ESG investing: a literature review

Søren Hvidkjær
Professor of Finance

September 2017

Report prepared for Dansif

Please address any correspondence to sh.fi@cbs.dk. I thank Lotte Jensen and Søren Larsen for their comment and suggestions.

1 Introduction

A fundamental question to investing based on Environmental, Social and Governance (ESG) considerations is how does such investing affect the value of the investor's portfolio? This report reviews the current evidence on this question. Specifically, how does ESG information affect the risk-return characteristics of an investor's portfolio, hence the formation of the optimal portfolio?[1]

The findings of the reviewed empirical studies of ESG investing can be summarized as follows:

1. Considerable evidence exists that so-called sin stocks exhibit outperformance relative to various benchmarks.
2. There is evidence that stock with high ESG ratings exhibit high future returns. The evidence is strongest in 1991-2004, while the returns of stocks with high ESG ratings do not appear to differ from benchmarks in 2005-2012. Some evidence suggests that returns again have been high since 2012.
    a. The evidence on investor returns to environmental screens is limited and the results are mixed.
    b. Investor returns to at least one social screen, namely employee satisfaction, were high during 1984-2011.
    c. Good governance firms as measured by the G-index had higher returns than poor governance firms in 1990-1999. However, the return difference disappeared in the subsequent period. Some evidence suggests that other measures of governance predict returns in the subsequent period, but the evidence is not conclusive.
3. Event studies indicate that the stock market does not respond positively to certain types of ESG/CSR initiatives taken by firms. While the results suggest that agency issues are a genuine concern, they also suggest that such concerns can be mitigated through sound corporate governance.
4. Active ownership by ESG investors can create value, both for shareholders and other stakeholders.

To appreciate the significance of the above results, it is central first to develop an understanding of the possible channels through which a signal can impact returns. Without such an understanding, the investor will have no way of developing an informed opinion of whether a useful signal in the past will hold in

[1] The review will focus narrowly on the effect of following ESG strategies on portfolio value. Other aspects include ethical, political and marketing considerations. While these aspects clearly can be important for an investor, they are outside the scope of the current paper.

the future, or whether new signals are likely to become relevant. Therefore, this report will begin with a discussion of the theoretical aspects of ESG investing. The report then provides a brief overview of the empirical literature on ESG and the types of questions addressed. Section 2 gives a detailed review of key articles in sin-stock investing, positive/negative ESG screening, event studies, active ownership, and meta studies. Section 3 concludes and provides some perspectives on the consequences of the results for investors.

1.1 Theoretical considerations to ESG investing

Plausible reasons exist for both outperformance and underperformance of ESG investing relative to conventional investing. In the following, we review the arguments as they pertain to an investor who (1) does not possess inside information about firm values and (2) does not engage in active ownership by directly exerting influence on management.[2] For such an investor, the central question to ESG investors is not whether ESG initiatives by firms create value, but whether any such value is properly recognized by the stock market.

In general, the main argument for **outperformance** of based ESG-strategies is, in essence, that the stock market underreacts to ESG information. That is, the value effects of a positive ESG event is not sufficiently recognized by the stock market, hence firms with such events tend to be undervalued and a strategy investing in these firms can obtain abnormally high returns.

The underreaction hypothesis is plausible, given that evidence exists of stock market underreaction in various situations. In particular, post-earnings announcement drift (Ball and Brown, 1968; and Bernard and Thomas, 1989) and momentum (Jegadeesh and Titman, 1993) are both among the most robust evidence against market efficiency and are consistent with market underreaction. Moreover, a reasonable hypothesis is that the stock market undervalues certain intangibles. The valuation of intangibles is typically more uncertain than tangibles and often intangibles do not appear directly on the balance sheet, hence they are less salient to investors. Evidence of underreaction to intangibles includes R&D costs, patent citations, advertising, and software development costs (see references p. 622 in Edmans, 2011). Likewise, ESG investments by firms are typically intangibles, and it is possible that the stock market underreact to the information in ESG-related initiatives.

A second reason for why high ESG stocks might outperform the market (and low ESG stocks) is that ESG investing has become more popular over time with investors. That is, a growing demand for a particular set of stocks can push up

[2] Section 1.1.1 discusses the considerations to active ownership.

the prices of those stocks, even in the absence of new fundamental information about the value of those stocks.

Demand effects are also a primary reason for why high ESG stocks might exhibit **underperformance** relative to low ESG stocks. Merton (1987) points out that when a large group of investors ignore certain stocks, say low ESG stocks, they can become undervalued. While this implies initial low returns, subsequently those stocks will have high returns relative to high ESG stocks. Even if the undervaluation is permanent, then a low stock price implies a high dividend/price ratio, hence higher returns, ceteris paribus.[3]

Also, firms in industries often shunned by ESG investors, such as tobacco and weapons industries, have incentives to practice very conservative accounting because their industries fall under considerable scrutiny from regulators (Berman, 2002; Hong and Kaperczyk, 2009). To the extent that investors do not account for this, it will lead to underreaction and hence subsequent high returns.

The argument for pricing effects of the ignored stock does not take firms' response into account. Low ESG firms may respond to falling investor interest, and thus lower stock price, by changing behavior. Heinkel, Kraus and Zechner (2001) analyze such a situation in a theoretical model. If firms can obtain a lower cost-of-capital by attracting more ESG-conscious investors by changing their behavior, then the effect on the cost-of-capital, hence on expected returns, is alleviated. Such actions could include making more environmentally friendly investments or improving employees' working conditions.

Almost perfectly elastic demand curves will lead to **no under- or overperformance** of high ESG stocks. ESG investing might not be sufficiently widespread to affect prices. Alternatively, arbitrageurs could offset any effect of ESG investors by trading in the opposite direction. Still, with costly arbitrage we would not expect any such fully offsetting effect, as arbitrageurs need to be compensated for their efforts (Grossman and Stiglitz, 1980).

However, if even ESG investing has no impact on prices, the individual investor might still experience negative financial effects of ESG investing for two reasons. First, a central tenet of finance is that diversification provides risk reduction without a reduction in expected returns (Markowitz, 1959). Thus, investment in a broad portfolio of assets provides the optimal risk-return trade-off, and any restriction in the investable universe leads to a worse trade-off. In the absence of any pricing effects, one would expect such lack-of-diversification effects would be largest if either entire industries are excluded (as opposed to excluding the worst ESG performers within an industry).

[3] A reinforcing effect comes from the breakdown of CAPM in segmented markets, hence the price of ignored stocks can be further depressed if they have high idiosyncratic volatility.

Secondly, low returns to ESG investing in the absence of price effects might occur if there are high costs to ESG screening. This is an especially pertinent challenge for passive, low-cost investors. Because of the diversification argument along with the existence of nearly informationally efficient market, an important lesson from finance is that most investors should pursue a passive, low-cost strategy. This entails obtaining the cheapest possible exposure to an asset class, but ESG requirements can be incompatible with such a goal because it involves selecting individual stocks.

In longer run, assuming that ESG investing will reach a high permanent level of penetration among investors, it is difficult to see how any outperformance could be sustained, for three reasons. First, the underreaction to intangible ESG information would disappear as many investors pursue strategies based on such information. This is similar to any other strategy based on other investors neglecting value-relevant information. Secondly, the popularity argument is based on growth in demand, thus temporary by nature. Thirdly, the Merton argument of ignored stocks become more relevant, the larger the group of investors is, who pursue ESG strategies. That is, the higher the level of ESG strategies among investors, the larger is the likely underperformance.

Thus, there are two pertinent questions for today's ESG investors. First, how close are we to a steady-state level of ESG investing? Secondly, how important is the Merton argument for stock prices, hence for future stock returns?

1.1.1 Active ownership and the interaction between ES and G

The above discussion assumes that investors are active investors, but passive owners. By contrast, active owners (also called activist investors) directly engage with management to change decisions regarding, say, matters related to ESG. Such activities allow for directly impacting firm value rather than seeking to identify undervalued firms. One possible channel for value creation is alleviating managerial myopia. That is, principal-agent issues such as managerial career concerns might lead managers to become focused on short-term results at the expense of long-term value creation. Active engagement by investors could thus allow management to take a long-term perspective.

ESG investing is linked to Corporate Social Responsibility (CSR), which in turn is linked to the debate of shareholder value vs. stakeholder welfare. Often, stakeholder management is fully consistent with the shareholder criterion. As Jensen (2001), a prominent critic of CSR, notes "we cannot maximize the long-term market value of an organization if we ignore or mistreat any important constituency (stakeholder)". However, there are clear principal-agent problems in CSR, as the manager may engage in CSR for personal benefit or driven by social preferences rather than maximizing shareholder value. As Tirole (2001)

notes, "management can almost always rationalize any action by invoking its impact on the welfare of *some* stakeholder. An empire builder can justify a costly acquisition by a claim that the purchase will save a couple of jobs in the acquired firm; a manager can choose his brother-in-law as supplier on the grounds that the latter's production process is environmentally friendly".[4] Such problems indicate that good corporate governance is essential in aligning CSR with maximization of shareholder value.

In short, principal-agent issues can cause managers exhibit myopia, thus not investing in value-creating ESG initiatives. On the other hand, principal-agent issues can also lead to value-destroying ESG investments. Therefore, G investing and active ownership ensuring good corporate governance arguably forms the basis of successful ES investing.

1.2 The empirical literature on ESG investing and the scope of the review

The literature ESG investing is part of a broader literature on how Corporate Social Responsibility relates to Corporate Financial Performance. This literature can be split into 4 categories, according to the methodology applied and questions addressed.

First, is the question of whether investors can form portfolios based on ESG signals that contain information affecting the risk-return characteristics of their portfolio, hence the formation of the optimal portfolio. The standard procedure is thus to identify a signal, say an ESG rating, and then construct a portfolio based on the signal in order to back-test whether it contains valuable information. Most of the articles reviewed in the current paper use this approach, as it is the most direct test of the relevance of ESG issues to the values of investors' portfolios.

Instead of forming portfolios based on individual stocks, some researchers study the performance of SRI mutual funds relative to conventional funds (or other benchmarks). A confounding factor here is that management fees affect results. Such fees can be both direct and indirect; hence it can be difficult to link the results to the underlying stock returns. Relatedly, portfolio manager skills differ which in turn necessitates making assumptions about the distribution of skills. Further, the performance-flow relationship makes it hard to establish systematic differences in performance over time.[5] Proponents of studies including mutual funds argue that such portfolios represent investable portfolios whereas one might in practice not be able to implement studies using individual stock portfolios. However, this is not an inherent flaw of studies based on individual

[4] The quotes from Jensen and Tirole are also reproduced in Renneboog, Horst and Zhang (2008), which contains a good review of the early literature on CSR/SRI/ESG investing.
[5] That is, successful funds tend to receive additional money from investors, but such additional money is often harder to invest at similarly high expected returns.

stock returns. Rather, the researcher can control investability via the weighting scheme of the portfolios.

A second strand of the literature uses event study methodology to study the stock market response to ESG-related news. This addresses how the stock market perceives the value of specific ESG initiatives; say the voluntary reduction in the emission of green house gasses. The evidence from event studies are reviewed in section 2.3 along with a discussion of advantages and drawbacks of such studies.

A third strand studies the relationship between CSR/ESG and accounting-based performance measures, such as ROA and ROE. Here, a researcher might find that high CSR firms are more profitable and conclude that CSR initiatives create value for shareholders. However, it is notoriously difficult to infer causality from correlations between different corporate variables. Indeed, as Roberts and Whited (2013) note, "arguably, the most important and pervasive issue confronting studies in empirical corporate finance is endogeneity". A firm with a high ESG score might have high profitability, but the high profitability might be driving the ability to invest in ESG rather than the ESG investments causing high profitability. To avoid such issues of reverse causality, one needs to identify exogenous variation in the ESG variable, rather than simply showing a correlation between the two variables. However, the causality issue is usually not properly addressed in the CSR literature.[6]

Moreover, finding that sound ESG decisions are sound business decisions does not imply that investors obtain superior returns from investing in these firms. The case for investors depends on whether the information is already priced into the stock. Only if the stock market systematically undervalues such information will the ESG investor obtain high returns.

Because of the difficulty of establishing causality and the lack of direct implications for investors, this review will not cover studies linking CSR and accounting performance.

A fourth strand of the literature attempts to obtain ex ante measures of firms' cost of equity capital. Absent frictions, equity investors' required returns are equal to the firms' cost of equity capital, thus past average returns are ex post measures of firms' cost of capital. Ex ante measures, by contrast, are usually computed from analysts' earnings estimates. The evidence indicates that high

[6] An exception is Flammer (2014) who uses a regression discontinuity design to show that narrowly passed CSR shareholder proposals leads to positive accounting performance. By contrast, Hong, Kubik and Scheinkman (2012) identify exogenous variation in financial constraints to show that less financially constrained firms engage in more CSR. Thus, the evidence suggests that causality can run in both directions, but more research is needed before making firmer conclusions.

ESG firms obtain a lower cost of capital (see, e.g., Chava, 2014), suggesting that investors are willing to accept lower returns when investing in high ESG firms. However, as with accounting measures, ex ante cost of capital do not directly measure investor returns, hence this review will not focus on these studies.

1.3 Some general observations on the articles reviewed

A few general observations can be made of the studies in this review relative to most other empirical asset pricing studies.

Many studies base inferences on a **very short time period** during which returns are measured. Several studies employ time spans of less than 10 years. This leads to at least two potentially severe problems. First, generally the asset pricing tests will have low statistical power. Test power is the probability that the test correctly rejects the null hypothesis when the alternative hypothesis is true. In the current context, the null hypothesis is that there is no difference in average returns between two sets of stocks, say high ESG and low ESG stocks, while the alternative hypothesis is that a return difference exists. Therefore, the low power becomes an issue when researchers are not able to reject the null of no return differences. The appropriate conclusion is then simply that one cannot reject the null. However, some researchers appear to confuse absence of evidence with evidence of absence. That is, even if average return differences of 3-4% per year are reported, some authors claim that no differences exist rather than they were not able to reject the null of no differences.

The second challenge when using short time spans is that results become sensitive to particular developments during the period. For instance, oil prices might have exhibited a particular trend up or down, or the economy might have been in an expansionary stage during the entire sample period. Statistical techniques are not necessarily able to fully account for such realizations of underlying factors.

The question of differences in returns between two sets of stocks uses methodology developed within the research area of finance. However, many of the articles discussed below are published in journals with otherwise little finance content. While many of the articles published in non-finance journals appear to be methodologically sound, a couple of differences exist relative to articles published in good finance journals. First, the econometric methods employed are usually standard, but more sophisticated methods are rarely employed. For instance, the standard method in controlling for any size and book-to-market effects is performing Fama and French (1993) 3-factor regressions. However, using characteristic-adjusted returns as in Edmans (2011) allows for interaction effects that would not be captured by 3-factor regressions, but this method has not been applied in articles in non-finance journals. Also, the

application of the methods lacks rigor in some instances. Secondly, some articles are simply not written in a neutral, objective language. One hopes that this does not reflect a lack of objectivity in the analysis.

To be clear, the majority of studies apply sound methods. However, because some studies do appear to exhibit biases, a thorough reading and independent interpretation of the results has been necessary, rather than relying on the authors' presentation.

2 Detailed review of individual articles

In this section, we will first review the literature on sin stock returns to evaluate the effect of sector exclusion of ESG investing. Then, we examine the evidence of using ESG ratings with negative and positive screening. The effects of general ESG screening is first examined, and results for E, S and G are examined in turn. In section 2.3, studies on stock market reactions to ESG events are reviewed. Section 2.4 discussed the results from a recent study of the financial effects of active ownership. Finally, in section 2.5 we discuss a recent meta study.

While the article reviews are based on careful readings, there is also a value for the reader in recognizing the quality of the publishing journal, as a wide dispersion exists in the quality of academic journals. Articles in higher ranking journals are usually of a higher quality, in part because of the rigorous referee process and in part because of the social construction aspect (they are perceived as high quality, thus more desirable outlets for authors). Within finance, the undisputed top journals are *the Journal of Finance, the Review of Financial Studies* and *the Journal of Financial Economics*. While inferences based on results in these journals tend to have higher validity than those from lower ranking journals, articles in lower-tier journals may simply address more narrow questions.

The Association of Business School's ranking of journals is used as a rough indication of journal quality. Table 2.1 replicates the ratings explanation from the ABS Academic Journal Guide 2015. In short, 4* journals are the leading journals while 1-rated journals are of the lowest perceived quality among rated journals. Hence, the ABS rating of each article is noted below. Since not all academic journals are rated, non-rated journals will be indicated by "ABS 0" below. ABS 0 journals are typically of a lower standard than ABS 1 journals.

Moreover, journals are rated within their particular field of study, but articles may span different fields. This is particularly true for studies on EGS investing. This reviewer finds that the finance aspects of the studies tend to be of higher quality when published in finance journals. Therefore, all citations also indicate whether the article is published in a finance (F), non-finance (NF) or partly finance (F/NF) journal.

The first reference to each article will thus be of the format *author names (year published, abbreviated journal title, ABS ranking, F or NF).*

Rating	Meaning of Quality Rating
4*	Journals of Distinction. Within the business and management field including economics, there are a small number of grade 4 journals that are recognised world-wide as exemplars of excellence. Their high status is acknowledged by their inclusion in a number of well-regarded international journal quality lists. The Guide normally rates a journal 4* if they are rated in the highest category by at least three out of the five non-university based listings – Financial Times 45, Dallas List, VHB, Australian Deans' List, CNRS. In addition, journals from core social sciences disciplines that do not appear in those listings may also be rated 4* on the grounds that they are clearly of the finest quality and of undisputed relevance to business and management. In the Guide of 2015, this applies to three journals from the fields of sociology and psychology.
4	All journals rated 4, whether included in the Journal of Distinction category or not, publish the most original and best-executed research. As top journals in their field, these journals typically have high submission and low acceptance rates. Papers are heavily refereed. Top journals generally have the highest citation impact factors within their field.
3	3 rated journals publish original and well executed research papers and are highly regarded. These journals typically have good submission rates and are very selective in what they publish. Papers are heavily refereed. Highly regarded journals generally have good to excellent journal metrics relative to others in their field, although at present not all journals in this category carry a citation impact factor.
2	Journals in this category publish original research of an acceptable standard. A well regarded journal in its field, papers are fully refereed according to accepted standards and conventions. Citation impact factors are somewhat more modest in certain cases. Many excellent practitioner-oriented articles are published in 2-rated journals.
1	These journals, in general, publish research of a recognised, but more modest standard in their field. Papers are in many instances refereed relatively lightly according to accepted conventions. Few journals in this category carry a citation impact factor.

Table 2.1

Source: ABS Academic Journal Guide 2015

The review will generally focus on articles published within the last decade. Earlier work might have made more fundamental contributions, but later work building on such contributions will usually include the original sample, thus contain the information in early work.

2.1 The sin stock evidence

Considerable evidence exists that so-called sin stocks exhibit outperformance relative to various benchmarks.

Hong and Kacperczyk (2009, JFE, ABS 4* F) is arguably the most prominent and most-cited article on the return effects of negative screening. Analyzing U.S. stocks, they find that the so-called sin stocks (defined in the paper as tobacco, alcohol and gambling firms) are held by relatively few institutional investors and followed less by financial analysts relative to a control group of stocks. This is consistent with the clientele hypothesis that social norms lead some investors to shun the sin stocks. Using the Merton (1987) arguments that stocks neglected by a large segment of investors will tend to have depressed prices, hence higher future returns, HK investigate the returns of sin stocks. They find that sin stocks outperform comparable by 3-4% per year.

One of the key strengths of the article is that the sample covers a long time span, namely 1926-2006 (although the sample period is shorter for institutional ownership and analyst coverage). This allows for more powerful tests than usually employed in the ESG literature. Still, the results do not appear to be very robust. For instance, the institutional ownership and analyst coverage results are not robust to inclusion of the market-to-book ratio as a control variable (tables 3A and 3C). The return tests are standard Fama-French (1993) factor regressions, and while alphas are significant in all specifications (1, 2, 3 and 4 factors), it is only significant at the 10% level in the standard 3-factor model in the 1965-2006 period (table 4A).[7] Moreover, the excess returns are defined relative to the returns of a COMP portfolio consisting of comparable stocks that belong to the Fama and French (1997) industry groups 2 (food), 3 (soda), 7 (fun), and 43 (meals and hotels). Therefore, the interpretation of the results is that sin stocks outperform a portfolio of food, soda, fun and, meals and hotel stocks. This is reasonable, but one might also ask whether alphas are significant when excess returns over the risk-free rate or a group of characteristics-matched stocks are used (see, e.g., Edmans, table 3, 2011). The Fama-MacBeth regressions in table 4B indicate that results are only significant when adding the control group.

HK are heavily critized by **Hoepner and Zeume** (2013) and **Adamsson and Hoepner** (2015). AH note on p. 4 that "*HK regress an equal- weighted portfolio of sin stocks on a value-weighted market benchmark. This implies that the outperformance could be driven by a small cap performance bias rather than sin stocks characteristics […]. This argument is founded on the empirical observation that small stocks outperform large stocks […]. The exceptionally good performance could hence be due to an over-weighting of small cap stocks and underweighting of*

[7] The Fama-French model is the most widely used asset pricing model for providing benchmark returns in empirical studies. The 3-factor model controls for returns driven by exposure to the market factor, a firm-size factor and a book-to-market factor. The 4-factor model also includes momentum. Alpha is the part of the average return that is not explained by the asset pricing model; hence alpha is the abnormal return.

large cap stocks." This argument, however, neglects the fact that HK use long-short portfolios; hence any bias towards small stocks is likely netted out by the short side. Indeed, the SIN-COMP portfolio in HK loads negatively on SMB in all specifications in table 4A. This indicates that the stocks in the SIN portfolio are larger than those of the COMP portfolio which is opposite of the AH claim.

This is not to imply that that equal- vs. value-weighting does not matter, but rather that while the effect might be greater among small stocks than among large stocks, it is not a manifestation of the small stock premium.

Finally, AH claim that the HK sin results disappear when using value-weighted returns. However, AH use a much shorter sample, namely 2002-2013. It is close to meaningless to use a 12-year sample to make such inferences. If one wishes to argue that the HK results disappear when value-weighting, one must use the same sample as HK.

Fabozzi, Ma and Oliphant (2008, JPM, ABS 2 F) show that sin stock exhibit high returns during 1970-2007 in several international markets. They analyze the returns of 267 stocks in alcohol, tobacco, biotech, defense, and adult entertainment (table 2) in 21 countries. The results are based on a one-factor model in which each stock return is adjusted by the corresponding national index, and the portfolio return is the equally weighted excess returns. The returns are highly significant both for the composite sin portfolio (table 3) and for each of the 6 industry portfolios (table 4). The sin returns are also economically significant with roughly 1% per month more than the market return. While results are robust with respect to the definition of sin, we do not know the robustness with respect to value-weighting or to alternative asset pricing models. Still, given the very high abnormal returns, it does not seem likely that a standard asset pricing model, such as the Fama-French model, would explain the high returns.

Trinks and Scholtens (2017, JBusEth, ABS 3 NF) show that sin stock exhibit high returns during 1991-2012 in several international markets. They select at the individual stock level (rather than excluding industries) and employ a broad definition of sin with 14 different issues, including meat and contraceptives. The sample is large, consisting of 1,634 stocks across 94 countries.

Tables 3 and 4 show strong outperformance: the value-weighted TotalSin portfolio outperforms the global FF-4 factor benchmark by 91-104 bp. per month – significant at the 1% level. Among the individual sin issues, tobacco has the strongest abnormal returns of 166 bp. per month.

While these results appear to be very strong, they also indicate that the global FF-4 factor model does not provide a reasonable benchmark. In table 3, results

are reported for the zero-investment portfolio of TotalSin minus the FF global market factor. This is an unusual procedure as the market factor appears on both sides of the regression. Indeed, the market beta of the zero-investment portfolio is -0.7; but this implies that the TotalSin portfolio has a beta of only 0.3. The low beta suggests that the FF global factor does not sufficiently capture average returns. Using the a priori expected beta of 1 and the average global factor return over the period of 0.44% per month (obtained from Ken French' data library), the abnormal return would be lower by around 31 bp. per month (-0.7x44). In table 3, the outperformance would then be 60 bp. per month. This is still quite large, but given the large standard errors, unlikely to be significant. Still, the large standard errors are another manifestation of the poor fit of the FF-4 factor model. Overall, this illustrates the difficulty of simply applying a global factor model to stocks listed across the globe. The approach of Fabozzi, Ma and Oliphant (2008) of using national market returns appears to provide superior benchmarks.

In sum, the results in Trinks and Scholtens (2017) suggest that returns of sin stocks are very high. However, results should be interpreted with caution given the lack of a good benchmark model.

Perhaps the most interesting and revealing study of exclusionary screening is **Hoepner and Schopohl** (2016, JBusEth, ABS 3 NF). HS study the performance of stocks excluded from the Swedish AP-funds and the Norwegian Government Pension Fund-Global (GPFG) during the period 2001-2015. The screening primarily norm-based rather than sector-based. That is, it is mostly conducted based on violations of international norms regarding environmental, human rights or labor rights issues or the production of controversial weapons. A sector-based screening is performed only for the tobacco industry for the GPFG in 2009-2015.

HS note in the abstract that the "portfolios of excluded companies do not generate an abnormal return relative to the funds' benchmark index". However, the results of the paper do not support this claim. Table 3 shows the main results for the six portfolios analyzed, namely the value- and equal-weighted portfolios of the excluded stocks in AP7, AP1-4 and GPFG. Panel A shows that the CAPM alphas for the six portfolios are *uniformly positive*, and the one of the alphas is significant at the 1% level while further two alphas are significant at the 10% level. Moreover, among the insignificant alphas, the smallest value corresponds to an abnormal return of 4.15% per year of the excluded stocks (for AP1-4 value-weighted). While one cannot reject with 90% confidence the hypothesis that this portfolio does not generate abnormal returns, calculations show that one can also not with 90% confidence reject the hypothesis that the abnormal returns are different from 10% per year. Indeed, the standard errors of the alpha

estimates for the AP1-4 portfolios are roughly twice as large as those of the AP7 portfolios, while the point estimates are similar between AP1-4 and AP7. Hence, we are not surprised that results for AP7 are significant while they are insignificant for AP1-4. Two issues seem to drive the high standard errors for AP1-4. First, the sample period is longer as AP7 covers nearly 14 years while AP1-4 covers just 9 years and 1 month. Secondly, the AP1-4 portfolio is very small compassing between 2 and 20 stocks while the AP7 portfolio includes between 19 and 54 stocks. Indeed, table 3 shows that the market factor explains much less of the time series variation in AP1-4 than in AP7. In short, the power of the test for the AP1-4 stocks is much lower than for AP7 stocks, hence we are not surprised to find significant results for AP7, but not for AP1-4 - even when the alpha estimates are very similar. Moreover, the standard errors of the GPFG portfolios are in-between, and indeed the level of significance is in-between those of AP1-4 and AP7.[8]

The differences in power across the 3 portfolios exist *by construction*: arising from the way that the authors construct the portfolios. It makes no sense to treat these portfolios equally when subsequently making inferences.

Most importantly, the key test with maximum power is conspicuously absent. One can easily construct the portfolio consisting of all excluded stocks across the three sub-portfolios. The abnormal returns for the equally weighted portfolio would be around 5% per year and that of the value weighted portfolio would be around 4% per year. Because of the added power, the alpha estimates for both portfolios are highly likely to be statistically significant.

Other studies that find outperformance of sin stocks include Statman and Glushkov (2009), Filbeck, Holzauer and Zhao (2014), and Humpfrey and Tan (2014). These studies are discussed below, as they also cover other ESG investing. A recent article is **Lobe and Walkshäusl** (2016, RMS, ABS 0 NF). That paper, appearing in a non-ABS rated journal, generally finds no effects. However, the sample selection appears peculiar, as the period covers only 1995-2007 even when the paper was not accepted for publication before October 2014, leaving a long gap. Finally, **Duran, Koh and Tan** (2013, PBFJ, ABS 2 F) provide some evidence based on seven Pacific-Basin markets that holdings and returns depend on the cultural norm of the country.[9]

[8] A similar line of argumentation holds for the 4-factors alphas in table 3B.

[9] In a non-academic study, **Junkin** (2015) estimates the historical costs of excluding tobacco stocks and other securities from CalPERS' portfolio to correspond to a present value between $3.7 and $8.3 bn.

2.2 Evidence based on ESG ratings using positive/negative screening

There is evidence that stock with high ESG ratings exhibit high future returns. The evidence is strongest in 1991-2004, while the returns of stocks with high ESG ratings do not appear to differ from benchmarks in 2005-2012. Some evidence suggests that returns again have been high since 2012.

Several companies provide ESG ratings for investors. Typically, the provider will rate firms along a number of dimensions, which can be grouped into the three ESG elements. It is then straightforward to construct portfolios based on a composite measure either within each ESG element or an overall ESG measure. The first such provider was KLD (now MSCI) that initiated ratings in 1990, thus providing the longest time span for asset pricing tests. The initial work using KLD data generally finds a positive relationship between ratings and returns. **Kempf and Osthoff** (2007, EFM, ABS 3 F) construct long-short value-weighted portfolios from the S&P 500 and DS 400 stocks in the period 1992-2004. They find significantly positive 4-factor alphas of around 5% per year using a 10% cut-off (table 5) of industry-adjusted ESG scores. **Statman and Glushkov** (2009, FAJ, ABS 3 F) generally confirm the findings of Kempf and Osthoff (2007) based on 1992-2007 data. Both studies also find that portfolios formed on community and employee relations exhibit the highest returns, while diversity, environment, products and human rights do not appear to affect returns (KO table 5, SG table 3).

Borgers, Derwall, Koedijk and ter Horst (2013, JEF, ABS 3 F) show that the ESG outperformance disappears after the initial sample period in Kempf and Osthoff. Also, they show evidence that the high returns in the initial sample period were due to market underreaction.

BDKH study the period 1991-2009 and use various ESG cut-off points to form long-short value- and equal-weighted portfolios. Four-factor alphas are positive and significant until 2004 after which they are close to zero and insignificant. The results are robust to changes in the ESG measure, including to industry-adjustments.

As benefits to ESG are initially intangible, the abnormal returns might appear because the market fails to fully incorporate intangible information. Prices are subsequently corrected as the intangible information eventually become tangible through higher earnings. Under this hypothesis we would expect that earnings of high ESG firms surprise positively, hence that returns around earnings releases are high and that actual earnings are higher than financial analysts' estimates. Indeed, BDKH find the ESG index is related positively to both earnings announcement returns and surprise in analyst' earnings forecasts until 2004. In the 2004-2009 period, the relationship disappears for earnings announcement

returns, while it is mixed (positive, negative, or no effect depending on the specification) for analysts' forecasts. Overall, the results suggest that investors became aware of potential benefits to ESG, hence the underreaction in the first period seems to have disappeared.

Halbritter and Dorfleitner (2015, RFE, ABS 1 F/NF) perform an analysis similar to the portfolio return part of BDKH. They do confirm the findings of initial outperformance followed by insignificant ESG returns in the subsequent period. Halbritter and Dorfleitner also investigate the returns to ESG scores provided by two other companies, ASSET4 and Bloomberg. Those samples only cover the more recent periods, namely 2003-2012 for ASSET4 and 2006-2012 for Bloomberg. Alphas are generally also insignificant for portfolios constructed based on the ASSET4 or Bloomberg data, confirming the KLD results for the later period. However, Halbritter and Dorfleitner also conduct Fama-MacBeth cross-sectional regressions (table 7). This testing procedure yields quite different results: While KLD scores remain insignificant, the ESG scores of ASSET4 and Bloomberg are both highly significant and positive. That is, firms with higher ASSET4 and Bloomberg ESG scores exhibit higher future returns. The Fama-Macbeth procedure entails using the full sample, applies a different weighting scheme, and utilizes controls based on characteristics rather than factor loadings. Still, results are usually consistent with those from factor regressions in other applications. Therefore, the difference observed in the current sample is unusual. The analysis in Halbritter and Dorfleitner does not allow for identification of the source of the difference in results between the Fama-MacBeth cross-sectional regressions and the portfolio factor regressions. It would be useful for future work to reconcile the results, for instance by analyzing the robustness with respect to changes in regressors or factors.

In sum, ESG scores were positively related to future returns in the 1990s, the effect seems to have disappeared in the 2000s. Still, the good news for ESG investors is that high ESG scores have not lead to lower future returns in the period up to 2012. This reviewer has not seen studies of risk-adjusted returns based on ESG scores in the period after 2012. However, **Larsen** (2016, F/I, ABS 0) reports a strong positive correlation between MSCI (formerly KLD) scores and realized returns during 2012-2016 (table 2). Moreover, Larsen finds that stocks with high scores exhibit a lower standard deviation of returns. This is promising for ESG investors, and it would be useful also to see alphas, i.e., returns adjusted for systematic risk.

Humphrey and Tan (2014, JBusEth, ABS 3 NF) use the KLD ratings and SIC codes to construct four SRI portfolios based on positive and negative screening. The sample period is 1996-2010. The two negative screens exclude based on tobacco, alcohol, gambling, weapons and nuclear. Relatively few stocks are

excluded in the negative screen, so there is a large overlap between the SRI and the benchmark portfolios, while the two positive screens select about one-third of the full-sample stocks. Table 3 shows that the SRI portfolio based on negative SIC screens underperform in the one-factor and four-factor models at the 5% level, while other three SRI portfolio returns are all lower than, but insignificantly different from, the benchmark returns.[10] Table 4 compares Sharpe ratios. All differences here are insignificant, even though the unscreened portfolio exhibits an annualized Sharpe ratio of 0.298 while the two SRI portfolios with positive screens have Sharpe ratios of 0.251 and 0.255.[11] Thus, the test power does not appear to be high.

While the sample period is of a reasonable length, it is surprising why HT do not start the sample in 1992 when KLD data became available. Indeed, Statman and Glushkov (2009) and **Filbeck, Holzauer and Zhao** (2014, JoI, ABS 0) both find that KLD-defined sin stocks outperform when the sample starts in that year.

2.2.1 Environmental screens

The evidence on investor returns to environmental screens is limited and the results are mixed.

Derwall, Guenster, Bauer and Koedijk (2005, FAJ, ABS 3 F) study the returns to a strategy based on Innovest Strategic Value Advisors' corporate eco-efficiency. They find that more eco-efficient firms exhibit higher stock returns than their less eco-efficient counterparts over the period 1995-2003. Given the very short time period, it is not surprising that the results are not robust. Specifically, long-short alphas are only significant in the 4-factor model (at the 5% level), whereas the 1-factor alpha is insignificant. Moreover, the authors backfill data as they only have Innovest scores starting in 1997. This is not standard procedure in finance research, as it can introduce severe look-ahead biases.[12] **Guenster, Bauer, Derwall and Koedijk** (2010, EFM ABS 3 F) correlate the Innovest eco-efficiency data with measures of operating performance and equity valuation. They find that eco-efficient firms become relatively more expensive, as measured by Tobin's q, during the sample period. This suggests

[10] The authors discount the underperformance of the SIC screened portfolios, noting that "We suspect, however, that the significant t statistic is more an artefact of the miniscule standard error [...] rather than denoting any real evidence of underperformance. These tiny standard errors are attributable to the almost identical return series of the two universes". Because few stocks are excluded the return series will be very similar; however, this will also affect the return differences, thus the *t*-statistic and the corresponding inferences are still valid.

[11] By comparison, Auer (2016, table 4) obtains a p-value of just 0.02 in a test for the difference in Sharpe ratios between 0.558 and 0.589 using a shorter sample period. The Auer study is discussed in section 2.2.3.

[12] The authors are aware of the issue and note that "results when we used real data for the 1997-2003 period are similar to those reported here." However, presumably results were weaker. Otherwise, there would be little point in backfilling.

that the return outperformance of the 2005 study was due to changes in valuation: either eco-efficient firms were initially undervalued or they became overvalued during the sample period.

The evidence on returns to an environmentally conscious investment strategy based on KLD scores is contradictory. Halbritter and Dorfleitner (2015, table 6) find that a long-short portfolio yields a 4-factor alpha of 6.6% per year during 1990-2001 – significant at the 1% level, while the strategy obtains negative, albeit insignificant, alphas in 2002-2012. In contrast to the positive results in the first part of the Halbritter and Dorfleitner period, Kempf and Osthoff (2005) and Statman and Glushkov (2009) find no evidence of outperformance based on KLD E scores during largely overlapping periods.

2.2.2 Social screens

There is evidence that investor returns to at least one social screen, namely employee satisfaction, were high during 1984-2011.

Edmans (2011, JFE, ABS 4* F) convincingly shows that firms with high employee satisfaction exhibit high future stock returns. As the paper notes, a "value-weighted portfolio of the '100 Best Companies to Work For in America' earned an annual four-factor alpha of 3.5% from 1984-2009, and 2.1% above industry benchmarks. The results are robust to controls for firm characteristics, different weighting methodologies, and the removal of outliers." The analysis appears to be carefully executed. In particular, characteristic-adjusted returns are used. In this procedure, returns of similar stocks are subtracted from the test stock before running factor regressions. This allows for controls for interaction effects which could otherwise result in the Fama-French model mispricing certain stocks. Also, Fama-MacBeth regressions (table 6) confirm results of the factor regressions. **Edmans** (2012, AMP, ABS 3 NF) extends the sample to cover 1984-2011 with almost identical results.

Employee satisfaction is an intangible variable, and it can benefit shareholders through employee motivation and retention. Edmans argues that the abnormal returns appear because the market fails to fully incorporate intangible information. Prices are subsequently corrected as the intangible information eventually become tangible through higher earnings. Consistent with this hypothesis, he finds that the realized long-term earnings growth of the "Best Companies" is higher than analysts' forecast 5 years earlier (relative to other companies). Moreover, the high returns persist up to 4 years after the portfolio formation date (table 9).

As noted earlier, Kempf and Osthoff (2005) and Statman and Glushkov (2009) found high returns to a strategy based on KLD scores on employee relations (and community). These results are thus consistent with Edmans (2011, 2012).

2.2.3 Governance screens

Good governance firms as measured by the G-index had higher returns than poor governance firms in 1990-1999. However, the return difference disappeared in the subsequent period. Some evidence suggests that other measures of governance predict returns in the subsequent period, but the evidence is not conclusive.

In a seminal paper, **Gompers, Ishii and Metrick** (2003, QJE, ABS 4* F/NF) construct a firm-level governance index, G-index, based on 24 provisions that weaken shareholder rights. A firm with weak shareholder rights would have adapted many provisions, hence have a high G-index, while a firm with strong governance would have a low G-index. For a sample of 1500 large U.S. firms during 1990-1999, they find that a portfolio long the 10% lowest G-index firms and short the 10% highest G-index firms earned an abnormal return of 8.5% per year.

Bebchuk, Cohen and Wang (2013, JFE, ABS 4* F) extend the sample period of Gompers, Ishii and Metrick to cover 1990-2008. They show that the abnormal returns are insignificant during 2000-2008. This suggests that the effect has disappeared after the original sample period. Still, because of the short time period, the asset pricing tests have low power. In fact, a modified governance index (the E-index) yields long-short returns yields a 4-factor alpha of 3.1% per year when value-weighted and 4.2% per year when equal-weighted. Thus, while insignificantly different from zero, the results would also be consistent with relatively high returns based on the governance index.

Neither GIM nor BCW find evidence that good governance firms are more risky. Therefore, the higher returns appear to be a market anomaly. Arguably, investors in the 1990s were simply not aware of the detrimental effects of the governance provisions, many of which had been implemented in the 1980s. However, because of the greater focus on governance in the 2000s, investors had become aware of the effects. Consistent with this learning hypothesis, BCW find that good governance firms tend to report more positive earnings surprises than poor governance firms in the 1990s, but that this relationship disappears in the 2000s.

While the results indicate that the pricing effect has disappeared after 2001, BCW show that good governance firms still tend to have higher valuation, profitability and growth (table 8) in the later period than do poor governance firms. While the correlation between governance and these performance measures does not resolve the direction of the causal relationship, they do imply that governance can be used as an indicator of performance.

Gu and Hackbarth (2013, RoF, ABS 4 F) show that the relationship between governance and stock returns identified in Gompers, Ishii and Metrick (2003) is

concentrated among high transparency firm, as measured by the distribution of analysts' forecasts. As Gu and Hackbarth note, this is consistent with the view that highly transparent firms are more valuable takeover targets, because acquirers can bid more effectively and identify synergies more precisely. By contrast, the results are not consistent with the view that outside investors cannot monitor opaque firms easily hence those firms would benefit more from good governance.

In the 2000-2011 period, along with the inability of the G-index to predict returns in the overall sample, the relationship is also insignificant for high transparency firms (table 6). However, it remains possible that a differential effect remains across levels of transparency (as all six long-short portfolios in table 6E exhibit higher returns for high transparency than for low transparency firms, no tests for significance of these differences are reported).

Based on 2004-2012 data of STOXX 600 (large European stocks) and ESG ratings from Sustainalytics, **Auer** (2016, JBusEth, ABS 3 NF) studies the effect of exclusionary screening on portfolio Sharpe ratios. During the sample period Sustainalytics rated stocks only in response to investor requests. While 520 of 892 stocks were rated at some point, no information is provided on how many stocks were rated at given points in time. Presumably, few stocks were rated at the beginning of the sample period.

The main result (table 4) is that the Sharpe ratio of the rated stocks increases when excluding stocks with poor Governance rating (significant at the 5% level), while exclusionary screening based on E and S does not affect Sharpe ratios.[13]

The usage of Sharpe ratios rather than alphas assumes that that the investor's overall portfolio consists of the benchmark portfolio (consisting of the rated stocks). If the investor holds a more diversified portfolio, the usual alpha analysis would be more appropriate. While no direct information is given on the risk-return of the excluded portfolios, one can infer from table 4 that the 20% lowest governance-rated stocks had a yearly average return of 1.7% less than the remaining 80%. One cannot compute differences in risk as this depends on the covariance structure, but the portfolio risk does fall when excluding the lowest G-rated stocks.

The sample period is very short, thus power of the tests is low, so the insignificant results for E and S screening is not surprising. Moreover, the period includes the financial crisis, which likely will dominate in particular the volatility

[13] Auer also notes that the portfolio of rated stocks has a higher Sharpe ratio than the full sample, but since we do not know whether the unrated stocks would have had relatively low or high ratings, we cannot conclude whether this indicates high or low returns for high ESG stocks.

estimates, thus the Sharpe ratios. Indeed, figure 3 shows that Sharpe ratios diverge during the financial crisis. Nonetheless, it is notable that Governance ratings appear to predict Sharpe ratios.

2.3 Stock market reactions to ESG events

Event studies indicate that the stock market does not respond positively to ESG/CSR initiatives by firms. While the results suggest that agency issues are a genuine concern, they also suggest that such concerns can be mitigated through sound corporate governance.

Event studies can provide powerful analyses of the value to ESG/CSR policies. A difficulty with long-term return studies is the assumption that we have identified the correct asset pricing model. Often, studies can be criticized on the grounds that an apparent abnormal return is caused by a latent risk factor, i.e., that the study does not properly account for differences in risk. Because event studies measure returns over short time windows of a few days, the (risk-driven) expected return component is typically negligible. Also, test power if often high because the volatility of returns is low over short time windows, thus making it easier to establish the statistical significance of any relationship.[14]

A few studies have used event studies to investigate the stock market response to changes in firms' environmental policies. **Fisher-Vanden and Thorburn** (2011, JEEM, ABS 3 NF) study the abnormal stock returns around announcements that firms have joined two voluntary environmental programs, namely the U.S. Environmental Protection Agency's Climate Leaders program and Ceres. The Climate Leaders corporate partners set specific 5-10 year goals for the reduction of greenhouse gasses, while membership of Ceres involves adapting environmentally conscious codes of conduct. The sample includes 117 announcements by large U.S. firms between 1993 and 2008. The results show a drop in market values by around 1% when firms join the Climate Leaders program – not only statistically significant, but also economically significant as it corresponds to an average loss of $3 billion! Moreover, when firms announce their specific reduction targets their market values drop by another 1%. By contrast, when firms join Ceres abnormal returns are insignificant.

Probit analysis shows that firms with a high number of shareholder resolutions toward climate action and firms with poor corporate governance score (as measured by the G-index discussed earlier) are more likely to join the Climate Leaders program. Moreover, returns are significantly more negative for firms with a poor corporate governance score, suggesting the existence of agency problems. In particular, as discussed in section 1.1.1 managers might seek to enhance their personal recognition or express their social preferences via ESG

[14] Higher volatility implies lower test power, and volatility increases with the square root of time.

initiatives to the detriment of shareholder value. As the authors conclude, "thus, it seems that firms are entering the Climate Leaders program despite the prospect of lowering shareholder value either because they are facing institutional pressures to do so, or because managers face less shareholder oversight, allowing them more discretion to make these types of voluntary environmentally responsible investment decisions."

Jacobs, Sinhal and Subramaniam (2010, JOM, ABS 4* NF) examine the stock market reaction to the announcement of various types of corporate environmental initiatives, including environmental business strategies, environmental philanthropy, voluntary emission reductions, econ-friendly products, renewable energy, and recycling. They mostly find insignificant results, except for voluntary emission reduction, for which they also find significantly negative returns consistent with Fisher-Vanden and Thorburn, and for environmental philanthropy, for which they find significantly positive returns.

Krüger (2015, JFE, ABS 4* F) studies the stock market response to 2,116 corporate event identified by KLD as either negative or positive along an ESG dimension. He finds a significantly negative response to negative ESG events. However, this is not surprising, as the news often imply negative cash flows, for instance a product recall or a court decision against the company. ESG efforts may strive to minimize the occurrence of such events, but measuring returns to negative events does not account for the costs of this minimization. Of course, the question for investors is whether the net effect of ESG efforts is positive.

The positive ESG events seem to be more forward looking, e.g., an investment in an ESG-related initiative. Therefore, these are more informative of the effects of ESG initiatives. Overall, the stock market reacts slightly negatively to positive ESG events.

Perhaps more interestingly, Krüger then interacts the response to measures of agency concerns. Krüger uses book leverage and liquidity as measures of agency concerns, following the free cash flow theory of Jensen (1986). Consistent with the agency theory, he finds a negative effect of liquidity and a positive effect of book leverage on returns around positive ESG events. Moreover, there is no interaction effect among negative ESG events. Because we would expect negative ESG events to be less related to agency conflicts, this reinforces the agency theory.

As in Fisher-Vanden and Thorburn, overall these interaction effects suggest that managers might seek to enhance their personal recognition via ESG initiatives to the detriment of shareholder value. Unlike Fisher-Vanden and Thorburn, Krüger does not directly interact the response with measures of corporate governance. It would be very interesting to see how ES initiatives interact with governance.

One critique of event studies is that other events may be happening within the event window that could confound the results. For instance, a positive ESG event could have been a company response to some negative news about the company. In that case, even if the ESG event is truly value creating, the stock market response may seem negative because of the confounding event.[15] While the relevance of such criticism is hard to evaluate, the stock market response does indeed appear to depend on earlier news. Specifically, Krüger finds a more positive stock market response to positive ESG events if KLD has registered an ESG concern within the last year for the company. This is especially true for employee relations, environment and human rights.

Flammer (2013, AMJ, ABS 4* NF) identifies events as eco-harmful or eco-friendly based on word searches in the Factiva database. As in the Krüger study, eco-harmful events could be accidents, such as oil spills, with clear cash flow implications. Indeed, the stock market reaction is negative. Contrary to Krüger, Flammer finds a positive stock market reaction to eco-friendly events. However, while many of the events clearly are environmentally friendly, such as recycling initiatives, other events are less clearly environmentally friendly. For instance, reports on February 14, 1991 regarding the Exxon settlement with Alaska following the 1989 oil spill appear to be treated as an eco-friendly event. Yet, the conservation group Defenders of Wildlife did not view the settlement as eco-friendly, noting on U.S. Newswire that "[w]e are flabbergasted that Gov. Hickel is attempting to settle claims against Exxon for less than $1 billion. Early on in this process, conservationists estimated that $5 billion would be needed. The cleanup, restoration, replacement and healing are far from being completed. Much oil still resides in subsurface sediments on many beaches, especially those outside Prince William Sound. Exxon has a corporate responsibility to pay the full costs for damage, response, and resolution." Surely, the settlement was positive news for Exxon's stock price, but hardly positive news for the environment. Such probable misclassification makes it difficult to interpret the findings.

2.4 Active ownership

There is evidence that active ownership by ESG investors can create value, both for shareholders and other stakeholders

The literature discussed so far have assumed that investors are active in selecting stocks, but that they are passive owners who do not directly try to influence management behavior. However, some ESG investors engage in active ownership via direct contact with senior management and by exercising ownership rights at shareholders' meetings.

[15] The Fisher-Vanden and Thorburn study is less prone to such criticism due to a shorter event-window, but it is also more narrowly focused than do the Krüger study.

Dimson, Karakas and Li (RFS, 2015, ABS 4* F) use a proprietary dataset from a large institutional ESG investor to study the effects of active ownership. The sample consists of 382 successful and 1,770 unsuccessful engagements for 613 U.S. publicly listed companies during 1999-2009. Their results suggest that ESG activism can be value enhancing. Specifically, they find that one-year cumulative abnormal returns are 7% on average following successful engagement, while there is no market reaction to unsuccessful engagements (figure 1). When categorizing engagements as either ES or G, they report a similar return response to the two types of engagements.

Also, ROA improve for successful engagement relative to unsuccessful ones, albeit mainly for ES engagements. Furthermore, pension activists and SRI funds increase their holdings after successful ES engagements, consistent with the clientele effect of Hong and Kacperczyk (2009) for sin stocks.

A drawback of the Dimson, Karakas and Li study is that the data is from a single institutional investor who chose to share the data. Had the institution anticipated different results, it would have less incentive to share. That is, other investors might be less successful pursuing active ownership. Nonetheless, the results demonstrate that active ownership on ESG matters can create value, both for shareholders and other stakeholders.

2.5 Meta studies

Several review articles and meta studies of ESG investing exist. The most comprehensive in terms of number of studies is **Friede, Busch and Bassen** (2015, JSusFinInv, ABS 0) who conduct a vote-count and meta-study of 2200 empirical studies. They find that a "large majority of studies reports positive findings" on the relationship between ESG and corporate financial performance (CFP). On this basis, they conclude that "the results show that the business case for ESG investing is empirically very well founded." However, there are at least three reasons to be skeptical towards this claim.

First, some authors appear to favor a positive effect of ESG rather than applying an objective scientific approach. There are examples of studies neglecting a particular relationship and examples of tendentious language. Biases in the review studies, that form much of the basis for the Friede, Busch and Bassen analysis, would aggravate any biases in the individual studies. Some review studies explicitly set out to make the business case for ESG, as for instance Clark, Feiner and Viehs (2015). They find a positive relationship in 94 of 110 studies

that they survey while only one has a negative relationship. However, this reviewer finds it hard to agree with many of the "positive" classifications.[16]

Secondly, the studies investigate the relationship between ESG and CFP of which stock returns is only one of many variables. The CFP is often an accounting variable such as ROA or ROE. However, as discussed in section 1.2, that sound ESG decisions are sound business decisions does not imply that investors obtain superior returns from investing in these firms. Of course, this depends on whether the information is already priced into the stock. Only if the stock market systematically undervalues such information will the ESG investor obtain high returns.

Thirdly, as also noted in section 1.2, it is notoriously difficult to infer causality from correlations between different corporate variables. A firm with a high ESG score might have high profitability, but the high profitability might be driving the ability to invest in ESG rather than the ESG investments causing high profitability. To avoid such issues of reverse causality, one needs to identify exogenous variation in the ESG variable, rather than simply showing a correlation between the two variables. This causality issue is usually not addressed in the CSR literature.

Friede, Busch and Bassen note [p. 226] that they "clearly find evidence for the business case for ESG investing. This finding contrasts with the common perception among investors. The contrary perception of investors may be biased due to findings of portfolio studies, which exhibit, on average, a neutral/mixed ESG–CFP performance relation." A plausible alternative interpretation is that investors are not biased, but that they understand that the findings from portfolio studies are more relevant for making the business case for ESG investing.

3 Conclusion and discussion of the findings

The literature on ESG investing has been prolific during the last decade. The literature is for the most part methodologically sound, but some authors do appear to wish to make the business case for ESG investing rather than applying a more dispassionate scientific approach. This has necessitated a rather careful evaluation of the relevant articles. Because there are many dimensions to ESG

[16] For instance, Fisher-Vanden and Thorburn (2011) is counted as positive, contrary to the authors conclusion that "the results indicate that voluntary environmentally responsible investments of this kind conflict with firm value maximization." Likewise, Krüger (2015) and Capelle-Blancard and Laguna (2010) are recorded as positive because the negative events yield negative stock market responses, but these are simply negative cash flow events and one would have to know the cost of lowering the probability of such events. Further, there is double-counting of positive studies based on the same data, such as Edmans (2011, 2012) and the governance index studies, while results on sin stocks are ignored.

investing, it is not surprising that no simple answer exists to the question of the profitability of ESG investing.

Overall, the most consistent finding in the current review is that sin stocks exhibit outperformance. This implies that *sector-based exclusions lower expected portfolio returns*. In turn, this creates potential conflicts for institutional investors and in particular within pension funds into which individual savers are locked. First, there could be a conflict between the personal values of the portfolio manager and her fiduciary responsibility towards savers. Moreover, it creates a conflict of interest between individual savers with different preferences and values. If solutions are to be found for such conflicts, pension funds would be well advised to acknowledge rather than ignore these conflicts.[17]

High expected returns to sin stocks imply that these firms experience a higher cost of capital. As such, the exclusions do achieve a presumed goal. If only some firms experience a higher cost of capital because, say, they are shunned by investors because of low ESG scores, then those firms will be at a competitive disadvantage. However, sin stocks are defined at the industry level, such as tobacco. Therefore, the effect of sin stock exclusions is to raise the general costs to firms in the industry without changing the relative competitiveness of the firms. Customers would thus bear the cost through higher prices. As such, portfolio exclusion is akin to levying a tax on the product, say a cigarette tax. The difference from an ordinary tax is that the beneficiaries of those "tax revenues" are the investors who choose to invest in tobacco firms. Presumably, this is not an intended consequence from ESG investing.

There is evidence that *stock with high ESG ratings exhibit high future returns*. The evidence is strongest in 1991-2004, while the returns of stocks with high ESG ratings do not appear to differ from benchmarks in 2005-2012. Some evidence suggests that returns again have been high since 2012.

The evidence on investor returns to environmental screens is limited and the results are mixed. There are also relatively few studies on the effect of social screens. However, investor returns to at least one social screen, namely employee satisfaction, were high during 1984-2011.

Good governance firms as measured by the G-index had higher returns than poor governance firms in 1990-1999. However, the return difference disappeared in the subsequent period. Some evidence suggests that other measures of

[17] A detailed discussion of possible solutions is outside the scope of this report. However, one possibility is to decompose ex post portfolio returns according to different sin dimensions, and offer savers the possibility of making offsetting actions, such as donating to a particular cause. An alternative would be to create various portfolios within the pension fond to which savers could choose to join.

governance predict returns in the subsequent period, but the evidence is not conclusive.

Using earnings announcement returns, studies suggest that the stock market initially underreacted to the information contained in ESG ratings and to governance information, but that this underreaction disappeared in the 2000s. If correct, then ESG investors should not expect outperformance based on portfolio construction with ESG ratings. On the other hand, *there is presently no empirical evidence to suggest that such portfolio construction will lead to lower performance* (except for sector exclusions).

Event studies indicate that *the stock market does not respond positively to ESG/CSR initiatives taken by firms*. While the results suggest that agency issues are a genuine concern, they also suggest that such concerns can be mitigated through sound corporate governance. These results are consistent with the theoretical considerations on the interaction effects of ES and G outlined in section 1.1.1 in that the value of ES initiatives depends on governance. The lesson for ESG investors is that *a high ES rating is a positive signal if the G rating is also is high, but a negative signal if the G rating is low*. Still, the evidence on this interaction effect is still limited and is based on event returns rather than longer-term portfolio returns. More studies are needed before making firmer conclusions regarding the investment case.

Finally, there is evidence that *active ownership by ESG investors can create value*, both for shareholders and other stakeholders. Specifically, successful ESG engagements by a large institutional investor into U.S. firms were followed by abnormal returns in the subsequent year. Likewise, accounting performance improved following successful engagements. Because of the proprietary nature of the engagement data, the evidence is still limited, but clearly encouraging for ESG investors pursuing active ownership.

Recent years have witnessed a spectacular growth in ESG investing. Therefore, any past evidence should be interpreted with additional caution. The nature of financial markets is such that as an investment strategy becomes widespread, abnormal returns eventually disappear or become negative. Simple conditioning information such as ESG ratings used by many investors is most susceptible to this effect. Therefore, the use of more subtle ESG information and active ownership might be a more effective strategy for the committed ESG investor in the long run.

References

Adamsson, H. and Hoepner, A.G., 2015. The 'Price of Sin' Aversion: Ivory Tower Illusion or Real Investable Alpha?.

Auer, B.R., 2016. Do Socially Responsible Investment Policies Add or Destroy European Stock Portfolio Value?. *Journal of Business Ethics*, *135*(2), pp.381-397.

Ball, R. and Brown, P., 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, pp.159-178.

Bebchuk, L.A., Cohen, A. and Wang, C.C., 2013. Learning and the disappearing association between governance and returns. *Journal of Financial Economics*, *108*(2), pp.323-348.

Berman, D., 2002. Why Sin is good: Tobacco, alcohol, and gaming stocks can add sizzle to your portfolio. *Money Sense*.

Bernard, V.L. and Thomas, J.K., 1989. Post-earnings-announcement drift: delayed price response or risk premium?. *Journal of Accounting Research*, pp.1-36.

Borgers, A., Derwall, J., Koedijk, K. and Ter Horst, J., 2013. Stakeholder relations and stock returns: On errors in investors' expectations and learning. *Journal of Empirical Finance*, *22*, pp.159-175.

Capelle-Blancard, G. and Laguna, M.A., 2010. How does the stock market respond to chemical disasters?. *Journal of Environmental Economics and Management*, *59*(2), pp.192-205.

Chava, S., 2014. Environmental externalities and cost of capital. *Management Science*, *60*(9), pp.2223-2247.

Derwall, J., Guenster, N., Bauer, R. and Koedijk, K., 2005. The eco-efficiency premium puzzle. *Financial Analysts Journal*, pp.51-63.

Dimson, E., Karakaş, O. and Li, X., 2015. Active ownership. *Review of Financial Studies*, *28*(12), pp.3225-3268.

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

Edmans, A., 2012. The link between job satisfaction and firm value, with implications for corporate social responsibility. *The Academy of Management Perspectives*, *26*(4), pp.1-19.

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

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

Fama, E.F. and French, K.R., 1997. Industry costs of equity. *Journal of Financial Economics*, *43*(2), pp.153-193.

Filbeck, G., Holzhauer, H.M. and Zhao, X., 2014. Using Social Responsibility Ratings to Outperform the Market: Evidence from Long-Only and Active-Extension Investment Strategies. *The Journal of Investing*, *23*(1), pp.79-96.

Fisher-Vanden, K. and Thorburn, K.S., 2011. Voluntary corporate environmental initiatives and shareholder wealth. *Journal of Environmental Economics and Management*, *62*(3), pp.430-445.

Flammer, C., 2013. Corporate social responsibility and shareholder reaction: The environmental awareness of investors. *Academy of Management Journal*, *56*(3), pp.758-781.

Flammer, C., 2015. Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science*, *61*(11), pp.2549-2568.

Friede, G., Busch, T. and Bassen, A., 2015. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, *5*(4), pp.210-233.

Gompers, P., Ishii, J. and Metrick, A., 2003. Corporate governance and equity prices. *The Quarterly Journal of Economics*, *118*(1), pp.107-156.

Grossman, S.J. and Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), pp.393-408.

Guenster, N., Bauer, R., Derwall, J. and Koedijk, K., 2011. The economic value of corporate eco-efficiency. *European Financial Management*, *17*(4), pp.679-704.

Halbritter, G. and Dorfleitner, G., 2015. The wages of social responsibility—where are they? A critical review of ESG investing. *Review of Financial Economics*, *26*, pp.25-35.

Heinkel, R., Kraus, A. and Zechner, J., 2001. The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis*, *36*(04), pp.431-449.

Hoepner, A.G. and Schopohl, L., 2016. On the price of morals in markets: an empirical study of the Swedish AP-Funds and the Norwegian Government Pension Fund. *Journal of Business Ethics*, pp.1-28.

Hoepner, A.G. and Zeume, S., 2013. Fiduciary duty and sin stocks: is vice really nice?.

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

Hong, H., Kubik, J.D. and Scheinkman, J.A., 2012. *Financial constraints on corporate goodness* (No. w18476). National Bureau of Economic Research.

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

Jacobs, B.W., Singhal, V.R. and Subramanian, R., 2010. An empirical investigation of environmental performance and the market value of the firm. *Journal of Operations Management*, *28*(5), pp.430-441.

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

Jensen, M.C., 1986. Agency costs of free cash flow, corporate finance, and takeovers. *The American economic review*, *76*(2), pp.323-329.

Jensen, M.C., 2001. Value maximization, stakeholder theory, and the corporate objective function. *Journal of Applied Corporate Finance*, *14*(3), pp.8-21.

Junkin, A., 2015. Comprehensive Divestment Analysis. Wilshire Consulting.

Kempf, A. and Osthoff, P., 2007. The effect of socially responsible investing on portfolio performance. *European Financial Management*, *13*(5), pp.908-922.

Krüger, P., 2015. Corporate goodness and shareholder wealth. *Journal of financial economics*, *115*(2), pp.304-329.

Larsen, S., 2016, Giver bæredygtighedsanalyser værdi i forvaltning af aktier? *Finans/Invest*, 2017(6), pp.6-12.

Lobe, S. and Walkshäusl, C., 2016. Vice versus virtue investing around the world. *Review of Managerial Science*, *10*(2), pp.303-344.

Markowitz, H.M., 1959, Portfolio Selection: Efficient Diversification of Investments. Wiley, Yale University Press, 1970, 2nd edn. Basil Blackwell, 1991.

Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, *42*(3), pp.483-510.

Roberts, M.R. and Whited, T.M., 2013. Endogeneity in Empirical Corporate Finance. Chapter 7 in the Handbook of Economics and Finance.

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

Statman, M. and Glushkov, D., 2009. The wages of social responsibility. *Financial Analysts Journal*, *65*(4), pp.33-46.

Tirole, J., 2001. Corporate governance. *Econometrica*, 61(1), pp.1-35.

Trinks, P.J. and Scholtens, B., 2015. The opportunity cost of negative screening in socially responsible investing. *Journal of Business Ethics*, pp.1-16.