension fund or health care. Furthermore, the employer can especially provide financial safety through increased wages. The findings of Taylor (1911) and Shapiro and Stiglitz (1984) that increased wages increase productivity are in line with the theory of Maslow (1943). As long as an employee is treated with respect and shown gratitude by his employer, the employment situation could have a positive effect on the esteem of the employee. Lastly, if the employee is exceptionally happy with his job and thinks that what he is doing is what he was kind of “born” to do, his self-actualization need is satisfied.

Management sometimes might start to wonder why employees are not as productive as they could be, although working conditions are good, fringe benefits existent and wages are reasonable. In other words, the basic physiological and safety needs are satisfied. McGregor (1960) tries to give an answer to this question. He builds upon the theories of Maslow (1943) and explains that employees that are satisfied in their basic needs, but not satisfied with their higher needs (social and egoistic needs), tend to demand higher wages as a compensation for the deprived needs. Management needs to understand this logic and try to satisfy the higher

needs as well, if it wants to avoid excessive wages to keep employees satisfied. Employees would rather have all their needs satisfied instead of earning more money, as a substitute for certain needs. The logic is strengthened by the findings of Abowd (1989), who shows a negative relationship between labor compensation and shareholder value. It might indeed be the case that the lack of satisfaction of higher needs leads to excessive wages demanded by employees that ultimately lead to a decrease in shareholder value. Employee satisfaction therefore could save costs due to less necessity to pay excessive wages.

Measures to satisfy the higher needs could be the decentralization and delegation of the organizational structure. This would free people from too close supervision and provide the employee with freedom to direct its own activities. Furthermore, employees would need to take more responsibility and thereby also satisfy their egoistic needs. Another suggestion by McGregor (1960) to increase higher level need satisfaction is the change of the performance appraisal process. In older systems the employee was evaluated as if it was a product and did not have any say in the whole process of performance assessment. The new approach should include the employee and give him the opportunity to set targets together with management. Management should, nonetheless, maintain in the leadership role in this process. With this approach, the employee once again needs to take more responsibility in appraising his own addition to the organization. The self-fulfillment and egoistic needs will be satisfied considerably more in the later approach.

The scoring that will be used in this study to assess each company's employee satisfaction level will try to capture especially the satisfaction of the safety needs mentioned by Maslow (1943). The wage level and whether the company provides e.g. pension fund, health care or insurance will be part of the scoring. This should reflect the individual's financial safety obtained from the employment. Following the logic of Maslow, a higher scoring should mean a higher satisfaction of the basic needs, which in turn should result in higher performance by the individual.

Mispricing of Intangible Assets

For employee satisfaction to have a positive effect on portfolio performance, it is not sufficient for it to be beneficial to the firm in financial terms. It also needs to be somehow mispriced in the short term by the markets. Since employee satisfaction can be seen as an

intangible asset, this is not unlikely to be the case. It is difficult for the markets to evaluate the impact of intangible assets on financial performance. This difficulty often arises due to lack of information on the stock market side. It is therefore worth it to discuss some evidence on mispricing of intangibles of different sorts.

Aboody and Lev (1998) analyze the relevance of intangibles in form of software capitalization and its amortization to find that these are significantly associated with future earnings. The capitalization of software can be seen as a random factor affecting the earnings. This randomness is difficult to be incorporated by analysts in their forecasts of earnings, which results in wrong forecasts of earnings. Eventually these wrong forecasts lead to wrong estimates of stock prices.

Stocks with high research and development (R&D) relative to the market value of equity seem to have excess returns over a period of three years according to Chan, Lakonishok and Sougiannis (2001). An explanation for this seems to be that those companies invest heavily in R&D although they have been past losers judging from their equity market value. In those companies, the managers seem to be quite confident about their research, since they continue with the R&D expenditures despite cost saving pressure. The market does not seem to have this information and therefore undervalues the company. Even having sufficient information and enough time to incorporate it into the stock price does not exclude the possibility of mispricing by the market, which is shown by Edmans (2011). Jenter, Lewellen and Warner (2011) strengthen the theory of managers having more information than the market, as they find that managers are able to predict to some extent future stock returns of their company up to 100 days into the future. Given these results, it is likely that the same predictive power is existent for outcomes of R&D. Furthermore, Chan, Lakonishok and Sougiannis (2001) show that the revision of the expectations of the markets tends to be slow, as can be seen by the excess return persistence over 3 years. However, evidence for a direct link between R&D spending and future stock returns is not given. Spending on advertising shows similar results, which is somewhat in line with the results obtained by Hanssens and Joshi (2010). They show in a conceptual and empirical way that advertising increases the firm value, as measured by market capitalization. Interestingly this impact does not only work through an indirect way via increased sales and consumers, but also directly through investors. Investors react positively to advertising before the advertising affects sales and consumers, which in turn increases the firm value. This direct effect on investors stands somewhat in contrast to

studies, which find that markets misprice intangible assets and are slow in adjusting their expectations based on the intangible asset (Edmans, 2011). In this case the market seems to incorporate the information before it materializes in increased sales. It is worth mentioning, however, that advertising is not necessarily an intangible as employee satisfaction is. It is therefore easier for the investors to quantify it and likely forecast the impact on e.g. sales. Furthermore, there seems to be a negative impact on the stock price when competitors of similar size advertise heavily.

Instead of R&D expenditure, Deng, Lev and Narin (1999) use patent citations and try to predict market-to-book ratios (M/B) and stock returns. They argue that only disclosing R&D spending is not enough for investors to evaluate the quality of the research done. Patent-related measures seem to reveal more information on the nature of research, e.g. whether it is basic or applied research, and about the outcomes of research. Indeed they find evidence that patent-related measures are positively and statistically associated with M/B and stock returns. The link to stock returns is weaker though than for M/B, because stock returns only reflect new information that is initially not known to investors, while M/B shows the growth expectations independent of the moment when the information arrives at the market. In other words, stock prices contemporaneously incorporate the information embedded in R&D intensity better, if not fully.

Ben Tanfous (2013) further finds that the combination of intangible assets, like R&D expenditures, advertising expenses and training of employees, has a positive impact on value creation, measured by market capitalization. The study was, however, regionally restricted to non-financial French companies. The link between the training of employees and value creation is also consistent with Taylor (1911). It will also be an important component of the scoring that will be used later on to evaluate the employment quality in each company. These findings are in line with Jacobson and Mizik (2003) that basically find that value creation through R&D (e.g. innovation) is not enough to guarantee financial success for the firm, which is also what Chan, Lakonishok and Sougiannis (2001) find. The created value by R&D needs to be appropriated by investing in advertising, reputation and brand effects. A firm that has great products through innovation also needs to advertise them, because the market might otherwise never really get to know those products.

The evidence on intangible assets to be mispriced by markets seems to be quite strong. This mispricing might lead to abnormal stock performance in the long run of companies that invest in intangible assets, given that those translate into improved tangible assets in the future. This research will also focus on an intangible asset, namely employee satisfaction. It differs therefore from advertising and R&D expenditures analyses. Following Ben Tanfous (2013) and his analysis of the impact of the training of employees on value creation, this factor is part of the scoring to assess the employment quality level of each company in the data set. Obtaining positive results in this study will also depend on the level of mispricing of the securities by the stock market in the short run.

Other SRI Screens and their Effect on Portfolio Performance

There exist several studies on the performance of SRI screens or fund performances that make use of SRI screens. These studies analyze positive and negative screens for different regions in the world. The results are ambivalent and can be divided into three categories. Some authors find enhancing effects of SRI screens on performance, others find no statistically significant effects or mixed effects and again others do find SRI screens to be detrimental for performance. Evidence for each of the possible outcomes will be discussed, beginning with the positive evidence.

In an early work on the subject of SRI, Moskowitz (1972) argues that socially sound investments are not necessarily financially unsound. According to him, corporations that are aware of socially important topics possess a special sensitivity that might enable them to outperform their competitors. He identifies certain companies that exert socially responsible behavior. His study is, however, not specific in terms of which social aspects are important. Hillman and Keim (2001) in contrast do make a distinction between social responsibility that affects stakeholders or not. The later being value destroying, while the first increases shareholder value. Also Brammer et al. (2006) divide social performance into three measures, namely community performance, environmental performance and employee performance.

Derwall, Guenster, Bauer and Koedijk (2005) find further empirical evidence in favor of SRI. As an SRI screen they use “eco-efficiency”, which is a measure for the economic value created by the company in relation to their waste production. In a best-in-class approach, two portfolios, one portfolio with high eco-efficiency companies and one with companies of

lower efficiency, are constructed and evaluated. The portfolio with the highly efficient companies outperforms the other portfolio and the results are robust, when including transaction costs or accounting for investment styles, industry specific factors or market sensitivity. On the one hand, this finding is somewhat unexpected, as it might be argued that companies that comply with environmental standards face costs that ultimately lead to higher prices for their products and decrease shareholder value. On the other hand, scoring high on eco-efficiency might improve the efficiency of input-output and lead to a competitive advantage. The impact on the stock prices, however, depends on the ability of the market to correctly price the possible future financial outcomes resulting from the environmental responsibility. Instead of using an environmental screen, Edmans (2011) and Edmans, Li and Zhang (2014) use a positive social screen in form of employee satisfaction and find positive SRI effects.

Many studies do not find either economically or statistically positive or negative effects of SRI. These results confirm the ideas that the markets do not price social responsibility. Hamilton, Jo and Statman (1993) indirectly analyze the performance of SRI screens by looking at mutual fund performances that apply SRI screens. They find that socially responsible mutual funds do not have statistically significant excess returns and also do not outperform non-socially responsible mutual funds. The sample of socially responsible mutual funds might be, however, rather small, since it only consists of 32 funds. According to the results, investors would therefore neither lose nor gain from investing in socially responsible funds. Since the mutual funds apply several SRI screens at the same time it is not clear whether specific SRI screens increase portfolio performance, while others have negative effects. This study will focus on a single SRI screen instead.

Focusing more on international evidence from Germany, the UK and the US, Bauer, Koedijk and Otten (2005) also do not find statistically significant differences in the returns between ethical and conventional mutual funds. A big difference in their analysis, compared to Hamilton, Jo and Statman (1993) is the usage of the Carhart (1997) multi-factor model, which allows for better evaluation of the active investing and mutual fund performance. In terms of investment styles, German and UK ethical funds seem to be invested more in small cap stocks than their US counterparts. Due to the analysis of several sub periods, the authors furthermore conclude that ethical funds had a so-called “catching up phase”, because the

ethical funds underperformed the conventional funds in initial sub periods. This could be caused by a learning phase that ethical funds had to undergo.

Extending the study by analyzing Canadian data, Bauer, Derwall and Otten (2007) find very similar results. No significant difference between ethical and conventional mutual fund performances is found. One major difference to previous studies is that the authors check whether ethical fund returns correlate more with ethical index returns or conventional index returns. The surprising result is a higher correlation to the conventional index returns, thereby indicating that the ethical component of ethical mutual funds might not be as high as expected.

Schröder (2007) criticizes somewhat the use of investment funds to analyze SRI performances and uses SRI indices to assess SRI screen performance instead. Two main reasons justify the use of indices instead of funds. First of all, it might be difficult to assess correctly the transaction costs of investment funds. Secondly, the impact of portfolio managers needs to be taken into account in investment funds. Both reasons might make it more difficult to attribute possible outperformance strictly to the SRI screens. This approach seems to allow a more direct measurement of the SRI screen effect. His results are, nonetheless, in line with a non-existent statistical difference between SRI equity index returns and conventional benchmark indices like MSCI indices. Guerrard (1997) obtains similar conclusions. He compares the returns of a screened equity universe with an unscreened universe for the period of 1987-1994 and does not find statistically different results between both. The differential for both universes is only 15 basis points for equally weighted annualized stock returns. Furthermore, he argues that some of the apparent positive effect of SRI in other studies might occur due to the fact that the exposure of the constructed portfolios to large cap and growth stocks is higher. These stocks might be the actual drivers of the performance. The social screens he uses are military, nuclear power, product (alcohol, tobacco and gambling) and environment.

Instead of only looking at companies that are positively screened by SRI screens, Statman and Glushkov (2009) attempt to analyze the returns of companies usually shunned by socially responsible investors, like firearm or tobacco producers. The screens used are similar to the ones used by Guerrard (1997). They find that the net effect of following both strategies is no significant difference between socially responsible and conventional investors. Using a best-

in-class approach to invest in socially responsible companies does increase returns. Not investing in companies that produce immoral products, however, puts the socially responsible investors in a disadvantage. These so called “sin stocks” seem to outperform other stocks and therefore compensate conventional investors for the outperformance of the socially responsible stock in the portfolio of responsible investors. The outperformance of “sin stocks” is in line with the evidence provided by Hong and Kacperczyk (2009).

In contrast to the above-mentioned studies, the following papers do find evidence against the idea of portfolio performance improvements through SRI. Renneboog et al. (2008) analyze in their paper SRI fund performance for several regions. SRI funds from the US, UK, many continental European countries and Asia- Pacific underperform their domestic benchmarks. The risk-adjusted returns are, however, mostly not statistically different from zero when they are compared with returns of matched conventional funds. The main models used are the usual CAPM model and the Fama-French-Carhart (FFC) model with four risk factors, which is mostly used for this type of analysis. They argue that one difficulty in comparing different results of different papers is sometimes that different models are used. Some studies just use the CAPM model, which does not account for any other risk factor, but the market premium. Furthermore, having only a small sample size, different sample periods and benchmarks, worsens the comparability of results. The authors give several possible reasons for each underperformance and outperformance of SRI funds compared to conventional funds. Underperformance might result due to the restricted investment opportunities that social screens produce. This is in line with the findings of Statman and Glushkov (2009), that not being able to invest in companies that produce immoral products, puts responsible investors at a disadvantage. Most SRI funds do not invest in so called “sin stocks”, although Hong and Kacperczyk (2009) find that these kinds of stocks outperform the market. SRI funds are therefore excluded from the opportunity to make financially attractive investments. Screening intensity can further increase underperformance of a fund. The more screens an SRI fund uses, the more restricted the investment opportunities get. In addition, companies that have high ethical standards can be overpriced by the stock market, because investors that are averse of unethical companies increase the demand for ethical companies and as a result raise their stock price. Reasons for a possible outperformance include the signaling effect of good environmental or social performance. Companies that score high on these subjects might signal good management within the company. This in turn might then result in better financial performance, when compared to competitors. SRI funds might also outperform,

since some screening is labor intensive and might result in relevant information advantages that conventional funds do not have.

Geczy et al. (2003) also find costs associated with SRI constraints. They analyze SRI mutual funds and conclude that the costs depend on, which asset pricing model is used and whether investors believe in the stock-picking abilities of fund managers or not. Using only the CAPM model and believes in low managerial skills results in lower differences. Models with higher exposure to size, value and momentum effects, like the FFC model, reveal a much higher cost of the SRI constraint. The cost is then around 30 basis points per month. The authors do, however, acknowledge that investors might obtain non-financial utility from investing socially responsible, which is not accounted for in their study and is difficult to be quantified. Therefore the results can be seen as lower bounds of the non-financial utility that investors need to get for them to be willing to invest socially responsible.

Instead of using fund data, Brammer et al. (2006) use firm level data, like Schröder (2007). They furthermore use disaggregated social responsibility data. The disaggregated data on the social performance allows distinguishing better, which part of social performance actually drives the financial performance in form of stock returns. The social performance is divided in three measures: community performance, environmental performance and employee performance. The composite measure of all three performance factors shows a negative relation with stock returns. A look at the three measures individually reveals though that the first two measures show a negative correlation with stock returns, while the employment factor is slightly positively correlated, which is in line with the results of Edmans (2011) and Edmans et al. (2014). These different effects for the three measures, suggests that a more detailed look at social performance components might make more sense than just looking at an overall measure of social performance. Hillman and Keim (2001) also suggest a more detailed look, instead of a general “score” for overall social performance.

Overall, it becomes apparent through this discussion that the evidence on the effects of SRI on financial performances is not necessarily clear. It would be tempting to say that the evidence might be tilted towards a more negative effect of SRI on performance, but specifically for the screens that only consider employment quality the evidence favor a more positive view (Brammer et al., 2006; Edmans, 2011; Edmans et al., 2014). This is one of the

reasons why a closer look on, and further investigation of employment satisfaction as an SRI screen might be worthwhile.

Although many studies focus on fund performances using SRI screens, this study will refrain from doing so as it is interested in testing a specific SRI screen. Since most funds use multiple SRI screens this would make it difficult to assess the performance of a single screen. Evidence that different SRI screens have different effects on performance further motivates this approach. Therefore, firm-level data instead of fund data will be used. Following the logic of e.g. Derwall et al. (2005), the portfolios to be tested will be constructed with a best-in-class approach. The model to be used is the FFC model, as it captures risk better than, for instance, the CAPM. Renneboog et al. (2008) also argue in this way.

6. Data and Summary Statistics

The following analysis is done for three distinct regions, which are the US, Europe (consisting of all countries that adopted the Euro in 1999) and Japan. The total sample is composed of 2,019 companies, which can be divided as follows into the different regions: 1,143 from the US, 446 from Europe and 430 from Japan. Some information on industry classifications is retrieved from the U.S. census bureau and will be explained more detailed later on. In order to continue the analysis, all companies need to be identified as either being part of a high human capital intense industry or a low one. The following paragraphs will explain the process of the industry division, the score construction to assess the employment quality and discuss the descriptive statistics of the resulting industry divisions.

Division of Companies

In order to determine whether an industry should be considered a human capital intense one or not, it is necessary to first decide on how to measure human capital overall. Early empirical studies use, for instance, school enrollment rates to assess the level of human capital in the economy (Barro, 1991). Aiming in the same direction and putting big importance on schooling, Barro and Lee (1996) use distributions of educational attainments to get to measures of average years of schooling. This should then reflect the best estimate for human capital levels. Mulligan and Sala-i-Martin (2000), however, suggest that using the years of schooling to obtain the human capital level might not be as appropriate as thought

earlier. They mention four reasons for why this is the case. First of all, workers of each education category are assumed to be perfect substitutes for workers of all other categories, which is not necessarily the case. Secondly, it assumes that workers with, for instance, 15 years of education are 15 times as productive as a worker with only one year of schooling, independent of wage rates. As a third reason they criticize that the elasticity of substitution across workers of different groups is assumed constant all the time. Lastly, one year of schooling is assumed to increase the skill by the same amount all the time. Instead of average years of schooling, Mulligan and Sala-i-Martin (2000) use labor income to construct a measure for human capital. The reasoning behind this logic is as follows. In case that a person studies something very useful from a production point of view, the wage of the person would be accordingly higher in the market than for somebody that maybe studied longer, but studied a field that is not as useful or not demanded. A worker that produces a lot is valued more and therefore earns more.

The labor income approach is favored in this study over the schooling years approach, not only due to easier data availability for the first, but also due to the convincing logic behind it. The market determines how valuable a worker is and this can be easily seen through a workers wage.

Following this logic, information on wages from the U.S. census bureau for each industry classification is retrieved. It published in December 2013 a list with data on the number of firms, number of establishments, employment, and annual payroll by enterprise employment size for the United States in 2011. All industries are listed according to their North American industry classification system (NAICS) code. For each industry classification the average annual payroll per employee is calculated. It has to be noted that this approach uses the US data as some kind of benchmark and assumes that wage distributions in the other two regions to be analyzed are similar to the one in the US. Since this study is more interested in the wages of an industry relative to another industry and not in absolute terms, the assumption seems to be legitimate.

In the next step, information on the industry classification is retrieved from datastream for all the companies in the whole Asset4 ESG Universe. The Asset4 ESG dataset from Thomson Reuters will be explained in more detail later on. For now it is only important to know that it consists of 4615 companies from 60 countries and should give a reasonably good worldwide

representation for the division of industries. As the industry classification of the companies in the datastream universe follows the Thomson Reuters Business Classification (TRBC), the different classification codes have to be matched to the NAICS codes retrieved from the U.S. census bureau. Another point to be considered is the fact that a company can be active in several industries at the same time. The main revenue driver e.g. might be gold ore mining, while the second and third highest revenue driver might be copper or nickel ore mining. This should be considered when attributing the payroll data from the U.S. census bureau to the individual companies. Not taking this information into account might give misleading results in the ultimate industry division.

To solve this issue, datastream offers the possibility to retrieve up to eight standard industrial codes (SIC) for each company according to the revenue level it obtains from a certain industry. In other words, this means that retrieving the “SIC 1” from datastream for a company gives the industry code in which the company gets the biggest share of its revenues from. The “SIC 8” gives then in contrast the least relevant revenue driver from the perspective of the company. All SIC codes for all companies in the Asset4 ESG dataset are retrieved for the analysis.

With the SIC codes given, each of them is matched to the corresponding NAICS code. This matching is done with the help of an excel file, which is also provided by the U.S. census bureau and contains SIC codes and their equivalent NAICS codes. Afterwards, the average annual payroll data obtained earlier is matched to the SIC codes from each company. By this time, each company is linked to several payrolls of each industry in which it generates revenues according to its eight SIC codes. Once again it is worth mentioning that it is assumed that the overall payroll distribution throughout the industries worldwide are the same as in the US and the US data from the census bureau is applied to all international companies in the Asset4 ESG universe. Companies with missing data are excluded from the analysis as well as companies where the matching from SIC to NAICS does not work properly.

An example makes the procedure more understandable. “3I GROUP PLC” is according to the datastream SIC 1 and SIC 2 active in “portfolio management” and “other financial vehicles”. For those two industries the correct NAICS code and the corresponding average annual payroll information from the U.S. census bureau are matched. An average of the annual

payrolls using four different weighting strategies is then calculated in order to take into account the possibility that for some companies their revenue is more equally distributed throughout all industries it is involved in, while others might be more specialized in a single industry. The different weighting methods can be seen in table 10 in the appendix. Once again, after having the four averages per company an overall average is constructed, which is the final payroll attributed to the company. As mentioned above, Thomson Reuters uses its own classification system (TRBC). Therefore the average of all companies in a certain TRBC category is taken to get the average annual payroll per employee in that specific TRBC. The TRBC industries that rank in the top 50% are then considered the human capital intense industries. In the following table (table 1), the resulting division of industries is shown.

Table 1: Division of Industries

High Human Capital Intensity	Low Human Capital Intensity
Investment Banking & Investment Services	Paper & Forest Products
Collective Investments	Diversified Trading & Distributing
Oil & Gas	Telecommunications Services
Biotechnology & Medical Research	Professional & Commercial Services
Banking Services	Construction & Engineering
Holding Companies	Office Equipment
Residential & Commercial REITs	Industrial Conglomerates
Insurance	Healthcare Providers & Services
Pharmaceuticals	Construction Materials
Electronic Equipments & Parts	Machinery, Equipment & Components
Communications & Networking	Containers & Packaging
Semiconductors & Semiconductor Equip.	Automobiles & Auto Parts
Uranium	Real Estate Operations
Natural Gas Utilities	Transport Infrastructure
Metals & Mining	Personal & Household Products & Services
Multiline Utilities	Water Utilities
Renewable Energy	Beverages
Electric Utilities & IPPs	Passenger Transportation Services
Oil & Gas Related Equipment and Services	Homebuilding & Construction Supplies
Coal	Leisure Products
Chemicals	Food & Tobacco
Media & Publishing	Household Goods
Software & IT Services	Food & Drug Retailing
Computers, Phones & Household Electr.	Textiles & Apparel
Aerospace & Defense	Hotels & Entertainment Services
Healthcare Equipment & Supplies	Other Specialty Retailers
Freight & Logistics Services	Diversified Retail

Scoring

In order to rank the companies on their employment quality, a score is formed using the Asset4 ESG dataset from Thomson Reuters. It is a dataset that is focused on environmental, social and governance (ESG) information, which covers over 4000 companies globally. Especially for SRI investing it is very valuable as it has more than 250 key performance indicators and more than 750 individual data points on ESG. The data is structured into four pillars: economic, environmental, social and corporate governance performance. For the purpose of this study, however, the focus will be on the social performance of the company. The social performance pillar of the dataset has seven subcategories, which in turn consist of several datapoints (table 7 in the appendix). The database already proved to be valuable in terms of its capability to be used to produce significant positive alphas. Eccles et al. (2013) use the database to form high and low sustainability company groups. An analysis of the stock performance of both groups then reveals that the high sustainability group outperformed the low sustainability group. It is therefore valid to assume that the database contains valuable information for investors that can be used to obtain alphas. This study is focused on an employment quality screen, hence only three of the above-mentioned subcategories remain relevant: employment quality, health & safety and training & development.

These three subcategories in turn consist of several factors. After analyzing the data availability for each factor in these subcategories for all available companies, only ten relevant factors remain. The criterion for including the factor in the analysis is the availability of that factor for at least 100 companies in each region (US, Japan, Europe) over a period of at least 9 years. The ten factors selected to form the scoring are shown in table 8 in the appendix. Six factors are selected from the employment quality subcategory, while one is selected from the health & safety subcategory and the other three from the training & development subcategory. The usage of a different dataset to determine in which companies to invest in is one of the differences to the study of Edmans (2011). While Edmans uses a list of “Best companies to work for”, which is publicly available, the Asset4 ESG dataset is not publicly available. It is necessary to have a subscription with Thomson Reuters to get access to it. This feature could result in an advantage for an investor, as it provides the investor with additional information that is not freely available to the rest of the market. Another difference to the Edmans approach is that several aspects of employment quality are used and an own

score is formed, while Edmans uses list inclusion in the “Best companies to work for”-list published by the Fortune magazine as a signal to invest in a company. This last point should not make a big difference though, since Fortune takes several aspects of employment quality into consideration, when the companies are listed on their list. Overall, it can be said that the approach at hand is more “raw” and does not include as many details and factors. For the construction of the “Best companies to work for”- list, more aspects are taken into account. This approach should, nonetheless, contain enough information to assess properly the employment quality in each company.

Out of the 10 factors used to form the score, two are of numerical nature, like the average salaries and benefits in U.S. dollars, while the other 8 factors are answers to yes/no questions. All factors are therefore transformed into z-scores to better combine them in an overall score. Especially for the yes/no questions a z-scoring approach is favorable as it brings more differentiation into the data. Asset4 ESG provides z-scores for some of the factors, but these are calculated using the global database. Since only three distinct regions are relevant in this study, the z-scores for each company are calculated manually, depending on all companies in its region. The transformation into z-scores follows the logic of Asset4 ESG. When a firm answers with a “yes” to the factor during a month, it is assigned a “1”, while a “no” results in a dummy value of “0”. It is not clear how Asset4 ESG deals in its z-scoring with companies that for some reason do not have data for that factor. These companies are therefore excluded from the z-score estimation by assigning them a dummy of “-“. This guarantees that they have no influence on the z-scoring of the other companies. The standard approach to calculate a z-score is then followed. The mean and standard deviation for all the companies in the region and in that specific month is used for that purpose. This procedure is repeated for every month. It is worth noting that one factor consists of two yes/no questions at the same time. In this case the factor equals a value of “0.5”, if only one of the two questions is answered positively.

An equally weighted average is then calculated for each month, based on all 10 z-scores that are attributed to the company. This average is the final score and the higher it is the more likely the company is to be included in the best-in-class portfolio for the month. In case that the company has less than 10 z-scores attributed to it, due to missing data, the score is still constructed with the fewer z-scores. The minimum requirement for constructing the final score is 8 z-scores. Otherwise the company is treated as not having any score at all.

Descriptive Statistics

All data for the descriptive statistics that will be discussed in this section are taken from datastream. The intangibles as a percentage of total assets, is constructed after data retrieval, given the data from datastream. The percentage is obtained by a simple division of “net (total) intangible other assets” by “total assets”.

The descriptive statistics of each industry division are discussed for each region and reveal some surprising insights for some of the characteristics (table 9 in the appendix). Beginning with a look at the market capitalization, companies being part of the high human capital industries are bigger than low human capital companies in the US and Europe. Given that the standard deviation is quite high though, the difference does not appear to be that big. In the US there is a difference in the mean of \$2.96 billion, while for Europe the difference is slightly higher at \$4.02 billion. In Japan both the differences in the mean and standard deviations are lower, indicating that the companies are closer to each other in terms of market capitalization. In fact, the maximum market capitalization in Japan is found in the low human capital industry.

The prices per stock seem to have more extreme variation in both the US and Japan. Means between the industry divisions are more different in both regions and the standard deviations are high. These results do seem to be influenced by extreme outliers, when looking at the maximum stock price in each region. In the US, the maximum value is \$100,185.46 per stock, which is attributed to Berkshire Hathaway. This obviously has a big impact on the descriptive statistics. Looking at the median reveals that the differences are not as big as suggested by the mean. For all three regions, the median does not seem to be of a big difference when comparing both industry divisions to each other.

The price/book (P/B) ratio is generally used by investors to see whether a stock is under- or overvalued. A low price/book ratio usually indicates that the stock price is undervalued. In the worst case it indicates that something fundamentally could be wrong. Each industry usually has its typical range for the ratio. Companies or industries relying more on human capital are expected to have higher price/book ratios as this might reflect the relative importance of labor in generating revenues. Higher intangibles, which could be expected from high human capital industries, should result in higher P/B ratios. It is therefore

surprising that a look at the mean differences between the industry divisions in the US and Europe indicates that the low human capital industries show higher P/B ratios. For the median it is even for all three regions. This observation might be an indication for a wrong division of the industries. It could be that some of the supposedly low human capital industries are in fact human capital intense. The analysis of the results from this study needs to take this into consideration. A misspecification of high and low human capital intense industries influences strongly the results.

The intangibles as a percentage of total assets show a pattern that is consistent with the P/B ratio and against the expectation. High human capital industries are expected to have more intangibles relative to their total assets, as they should have more intangibles, such as patents and copyrights. These higher levels of intangibles are usually the result of higher human capital. However, the low human capital industry companies show higher percentages in the US and Europe. A look at the median confirms this pattern. For Japan, the difference is almost non-existent, although it still shows higher values for the low human capital industry.

Overall, it is worth keeping in mind from this analysis that out of all regions, Japan seems to have the most homogeneous structure out of all three regions, when comparing the two industry types. Furthermore, the P/B ratio and the ratio of intangibles to total assets might raise the concern that the divisions of the industries are not correct. This might have strong implications for the main analysis later on, as it is a cornerstone for the performance evaluation of the portfolios.

7. Methodology

The above-mentioned scoring process is repeated on a monthly basis in order to form for each month the best-in-class portfolio. Given the different data availability for each region, the following time horizons are analyzed for the regions: the US has a time horizon from January 2002 until December 2012, Japan is analyzed from April 2003 until March 2013 and Europe from January 2004 until December 2012. Since the data updates of the Asset4 ESG database are not necessarily always on the first day of the month and could be sometimes as late as the 6th of the corresponding month, the monthly data is retrieved always for the 6th day of the month. In this way it is always ensured that the most updated information is obtained for the month. By doing this, the market is allowed to a small extent to already incorporate

the employment quality information in its stock pricing, when for some months, the data is published before the 6th of the month. The time given to the market to adjust is, however, not as much as in the paper of Edmans (2011), who gives the market one month to adjust.

The best-in-class portfolio out of each industry type consists of 20 companies with the highest overall scoring. This approach is favored over the percentile approach, because equally sized portfolios should be compared to each other. Given that after the division of the industries, some industry portfolios are bigger than the others, a percentile approach would have resulted in different sized portfolios. It has to be noted that during some months, some companies have the exact same score. In this case it could occur that the best-in-class portfolios are slightly bigger than only 20 companies. It is also worth noting that using 20 companies to construct the portfolios is a subjective choice. In this study, 20 companies are assumed to be the least amount necessary to guarantee a decent diversification level.

The model used to assess the performance of the portfolio is the commonly used Carhart (1997) four-factor model. Using this model should usually guarantee that possible outperformance does not derive from certain risk factors. These risk factors are the excess market return, a market capitalization factor, a book value to market value (B/M) factor and a momentum factor. The market capitalization factor is basically the return difference between a small capitalization portfolio and a large capitalization portfolio. It should ensure that the outperformance is not driven by the fact that many small companies are included in the portfolio. The book value to market value factor in turn is the return difference of a value portfolio with high B/M and a growth portfolio with low B/M. This factor ensures that performance is not caused due to high amounts of growth or value stocks in the portfolio. Lastly the momentum factor is the return difference between a portfolio of past “winners” and a portfolio of past “losers”. The precise construction of the individual factors follows the logic of Fama and French (1993) and Carhart (1997). The factors are retrieved from the online Kenneth French data library. For each region, the corresponding factors are used, the US factors for the US portfolios and so on.

The model to be estimated then looks as follows:

$$r_{i,t} - r_{f,t} = \alpha + \beta_1(r_t^m - r_{f,t}) + \beta_2 r_t^{smb} + \beta_3 r_t^{hml} + \beta_4 r_t^{mom} + \varepsilon_{i,t}$$

where:

$r_{i,t}$ = return on the portfolio; $r_{f,t}$ = one-month U.S. T-bill rate; r_t^m = return on a region's value weighted market portfolio; r_t^{smb} = market capitalization factor; r_t^{hml} = B/M factor; r_t^{mom} = momentum factor

The stock price data is retrieved from datastream and transformed into logarithmic returns. Although some of the companies ceased to exist during the time horizon, they are still included in the analysis to avoid any sort of survivorship bias.

In addition to taking the difference of the constructed portfolio return and the one-month U.S. T-bill rate as the dependent variable, also the difference to the overall corresponding high or low human capital industry portfolio return is used. Using this second approach should show whether the strategy produces higher and more significant alphas for the high human capital industry than the low human capital industry. One would expect the strategy to work better in industries where human capital is relatively more important, since having a better workforce should be valued more in those industries.

The decision to use the four-factor model from Carhart (1997) instead of the simple CAPM model or the three-factor model of Fama and French (1993) is simply due to the better risk adjustment. It is empirically shown by Carhart that the three-factor model does not account for the momentum strategy that is found by Jegadeesh and Titman (1993). Carhart then suggests adding this factor to the Fama and French model.

The portfolios will be formed both in an equally weighted fashion as well as in a value weighted way. The value-weighted portfolios will be constructed using the market capitalizations of each stock for the respective months in which the portfolio is constructed. The market capitalization data, however, is updated on a year-end basis, as Datastream only provides information of shares outstanding in this time interval. According to Plyakha, Uppal and Vilkov (2014), equal-weighted portfolios outperform their value- and price-weighted counterparts in terms of average return, four-factor-alpha, Sharpe ratio and certainty-equivalent return. The higher alpha of the equal-weighted portfolio seems to come from the

monthly rebalancing that is necessary to remain equal weights. Analyzing two kinds of portfolio weighting will serve as a robustness test for the results.

8. Results

The results of the regression analysis are presented and discussed in this section. The discussion is organized per benchmark and starts with the risk-free rate as a benchmark followed by the industry returns as benchmarks. For both benchmarks, equal- and value-weighted portfolios are analyzed.

Table 2: Risk-adjusted returns of equal-weighted best-in-class portfolios (risk-free rate)

Region	Low human capital					Adj. R ²	Obs. No.
	α	$\beta_{\text{mkt-rf}}$	β_{smb}	β_{hml}	β_{mom}		
<i>US</i>	-0.340	1.018***	-0.115	0.440***	-0.143*	0.575	132
<i>Europe</i>	-0.274	0.643***	-0.416*	0.244	-0.125	0.508	108
<i>Japan</i>	-0.363	0.888***	-0.414**	0.154	-0.206*	0.478	120
Region	High human capital					Adj. R ²	Obs. No.
	α	$\beta_{\text{mkt-rf}}$	β_{smb}	β_{hml}	β_{mom}		
<i>US</i>	-0.140	1.038***	-0.291*	0.110	-0.048	0.574	132
<i>Europe</i>	-0.171	0.534***	-0.686**	0.940***	-0.388***	0.537	108
<i>Japan</i>	-0.743	1.078***	-0.324*	0.366*	-0.314**	0.545	120

This table reports regression results of monthly returns of equal-weighted best-in-class portfolios using Carhart's (1997) four-factor model:

$$r_{i,t} - r_{f,t} = \alpha + \beta_{\text{mkt-rf}}(r_t^m - r_{f,t}) + \beta_{\text{smb}}r_t^{\text{smb}} + \beta_{\text{hml}}r_t^{\text{hml}} + \beta_{\text{mom}}r_t^{\text{mom}} + \varepsilon_{i,t}$$

where $r_{i,t}$ is the return on equal-weighted best-in-class portfolios in month t for region i and $r_{f,t}$ is the return of the risk-free rate. α is the intercept that captures the abnormal risk-adjusted return. r_t^m , r_t^{smb} , r_t^{hml} and r_t^{mom} are, respectively, the Fama and French (2012)'s regional factors on market, size, value and momentum. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. The sample period is for the US from January 2002 to December 2012, for Europe from January 2004 to December 2012 and for Japan from April 2003 to March 2013.

Beginning with the equal-weighted best-in-class portfolio with the risk-free rate as benchmark (table 2), the most important finding is that none of the alphas is significant. A different result was expected also for the signs of the alpha coefficients. Instead of being completely negative, the expectation for the signs was to be positive for both industries and when comparing both industries, the alphas for the high human capital industry should be higher than the ones for the low human capital industry. Having all alphas negative means

that none of the constructed best-in-class portfolios outperformed a simple investment at the risk free rate. Although the coefficients are insignificant, this is a bad sign for the portfolio performance in general.

The significance for the market risk premium factor was to be expected, since the regression has the risk free rate in the dependent and independent variable side. The high values of the factor for each portfolio indicate that the portfolios are moving with the market and are cyclical. The SMB factor is negative for all portfolios and significant in 5 out of 6 cases. This suggests a bias of all the portfolios towards large stocks, which in turn might be the result of the Thomson Reuters Asset4 ESG dataset to be overall biased towards large stocks. The HML factor is positive for all portfolios and significant in 3 out of 6 times. All portfolios are tilted towards the inclusion of value stocks and indicate once again that the overall dataset might be the reason for this. For the MOM factor, 4 coefficients are significant and all of them are negative. The negativity of the coefficients results from portfolio biases towards past “losers”. Overall, all portfolios throughout all regions and industries show the same signs although revealing differing significance levels.

The regression results for the value-weighted best-in-class portfolios (table 3) are very similar in terms of the coefficient signs and significance levels. The alphas are again negative and not significant for any constructed portfolio. The low human capital portfolios for the US and Europe show different coefficient signs on the MOM factor. They are slightly positive, but not significant. The European and Japanese high human capital portfolios seem to be best captured by the four risk factors, as they are all significant. The interpretation of signs follows the previous logic. The results strengthen the idea that the SRI screen performs bad and is worse than the performance of the risk-free rate. The weighting method of the portfolios does seem to be irrelevant for this finding.

The regression results using the industry returns as a benchmark differ from the previously obtained results with the risk free rate. For the equal weighted portfolios (table 4) the alphas stay insignificant most of the time and only the one for the US region in the low human capital industry is significant at the 10% level. The coefficient is -0.284 and suggests that the portfolio underperforms the low human capital industry. The alphas for the other two regions in this type of industry are also negative. For the high human capital industry, however, two alphas (US and Europe) are positive and do seem to outperform their respective industry.

Since they are not statistically significant, this result has, nonetheless, only small relevance. The coefficients for the market risk premium are smaller, suggesting a smaller movement with the market, as was to be expected, because the benchmark changed. The bias towards large stocks is still existent, while the HML factor shows a mixed picture throughout the regions/industries and is only significant in 1 out of 6 portfolios. The MOM factor shows two significant biases towards past “losers” in the portfolios of Europe and Japan for the high human capital industry. The other 4 portfolios show insignificant mixed biases for the MOM factor. Overall, it is also worth mentioning that the adjusted R^2 drops sharply when using industry returns as benchmarks. The models have therefore less explanatory power for the return differences to the industry returns than the risk free rate. This difference in adjusted R^2 should be mainly driven by the previously high significance of the market risk premium factor, which is not as significant anymore.

Table 3: Risk-adjusted returns of value-weighted best-in-class portfolios (risk-free rate)

Low human capital							
Region	α	β_{mkt-rf}	β_{smb}	β_{hml}	β_{mom}	Adj. R^2	Obs. No.
<i>US</i>	-0.078	0.806***	-0.404***	0.354***	0.009	0.495	132
<i>Europe</i>	-0.145	0.550***	-0.642***	0.165	0.048	0.449	108
<i>Japan</i>	-0.107	0.734***	-0.426***	0.193	-0.077	0.445	120
High human capital							
<i>US</i>	-0.101	1.028***	-0.453***	0.125	-0.093	0.544	132
<i>Europe</i>	-0.028	0.595***	-1.005***	0.889***	-0.363**	0.547	108
<i>Japan</i>	-0.541	1.100***	-0.430**	0.440**	-0.311**	0.550	120

This table reports regression results of monthly returns of value-weighted best-in-class portfolios using Carhart’s (1997) four-factor model:

$$r_{i,t} - r_{f,t} = \alpha + \beta_{mkt-rf}(r_t^m - r_{f,t}) + \beta_{smb}r_t^{smb} + \beta_{hml}r_t^{hml} + \beta_{mom}r_t^{mom} + \varepsilon_{i,t}$$

where $r_{i,t}$ is the return on value-weighted best-in-class portfolios in month t for region i and $r_{f,t}$ is the return of the risk-free rate. α is the intercept that captures the abnormal risk-adjusted return. $r_t^m - r_{f,t}$, r_t^{smb} , r_t^{hml} and r_t^{mom} are, respectively, the Fama and French (2012)’s regional factors on market, size, value and momentum. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. The sample period is for the US from January 2002 to December 2012, for Europe from January 2004 to December 2012 and for Japan from April 2003 to March 2013.

The value-weighted portfolios with industry return benchmarks also show low significance for their alphas (table 5). The European region has the only significant alpha at the 10% significance level and a coefficient of -0.330. This means that the European low human

capital portfolio significantly underperforms the low human capital industry, which is not necessarily in line with the expectations. Using the industry as a benchmark, the initial expectation was that the low human capital best-in-class portfolios should still outperform their industry peers. High human capital best-in-class portfolios should, nonetheless, also outperform their industry peers, but by a higher amount than the low human capital portfolios. None of the logics are confirmed when looking at both, equal- and value-weighted portfolios regression results. Three high human capital portfolio alphas of the regressions using industry returns are positive, while the other three are negative, with all alphas being insignificant. The SMB factor shows consistent results with the previously discussed models and shows a bias within all portfolios towards large stocks. The other factors do have mixed results and do not seem to follow a specific pattern.

Table 4: Risk-adjusted returns of equal-weighted best-in-class portfolios (industry return)

Low human capital							
Region	α	$\beta_{\text{mkt-rf}}$	β_{smb}	β_{hml}	β_{mom}	Adj. R ²	Obs. No.
<i>US</i>	-0.284*	-0.070*	-0.118*	0.066	-0.016	0.039	132
<i>Europe</i>	-0.208	0.018	-0.174*	-0.094	0.064	0.034	108
<i>Japan</i>	-0.008	-0.037	-0.364***	-0.047	-0.069	0.191	120
High human capital							
<i>US</i>	0.125	-0.154***	-0.134	-0.234***	0.038	0.221	132
<i>Europe</i>	0.295	-0.081*	-0.243**	0.182	-0.123**	0.101	108
<i>Japan</i>	-0.076	0.115**	-0.370***	0.025	-0.138**	0.244	120

This table reports regression results of monthly returns of equal-weighted best-in-class portfolios using Carhart's (1997) four-factor model:

$$r_{i,t} - r_{ind,t} = \alpha + \beta_{\text{mkt-rf}}(r_t^m - r_{f,t}) + \beta_{\text{smb}}r_t^{\text{smb}} + \beta_{\text{hml}}r_t^{\text{hml}} + \beta_{\text{mom}}r_t^{\text{mom}} + \varepsilon_{i,t}$$

where $r_{i,t}$ is the return on equal-weighted best-in-class portfolios in month t for region i and $r_{ind,t}$ is the equal-weighted return of the overall industry (high or low human capital). α is the intercept that captures the abnormal risk-adjusted return. $r_t^m - r_{f,t}$, r_t^{smb} , r_t^{hml} and r_t^{mom} are, respectively, the Fama and French (2012)'s regional factors on market, size, value and momentum. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. The sample period is for the US from January 2002 to December 2012, for Europe from January 2004 to December 2012 and for Japan from April 2003 to March 2013.

A direct comparison of results with the ones obtained by Edmans (2011) and Edmans et al. (2014) is not possible, as they apply the SRI screen on overall markets, while this study focuses on the SRI screen application on certain industries. Another important factor that

needs to be considered when comparing the results is that the sample period is different. While Edmans (2011) has a sample period from 1984-2009, the sample period in this study is from around 2002 until 2013 (depending on the region), which is overall smaller and includes more recent data. Nonetheless, the results do somewhat stand in contrast to the initially expected results. The alphas are mostly insignificant and do not necessarily show the expected sign for its coefficients.

Table 5: Risk-adjusted returns of value-weighted best-in-class portfolios (industry return)

Low human capital							
Region	α	β_{mkt-rf}	β_{smb}	β_{hml}	β_{mom}	Adj. R ²	Obs. No.
<i>US</i>	-0.231	-0.070	-0.129*	0.130*	0.030	0.068	132
<i>Europe</i>	-0.330*	0.018	-0.102	0.047	0.131**	0.040	108
<i>Japan</i>	-0.154	-0.102***	-0.096	0.093	-0.008	0.062	120
High human capital							
<i>US</i>	-0.203	-0.022	-0.015	-0.107	-0.060	0.006	132
<i>Europe</i>	0.036	0.050	-0.089	0.109	-0.161***	0.159	108
<i>Japan</i>	-0.118	0.130***	-0.305***	0.018	-0.146***	0.322	120

This table reports regression results of monthly returns of value-weighted best-in-class portfolios using Carhart's (1997) four-factor model:

$$r_{i,t} - r_{ind,t} = \alpha + \beta_{mkt-rf}(r_t^m - r_{f,t}) + \beta_{smb}r_t^{smb} + \beta_{hml}r_t^{hml} + \beta_{mom}r_t^{mom} + \varepsilon_{i,t}$$

where $r_{i,t}$ is the return on value-weighted best-in-class portfolios in month t for region i and $r_{ind,t}$ is the price-weighted return of the overall industry (high or low human capital). α is the intercept that captures the abnormal risk-adjusted return. $r_t^m - r_{f,t}$, r_t^{smb} , r_t^{hml} and r_t^{mom} are, respectively, the Fama and French (2012)'s regional factors on market, size, value and momentum. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. The sample period is for the US from January 2002 to December 2012, for Europe from January 2004 to December 2012 and for Japan from April 2003 to March 2013.

Overall, the results show that the proposed SRI screen does not seem to have an impact on portfolio performance, independent of weighting method for the best-in-class portfolios. The hypothesis of the SRI screen to perform better when applied to human capital intense industries compared to less human capital intense industries cannot be supported by the regression results. 22 out of 24 alphas are insignificant, which shows the likely non-existent alpha for this constructed SRI screen strategy. Although mostly not significant in statistical terms, the alpha coefficients for the high human capital regressions are mostly more positive than their low human capital counterparts for all regressions in the US and European region.

This is somewhat in line with the hypothesized effect of the SRI screen, which was that the SRI screen should work better when applied to human capital intense industries compared to other industries. Since these coefficients are insignificant and the coefficients for the region of Japan show a mostly reversed pattern, these observations seem to have weak relevance. It therefore looks like the screen does not result in the best-in-class portfolios to outperform their respective industries, as initially expected.

9. Discussion and Implications of Results

Since the results are different to the initially expected outcome it is worth analyzing why this might be the case. There are several reasons that influence the presented results. First of all, it has to be clarified that there could be two main reasons, which in turn might have sub reasons explaining the outcome. The two main reasons are that the proposed hypothesis is correct, but that this study fails to prove it or that the hypothesis is wrong and therefore cannot be proven anyhow.

Assuming that the proposed hypothesis is correct, the failure to provide evidence for it could be the result of some methodological misspecification. In this study, the division of the several companies into the high and low human capital intense industries might be wrong. The descriptive statistics (table 9 in the appendix) of the two types of industry indeed do show some unexpected characteristics. For the US and Europe, the companies defined as low human capital industries in fact show a higher percentage of intangibles in relation to the total asset amount. Usually it would be expected to be the way around, since companies with a high amount of intangibles are more likely to depend on high human capital. For these types of companies, human capital should be one of the key assets to guarantee economic success. The Japanese region also shows in similar way unexpected values for the percentage of intangibles relative to the total assets, although it is less profound and the descriptive statistics for that region seem overall more homogeneous. As the division of the industries forms the base for this study, a misspecification of the industries would have an essential effect on the regression results and most likely nullify any expected effect proposed by the hypothesis. Looking at the payroll per company and industry might be a bad approach for the division of the companies and another approach could lead to more favorable results to strengthen the hypothesis. Instead of using payroll data, the use of years of schooling per industry could result in a more accurate division into high and low human capital industries

as discussed earlier. The availability of data on this matter can, however, represent a constraint. Furthermore a more extreme division of the industry composition might also lead to more favorable results. This means that instead of defining 27 out of the 54 industries as high or low human capital industries, only the top and lower 10 or 20 industries could be defined in this way, while excluding the other industries. A side effect would be a smaller sample size for each region.

Another vulnerable point in the methodology of this study is the overall construction of the score value. The constructed score might fail to assess the companies properly and therefore lead to the wrong identification of companies with favorable employment conditions. This failure of identification then results in constructed portfolios that are not conform with the positive SRI screen idea based on employee satisfaction. The usage of ten variables for the scoring might not fully capture the employee satisfaction in each company. The measure used by Edmans (2011), which is Fortunes' "100 best companies to work for in America"-list, seems to be better at identifying this type of companies at least in the US. The scoring for Fortunes' list is constructed by two thirds of a 57 question survey of the employees and the other third is the result of an evaluation provided by the Great Place to Works® Institute in San Francisco. The fact that the list inclusion is based to a great extent on the answers of the actual employees might play an important role in the lists identification power of employee friendly companies.

For this study, the size of the constructed portfolios is equal to 20 stocks, which is a subjective choice that is based on the idea of a minimum number of stocks necessary to obtain a decent diversification. Since each portfolio is based on either low or high human capital industries this might already affect the diversification characteristics of the portfolios in a negative way, since the exposure tends to be towards a certain type of industry. Meaning that 20 stocks could be too few to have a sufficiently diverse portfolio. Increasing the size of the portfolios might result in more diversified portfolios and increase overall performance of the portfolios. A comparison to the size of portfolios in the study of Edmans (2011) reveals that he uses significantly bigger portfolios with around 50 stocks.

As mentioned above, the market needs to undervalue the employee satisfaction initially in the first place for an investor to profit from an upward stock price adjustment in the long run. This study uses a dataset that is less publicly available than the one used by Edmans and

should therefore increase the likelihood of the undervaluation by the market, which in turn should enhance the portfolio performance even more. Furthermore, in contrast to Edmans approach, the market is not given one month to adjust and price in the information on employee satisfaction. This again should increase the supposedly positive effect on the estimated portfolio performance. It could be, however, that despite the apparently private nature of the information, investors price in the information on employee satisfaction quite accurately. This in turn would decrease the initial undervaluation of the stock by the market. It is discussible as whether to see information provided by Thomson Reuters compared to information publicly available from Fortune as being more private. Although the information provided by Fortune is indeed more publicly available, it does not mean that big investors do not also have access to the Thomson Reuters Asset4 ESG dataset. Especially investors that are focused on making ESG investments and use many SRI screens might actually rely more on a source like Thomson Reuters than Fortune.

Related to the point of the correct pricing of stocks by the market, an aspect that is important in order to assess the results is the sample period. Given that this study uses recent data, the non-existence of significant positive alphas could be explained by a learning effect of investors and the stock market with regard to the incorporation of employee satisfaction as a relevant value driver in companies. Derwall et al. (2011) argue that investors learn from past mistakes. They show with time-varying monthly alphas, using the Carhart four-factor model, that the performance of portfolios based on strong employee relations declines rapidly in more recent years. From 1992-2002 their obtained annual abnormal return is 5.62% and significant at the 10% level, while their abnormal annual returns for the period of 1992-2008 is 2.81% and not significant. This basically means that once investors notice that they mispriced some information in the past that influences the stock price performance, they tend to not repeat these mistakes and incorporate the information better in their stock evaluations in the future. Alphas that were existent in the past might have disappeared through time as the awareness for the basis of these alphas increased. Also Borgers et al. (2013) find similar results supporting the learning effect of investors. They use trading strategies based on a stakeholder-relations index and find these to have generated risk-adjusted returns that were economically and statistically significant in a period from 1992-2004 and mostly not significant in the period of 2004-2009. Since the sample period of this study is from around 2002 until 2013 (depending on the region), the insignificant alphas might be explained by the learning effect of investors.

Instead of assuming that the logic and hypothesis proposed is correct, it could be the case that it is wrong or non-existent. Applying this sort of positive SRI on different types of industries might not result in different or even improved portfolio performance at all. The importance of employee satisfaction or high human capital for a company might be independent of its industry affiliation and is maybe more company specific. It could be the case that for both industry groups (high and low human capital) there are companies that score highly on employee satisfaction, which in turn results in higher motivated employees with increased performance that ultimately might result in better stock performance in the future (given that the market undervalues this fact). As this study is more focused on the SRI screen effect on certain industry portfolios and not the overall economy, the regression results do not necessarily oppose the results obtained in the studies of Edmans. The SRI screen might still enhance portfolio performance when applied to an entire economy portfolio.

The fact that the results obtained in this study do not suggest any regional differences is standing somewhat in contrast to the results obtained by Edmans et al. (2014). They find that the risk-adjusted returns of these types of portfolios depend on how flexible the labor market is in the specific region. The more flexible the labor market is, the more valuable is employee satisfaction as a tool for motivation, retention and recruitment of labor. Given that the US and Japan have more flexible labor markets than Europe, their corresponding alphas for the risk-adjusted returns should be significantly positive. The alphas for Europe should be at least less positive, if not insignificant. Most of the alpha coefficients obtained in this study, however, show negative signs for all three regions. An exact comparison of the results is once again not possible, as this study focuses more on the application of the SRI screen on different industries and not whole markets. Furthermore, it is worth mentioning that the alphas of Edmans et al. (2014) are also only significant for Chile, Japan, Sweden and the US (out of 14 countries) for equal-weighted portfolios. For value-weighted portfolios only Chile, Denmark, the UK and the US show significant alphas. The authors argue that the insignificance of the alphas result from small sample sizes, having an average observation number of 123, when excluding the US data.

The sample size of this study is quite small when comparing to the size of the previously mentioned studies. With 108 observations for Europe, 120 for Japan and 132 for the US, the insignificance of some alphas might be explained by this fact. A look at the coefficients without looking at the significance level might give some further insights (table 6). For the

regions of the US and Europe, the alpha coefficients for the high human capital industry are always higher than for the low ones except for one portfolio, which is the value weighted US high human capital alpha with the risk free rate as benchmark (Low HC: -0.078 vs. High HC: -0.101). This might indicate that the strategy indeed works better for the high human capital industry, as initially suggested. For the Japanese region, however, the pattern is mostly reversed and the constructed portfolios for low human capital industries “outperform” their counterparts. The only exception is the value weighted high human capital alpha with the industry return as benchmark (Low HC: -0.154 vs. High HC: -0.118).

Table 6: Alpha coefficients

	Risk free rate			Industry return		
	Equal-weighted					
	Region	US	Europe	Japan	US	Europe
<i>Low HC</i>	-0.340	-0.274	-0.363	-0.284	-0.208	-0.008
<i>High HC</i>	-0.140	-0.171	-0.743	0.125	0.295	-0.076
	Value-weighted					
<i>Low HC</i>	-0.078	-0.145	-0.107	-0.231	-0.330	-0.154
<i>High HC</i>	-0.101	-0.028	-0.541	-0.203	0.036	-0.118

This table reports an alpha coefficient comparison between the low and high human capital industries. The significance level is not shown.

10. Conclusion

This paper analyzes the effects of a positive SRI screen based on employment quality on two different industry portfolios, namely high and low human capital intense industries. Using a Carhart four-factor model the obtained alphas are mostly insignificant and negative throughout each analyzed region (US, Europe and Japan). The results also seem to be independent of the weighting method for the constructed portfolios (equal- and value-weighted) or the benchmark used to run the regressions (risk-free rate and industry returns). Only 2 out of 24 alphas are significant and negative, instead of being positive as previously hypothesized. The expectation of the positive effect is based on the logic that better employment quality results in higher human capital in a company, which in turn results in higher tangible results in the long run, like increased earnings. Assuming that the stock

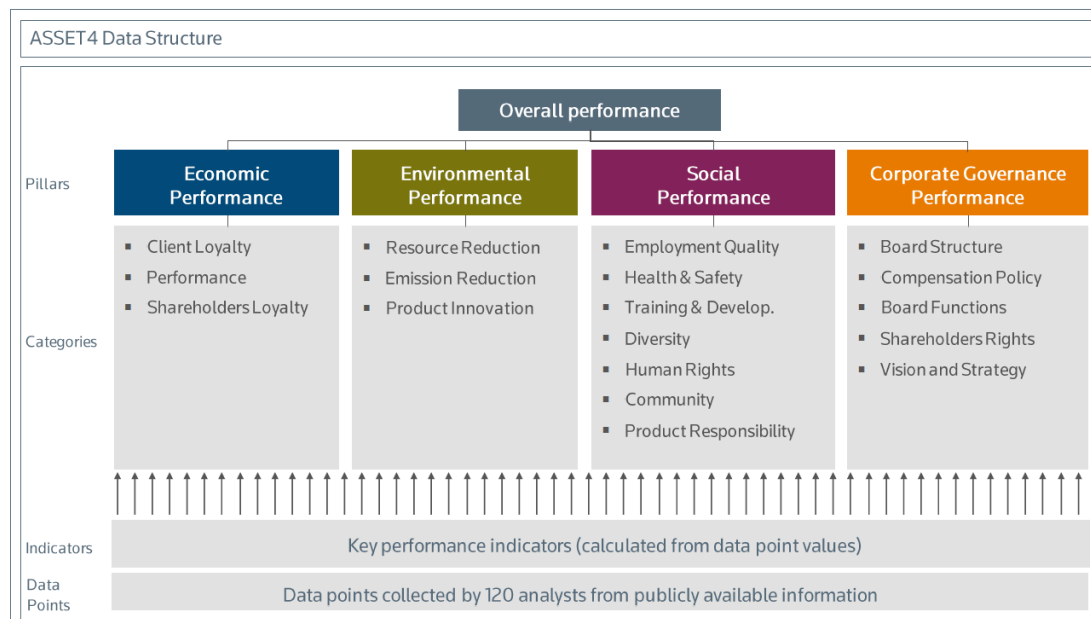
market finds it difficult to value employment quality, an undervaluation in the short run and a correction of this mispricing in the long run would then result in the outperformance from the constructed portfolios. Since human capital is likely to be seen more as a key asset in the human capital intense industry, higher and more significant results were expected from the application of the SRI screen on the high human capital industry portfolio.

Although these results were not expected initially and somewhat not in line with the results obtained by Edmans (2011), they are in line with the idea of a learning effect of investors and the stock market. Investors learn from their mispricing mistakes in the past and can mitigate these mistakes in the future (Derwall et al., 2011; Borgers et al. 2013). Employment quality might still be an important driver of the future performance of a company, but investors apparently have learned over time to use the available information on employment quality more appropriate. This might allow them to make better evaluations of companies, resulting in less mispricing or undervaluation of stocks in the short run. One essential prerequisite for the SRI screen to be profitable could therefore not be given anymore, which makes it impossible to benefit from an SRI screen based on employment quality in financial terms. Since the attention given to SRI increased significantly in the past, this seems to be a plausible reason explaining the non-existence of significant positive alphas.

This research has certain limitations, which could serve as basis for further research in the area of SRI applications on specific industries. The division of industries in this study is based on the payroll per company and industry. A different approach to the division of the companies into the high and low human capital industries could lead to different results. Years of schooling as a criterion might be an alternative approach. The score constructed in this research might not capture accurately the employment quality in the companies. The usage of a different scoring method or overall different database could shed some light on this issue. Although the research covers parts of three main continents, the study could also be extended in geographical terms. Lastly, the portfolio size is subjectively chosen. It could be the case that the size of the portfolio (20 companies) is not sufficient to have enough diversification. Choosing bigger portfolios might diminish this risk and lead to different results. It is also worth mentioning that with regard to an implementation of this strategy in the real world, transaction costs would have to be considered, which would worsen the overall results.

Appendix

Table 7: ASSET4 ESG data structure



Source: Thomson Reuters website

Table 8: Scoring components

Subcategory	Factor Description	Type of data	ASSET4 ESG code
Employment Quality	Average salaries and benefits in US dollars (Salaries and Benefits (US dollars) /Total Number of Employees).	Numerical values transformed in z-scores	SOEQO01S
	Does the company claim to provide a bonus plan to most employees?	Y / N transformed in z-scores	SOEQDP0201
	Does the company have a competitive employee benefits policy or ensuring good employee relations within its supply chain? AND Does the company have a policy for maintaining long-term employment growth and stability?	NN / NY/ YN / YY transformed in z-scores	SOEQD01S
	Has the company won an award or any prize related to general employment quality or "Best Company to Work For"?	Y / N transformed in z-scores	SOEQO05S
	Does the company claim to provide its employees with a pension fund, health care or other insurances?	Y / N transformed in z-scores	SOEQDP025
	Total salaries and benefits divided by net sales or revenue.	Numerical values transformed in z-scores	SOEQO02S
Health & Safety	Does the company have a policy to improve employee health & safety within the company and its supply chain?	Y / N transformed in z-scores	SOHSD01S
Training & Development	Does the company claim to provide regular staff and business management training for its managers?	Y / N transformed in z-scores	SOTDDP024
	Does the company have a policy to support the skills training or career development of its employees?	Y / N transformed in z-scores	SOTDD01S
	Does the company claim to favor promotion from within?	Y / N transformed in z-scores	SOTDDP023

Table 9: Descriptive statistics of industries per region

Industry	Obs.		Mean		Median		St. Dev.		Min.		Max.		
	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	
US	Market Cap. (in bn. USD)	441	702	10.48	13.44	3.84	4.45	24.06	29.13	0.16	0.08	272.27	363.07
	Price (in USD)	441	702	33.11	184.10	27.94	29.44	32.29	3,815.75	2.80	0.50	584.47	100,185
	Price/Book	441	702	4.90	2.96	2.91	2.34	20.04	6.04	-124.08	-61.30	301.45	76.09
	Intangibles % of tot. assets	441	702	25.9%	17.4%	22.9%	9.3%	20.3%	19.5%	0.0%	0.0%	88.5%	81.7%
Europe	Market Cap. (in bn. USD)	206	240	9.87	13.89	4.31	4.97	14.65	23.98	0.45	0.16	96.41	209.11
	Price (in USD)	206	240	37.23	33.47	28.06	22.88	34.65	33.86	0.82	0.52	297.46	212.34
	Price/Book	206	240	2.93	2.13	2.01	1.67	3.47	1.60	-4.30	-1.94	26.76	12.35
	Intangibles % of tot. assets	206	240	23.0%	15.9%	19.7%	9.2%	18.0%	17.8%	0.0%	0.0%	69.7%	75.8%
Japan	Market Cap. (in bn. USD)	206	191	7.08	7.03	3.14	3.83	12.63	9.09	0.66	0.79	135.72	76.15
	Price (in USD)	206	191	24.88	119.06	12.72	11.52	50.86	688.26	1.94	1.43	605.30	6,434.27
	Price/Book	206	191	1.59	1.82	1.42	1.38	1.28	2.00	-7.31	-2.21	6.53	20.99
	Intangibles % of tot. assets	206	191	3.6%	3.4%	1.9%	1.3%	5.0%	5.4%	0.0%	0.0%	37.5%	30.5%

Table 10: Weighting of payroll data for average payroll calculation per company

Weighting method	Weight given to different SIC codes							
	SIC 1	SIC 2	SIC 3	SIC 4	SIC 5	SIC 6	SIC 7	SIC 8
1	4	1	-	-	-	-	-	-
2	8	7	6	5	4	3	2	1
3	15	10	7.5	5	4	3	2	1
4	1	0	0	0	0	0	0	0

References

- Aboody, D., & Lev, B. (1998). The value relevance of intangibles: the case of software capitalization. *Journal of Accounting Research*, 161-191.
- Abowd, J. M. (1989). The effect of wage bargains on the stock market value of the firm. *The American Economic Review*, 774-800.
- Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics*, 543-569.
- Barro, R. J. (1991). *Economic growth in a cross section of countries* (No. w3120). National Bureau of Economic Research.
- Barro, R. J., & Lee, J. W. (1996). International measures of schooling years and schooling quality. *The American Economic Review*, 218-223.
- Bauer, R., Derwall, J., & Otten, R. (2007). The ethical mutual fund performance debate: New evidence from Canada. *Journal of Business Ethics*, 70(2), 111-124.
- Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7), 1751-1767.
- Borgers, A., Derwall, J., Koedijk, K., & Ter Horst, J. (2013). Stakeholder relations and stock returns: on errors in investors' expectations and learning. *Journal of Empirical Finance*, 22, 159-175.
- Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. *Financial Management*, 35(3), 97-116.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.
- Chan, L. K., Lakonishok, J., & Sougiannis, T. (2001). The stock market valuation of research and development expenditures. *The Journal of Finance*, 56(6), 2431-2456.
- Deng, Z., Lev, B., & Narin, F. (1999). Science and technology as predictors of stock performance. *Financial Analysts Journal*, 20-32.
- Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The eco-efficiency premium puzzle. *Financial Analysts Journal*, 51-63.
- Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of Banking & Finance*, 35(8), 2137-2147.

- Eccles, R. G., Ioannou, I., & Serafeim, G. (2013). The impact of corporate sustainability on organizational processes and performance. *Management Science*.
- Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics*, 101(3), 621-640.
- Edmans, A., Li, L., & Zhang, C. (2014). *Employee Satisfaction, Labor Market Flexibility, and Stock Returns Around The World* (No. w20300). National Bureau of Economic Research.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56.
- Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105(3), 457-472.
- Geczy, C. C., Stambaugh, R. F., & Levin, D. (2003). *Investing in socially responsible mutual funds*. Rodney L. White Center for Financial Research, The Wharton School, University of Pennsylvania.
- Gorton, G., & Schmid, F. A. (2004). Capital, labor, and the firm: A study of German codetermination. *Journal of the European Economic Association*, 2(5), 863-905.
- Guerard Jr, J. B. (1997). Is there a cost to being socially responsible in investing?. *The Journal of Investing*, 6(2), 11-18.
- Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal*, 62-66.
- Hillman, A. J., & Keim, G. D. (2001). Shareholder value, stakeholder management, and social issues: what's the bottom line?. *Strategic management journal*, 22(2), 125-139.
- Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1), 15-36.
- Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1), 65-91.
- Jenter, D., Lewellen, K., & Warner, J. B. (2011). Security issue timing: What do managers know, and when do they know it?. *The Journal of Finance*, 66(2), 413-443.
- Joshi, A., & Hanssens, D. M. (2010). The direct and indirect effects of advertising spending on firm value. *Journal of Marketing*, 74(1), 20-33.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological review*, 50(4), 370.
- McGregor, D. (1960). The human side of enterprise. *New York*, 21, 166.

- Mizik, N., & Jacobson, R. (2003). Trading off between value creation and value appropriation: The financial implications of shifts in strategic emphasis. *Journal of Marketing*, 67(1), 63-76.
- Moskowitz, M. (1972). Choosing socially responsible stocks. *Business and Society Review*, 1(1), 71-75.
- Mulligan, C. B., & Sala-i-Martin, X. (2000). Measuring aggregate human capital. *Journal of Economic Growth*, 5(3), 215-252.
- Plyakha, Y., Uppal, R., & Vilkov, G. (2014). Equal or Value Weighting? Implications for Asset-Pricing Tests. *Implications for Asset-Pricing Tests (January 15, 2014)*.
- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3), 302-322.
- Schröder, M. (2007). Is there a difference? The performance characteristics of SRI equity indices. *Journal of Business Finance & Accounting*, 34(1-2), 331-348.
- Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review*, 433-444.
- SIF, 2012. Report on Sustainable and Responsible Investing Trends in the United States 2012. The Forum for Sustainable and Responsible Investment. <<http://www.ussif.org/>>.
- Statman, M., & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, 33-46.
- Tanfous, M. B. (2013). The Contribution of Intangibles to the Value Creation. *Journal of Business Studies Quarterly*, 5(1), 32.
- Taylor, F. W. (1911). The principles of scientific management. *New York*, 202.
- Zingales, L. (2000). In search of new foundations. *The journal of Finance*, 55(4), 1623-1653.