

The expected returns of ESG excluded stocks. The case of exclusions from Norway's Oil Fund

Erika Berle, Wanwei (Angela) He and Bernt Arne Ødegaard*

April 2022

Abstract

What are the consequences of widespread ESG-based portfolio exclusions on the expected returns of firms subject to exclusion? We consider two possible theoretical explanations. 1) Short-term price pressure around the exclusions leading to correction of mispricing going forward. 2) Long term changes in required returns. We use the exclusions of Norwegian Government Pension Fund Global (GPF – “The Oil Fund”) to investigate. GPF is the world’s largest SWF, and its ESG decisions are used as a model for many institutional investors. We construct various portfolios representing the GPF exclusions. We find that these portfolios have significant superior performance (alpha) relative to a Fama-French five factor model. The sheer magnitude of these excess returns (5% in annual terms) leads us to conclude that short-term price pressure can not be the only explanation for our results, the excluded firms expected returns must be higher in the longer term.

Keywords: ESG investing; Exclusions; Oil Fund

JEL Codes: G10; G20

Introduction

“Sustainability concerns” may be a good way to characterize recent trends in society, concerns which have spread to financial markets. In the investment community, the early term “Socially

*All authors are from the Business School at the University of Stavanger (HHUiS), Stavanger, Norway. Corresponding author: Bernt Arne Ødegaard. Email: bernt.a.odegaard@uis.no. Postal adress: UiS Business School, EOJ Building, University of Stavanger, NO-4036 Stavanger, Norway. Declaration of Interest: none.

Responsible Investing” has been replaced with the current concern with “ESG-aware investing,” where ESG is concerned with environmental, social and governance considerations of a firm’s decisions. This practitioner interest has been followed by academic interest. Most of the academic work is empirical, looking for links between ESG characteristics and company performance. This literature has grown so quickly that a recent survey (Coqueret, 2021) lists 10 *meta* studies of the issue, the last of which (Whelan et al., 2021) surveyed over 1,000 studies produced in the period 2015–2020. There is no clear conclusion from these studies.

Much of the empirical work on ESG does not really discuss why the focus on ESG, and the removal of part of a stock’s potential investors through exclusion, should induce differences in expected returns. Intuitively, exclusions should reduce the feasible set of investment portfolios, worsening the risk/return trade-off.¹ We are, however, seeing an emergent theoretical literature on the equilibrium consequences of ESG concerns. Pástor et al. (2021) and Pedersen et al. (2021) are the best known examples. Key to these and related models is that investors have preferences over both the monetary return from an investment, and investments’ ESG characteristics. For example, investors can feel satisfaction in not supporting gun violence through ownership of gun manufacturers (negative screening). On the other hand, investors may want to support clean energy generation, and get an extra utility from investing in such energy companies. These models support differences in expected returns depending on the ESG characteristics of a stock. Pástor et al. (2021) argues that expected returns is decreasing in ESG quality. The model of Pedersen et al. (2021) has similar implications, but less clear cut predictions, as the equilibrium depends on the relative proportions of different types of investors, and one can end up with both higher and lower expected returns for high quality ESG stocks. An alternative interpretation of these effects is provided by Cornell (2021), who view high returns for high quality ESG firms as a transitory phenomenon due to changing preferences affecting costs of capital.

These theoretical models predict ESG-related differences in long term expected returns. An alternative explanation of return differences is more concerned with short-term effects. Company exclusions tend to be clustered in time, either by the example of an exclusion by e.g. the GPFM being followed by competing funds or by news about a company leading to a reassessment of ESG quality. For either of these reasons, we may see many institutional

¹Indeed, this kind of reasoning was behind some of the early empirical work on “sin” stocks (Hong and Kacperczyk, 2009; Chava, 2014), which found evidence of superior returns for industries typically excluded by ESG screens (tobacco, weapons).

investors selling their stakes simultaneously, leading to a negative price pressure on the stock. If the expected cash flows of the firms are unchanged, this is temporary mispricing expected to be reversed. Such reversals will lead to higher returns going forward. In our analysis, we will consider these two explanations (changes in long term expected return versus short-horizon price pressure effects).

In this paper we look at the exclusion issue using data from Norway's Government Pension Fund Global (GPFG – "The Oil Fund"). GPFG is the world's largest SWF, with an equity portfolio valued at one trillion USD at the end of 2021. The GPFG ESG decisions are used as a model for many institutional investors. The GPFG has followed a policy of excluding stocks using on an expanding set of ESG criteria, starting with exclusions based on production of cluster ammunition and nuclear armaments. We consider the list of the GPFG's exclusions a list of "low quality ESG" companies. We construct various portfolios representing these excluded stocks and measure the (abnormal) performance of these portfolios. The portfolios are constructed for all excluded stocks, also for separate reasons for exclusion, for example we distinguish product-based exclusions and conduct-based exclusions.

We measure performance as alpha relative to the global five factor model of Fama and French (2017), and find significantly positive alphas for most variations we investigate. For example, the equally weighted portfolio of all excluded stocks has a statistically significant alpha of almost 5% (in annual terms). An estimation for the US part of the exclusion portfolio has an even higher alpha (5.7% in annual terms).

These magnitudes lead to us conclude that the long term expected returns of excluded firms must have increased. Any short term price pressure can not lead to return effects of this magnitude. In fact, there are extant event study evidence (Atta-Darkua, 2020; Eriksen et al., 2020) that puts the short term reaction to the GPFG's exclusion announcements at around 1.5%. We, therefore, conclude that much of the estimated alpha is due to changes in long term expected returns.

One way to interpret the results is to think in terms of the equilibrium cost of capital. Firms with such a bad ESG rating that they are excluded from the GPFG may be facing an uphill struggle in raising capital for new investments. They therefore have to offer higher returns to the investors willing to "dirty their hands" by providing capital. From society's point of view, this is a good thing, the higher capital cost will actively discourage investment in low rated ESG projects, as fewer projects will be able to sustain such high returns.

The structure of the paper is as follows. We first, in Section 1 give an overview of the is-

sues and discuss the literature before giving some background on the Norwegian Government Pension Fund Global (GPF) in Section 2. Section 3 discusses the data sources and gives some summary statistics. Section 4 contains the result, describing the returns of the exclusion portfolios, and tests for significance. We finally offer a short conclusion. A separate Appendix provides some additional analysis.

I Literature

We are looking at the equity investment decisions by institutional investors, how they are affected by environmental, social, and governance (ESG) considerations, and the equilibrium implications for stock returns. While the concept of ethical investing has a long history (Liang and Renneboog, 2017), it is in the last fifteen years or so that the ESG viewpoint has moved to the forefront. Mutual funds marketed as “socially responsible” and “sustainable” has seen large inflows, to the extent that today, one third of U.S. assets under management are subject to a sustainable investment strategy (SIF, 2020). Regulation is also a driver of the increased ESG focus. The best known example is the EU introduction of a taxonomy of sustainable activities.

From the large institutional investors point of view, ESG considerations will affect all of their portfolio decisions. The investors investment universe needs ranking in the ESG dimension, which will affect over- and under-weighting decisions. For low ESG ranked stocks, an institutional investor will react by either dialogue or divestment. The most common reaction from institutional investors is dialogue, either directly or through voting at the annual meeting. Institutional investors argue that dialogue is a better way of achieving change. There is also research pointing to the value effect of dialogue (Lewellen and Lewellen, 2022).

Exclusion is chosen in only a minority of cases and is viewed as a reaction of last resort. Even if it is a last resort, the number of stocks seeing widespread exclusions is increasing. We will discuss the causes of exclusions in the context of the oil fund, let us instead discuss some theoretical aspects of exclusions and ESG rankings.

To simplify the discussion, let us label as “green” the stocks with high quality ESG ranking and as “brown” those with low quality ESG ratings. The intuitive argument of Pástor et al. (2021) is that when there is a subset of investors that gets utility from green stocks beyond the pure monetary return, green stocks can sustain lower returns, which provides the prediction that brown stocks will have higher expected returns than green stocks. There is, however, a

tradeoff here. The higher expected returns for brown firms also mean that costs of capital for these firms are higher. Thus, when financing new investments, the brown firms will face a steeper hurdle rate than green firms. These brown firms will then have an incentive to become greener to access cheaper capital. In equilibrium, this will be a true tradeoff, and future investments will be greener.² This tradeoff is behind the ambiguity in theoretical predictions. For example, the more ambiguous results of Pedersen et al. (2021).³ The theory thus leaves the field open for some empirical input to the debate, which we aim to provide.

We are, however, not the first to provide such empirical evidence. Let us discuss some of the relevant literature. One strand of the research literature concerns what is called “sin stocks,” which looks at the stock returns of industries such as gambling. Studies such as Hong and Kacperczyk (2009), Chava (2014), and Bolton and Kacperczyk (2021), among others, find evidence that sin stocks outperform comparable stocks. Specifically, Hong and Kacperczyk (2009) show that sin stocks, which in their study are stocks involved in producing alcohol, tobacco and gambling, have significantly positive abnormal returns based on both a Fama-French four-factor model (Fama and French, 1996; Carhart, 1997), and a cross-sectional regressions analysis (Fama and MacBeth, 1973). In a similar vein, Chava (2014) looks at the effects of environmental concerns and argues that the stocks excluded by environmental screens have higher cost of capital and higher expected returns. More recently, Bolton and Kacperczyk (2021), find evidence that stocks with higher carbon emissions (both in terms of levels and innovations) earn higher returns. However, mixed evidence of such “sin” return differences is reported by other studies. For example, while Fabozzi et al. (2008) examines 21 countries and show that sin stocks significantly outperform in 19 and mildly underperform in 2 countries, Blitz and Fabozzi (2017) revisit the sin anomaly and show that the alpha estimates become economically small and statistically insignificant after controlling for the five Fama-French factors. Outside of the US, Durand et al. (2013) find that sin stocks significantly underperform in seven Pacific-basin markets. Overall, the evidence for the existence of a sin premium is still ambiguous.

We are also not the first to use the exclusions of the Norwegian GPF as objects of study. These existing studies can be grouped by the question they ask. First, two recent studies, Attardkua (2020) and Eriksen et al. (2020) looks at the short-term price reactions to exclusion

²Note that this opportunity to finance green projects cheaper provides an incentive for “green-washing” by misrepresenting the green credentials of projects.

³This theoretical tradeoff is understood by institutional investors, as evidence of this we point to a summary of these theoretical models in a working paper from the GPF (NBIM, 2021b).

announcements by the oil fund (i.e. these are event studies). In both cases they estimate negative announcement price effects. Second, several papers look beyond the immediate market reaction and investigate the returns of the stocks excluded by the GPFG. Beck and Fidora (2008) and Dewenter et al. (2010) were early studies. More recent is Hoepner and Schopohl (2018), which analyze the exclusions from the GPFG and the Swedish AP-funds. They find no significant return differences relative to the funds' benchmark portfolios. Their focus on the portfolios' long-term performance is similar to ours, but we employ a longer sample and different empirical methods. We will return to these studies after we present our results, point out differences, and relate these studies' results to ours.

2 The oil fund and the fund's exclusions

In this section we provide some background information on Norway's GPFG, and the fund's evolving ESG and exclusion policies.⁴

The purpose of the fund is to manage Norway's considerable resource wealth stemming from the oil industry. The fund is an attempt to avoid the consequences of the “resource curse” — the adverse effects of a sudden increase in natural resource wealth (Ross, 1999). Effectively, the fund translates the oil and gas in the North Sea into a well-diversified financial portfolio invested outside of Norway to avoid overheating the Norwegian economy. The first oil revenues were transferred into the fund in 1996. Initially, the fund invested in treasury securities, but it was soon realized that the size of the revenues channelled into the fund would make it necessary to diversify the asset mix. In 1998 the funds' portfolio was split into 40% equity and 60% fixed income securities. The equity fraction has since increased to its current level of 70%, and several other asset classes, such as real estate and infrastructure investments, have been added. At the end of 2021, the fund's market value was 12,340 billion NOK (NBIM, 2021a).

In our discussion, we will concentrate on the equity part of the portfolio. The equity part of the GPFG was valued at 8,878 billion NOK (1,014 billion USD) at yearend 2021. At the time, the fund's portfolio contained 9,338 stocks across 65 countries.

The fund is managed by Norges Bank (the central bank of Norway) on behalf of Nor-

⁴For more information we refer to NBIM's recent survey of their ESG history (NBIM, 2020). For more academic views of the fund, we refer to Chambers et al. (2012, 2021) and the evaluations of the fund's performance: Ang et al. (2009), Ang et al. (2014) Dahlquist and Ødegaard (2018) and Bauer et al. (2022).

way's Ministry of Finance (which is instructed by the Norwegian Parliament). The fund can thus be viewed as being owned by the people of Norway. The Ministry attempts hands-off management of the fund by limiting instructions to an investment mandate (Ministry of Finance, 2021). For our purposes, the most important part of this mandate is that the Ministry of Finance specifies a *target portfolio*, a weighted average of the developed world's stock markets, close to a world portfolio, together with a maximal allowable tracking error (the difference between the return of the target portfolio and the GPFG portfolio). This construction ensures that the fund should be thought of as a "near index fund"⁵ The mandate by the Ministry instructs the fund to have an active strategy that attempts to achieve returns above those of the target portfolio within specific risk limits.

Exclusions of companies from the fund's equity universe will lead to tracking error and are thus a cost for the GPFG. Exclusions still happen, though, and are the subject of this article. To understand where exclusions are coming from it is useful to consider some political issues. By adding equities to the GPFG asset mix, the Norwegian Parliament effectively became part-owners of thousands of companies world-wide. As an owner, one is arguably party to the actions of companies one owns. This can quickly become a political issue.

The first ethically motivated exclusion took place in 2002 of Singapore Tech, a producer of anti-personnel mines (Ministry of Finance, 2002). The first specific mention of Singapore Tech was raised in a discussion in the Parliament in 2001 by human rights organizations and Christian democratic and social-democratic political parties. Singapore Tech was the only company mentioned by name, but the broader discussion raised the question of a need to ensure ethical guidelines for the fund's investments. Up until then, the fund had no ethical guidelines impacting investment strategies. The question in 2001 was whether the investment in Singapore Tech was a direct breach of Norway's obligations towards human rights.

In the autumn of 2002, the Norwegian government appointed a public committee to propose ethical guidelines for the fund. The report concludes that although the question of participation raises difficult questions, the committee assumes that owning shares or bonds in a company that can be expected to commit gross unethical acts can be considered as complicity in these actions (Graver et al., 2003). In the revised national budget of 2004, ethical guidelines were established aligned with the recommendations in the report.

The Council on Ethics was established 19. November 2004. Its primary function is to

⁵Using standard classifications of mutual funds, (Dahlquist and Ødegaard, 2018, pg 91) shows that the GPFG's active share is so low that it would be classified as an index fund.

advise the Norges Bank on the observation and exclusion of companies from the fund. The ethical guidelines are determined by the Ministry of Finance and contain both product-based exclusions (currently including tobacco, cannabis, certain types of weapons, and coal), and conduct-based exclusions (currently including human rights abuses, environmental damage, unacceptable levels of greenhouse gas emissions, corruption, sale of weapons to specific states) (Etikkrådet (Council of Ethics), 2005). The threshold for exclusion is high. Only companies representing an unacceptable high future ethical risk to the fund are excluded. Notably, the opportunity to exercise ownership rights instead of exclusion could be a more suitable alternative to reduce the risk of continued norm violations. Thereby, the action to exclude is founded upon a discussion with the Fund (Ministry of Finance, 2021). The Ethical Council publishes their announcement after Norges Bank has agreed. The process provides the fund time to divest before the information is official and avoids unnecessary market movements during the divestment period.⁶ Through continued dialogue with the excluded firms, the Ethical Council can revoke the decision to exclude in the event of a change in operations for the excluded company.

The ethical guidelines were again revised in 2020 (Mestad et al., 2020). Based on their conclusions and the fund's response, the general consensus is that the guidelines have served their purpose (Norges Bank, 2020). Particularly the fund highlights that the ethical guidelines serve to reduce non-financial risk, as this type of risk cannot be diversified away. The revision suggests a further broadening of the exclusion criteria to reflect the developments in the last 15 years. An example is the inclusion of deadly autonomous weapons. Some discussions in the report and the answer from the fund are interesting. The report suggests the inclusion of a new criterion for excluding companies that sell military equipment to states that use this for serious and systematic violations of humanitarian law. Whereas Norges Bank agrees with the sentiment of the request, they highlight that the fund does not exclude countries but companies. Thereby, there must be broad and conclusive company evidence for such actions to make the guidelines effective. Otherwise, this type of exclusion will be based on foreign politics rather than individual company actions.

⁶The time frame Norges Bank has had to implement their selloff has varied. An early mandate for the ethical council (Etikkrådet (Council of Ethics), 2006, pg 9) explicitly gave Norges Bank two months to sell their stake before the exclusion was announced. This mention of an explicit time is no longer in the more recent mandates, it is now just specifying that the ethical council will make their announcement after Norges Bank's announcement of the divestiture — which clearly means the fund has ample opportunity to sell their stake before anything is public.

In addition to exclusions made by the Ethical Counsel, the fund conducts its own risk-based divestments NBIM (2020). These are divestments based on the fund ESG risk management. Risk-based divestments are not published, but the underpinnings of such decisions are transparent. These divestments will not be analyzed in our study.

3 Data and methods

The prime source of data is announcements from the Ethical council and GPFPG. From these announcements we construct a history of companies excluded, with the key dates those of the GPFPG news release. Throughout the 2005-2022 period, a total of 189 companies have been excluded for shorter or longer time periods. In Table 1 we break down the official reasons for exclusion. The majority of exclusion justifications are product-based, with production of coal the majority of these. The excluded stocks are distributed across 33 countries. The country with the largest number of exclusions is the US, with 52 exclusions. The next are China and India, with 16 and 13 exclusions, respectively.⁷

For the identified companies we gather stock market data from Refinitiv, daily prices and shares outstanding. We also gather exchange rates, from Yahoo Finance. Of the 189 excluded companies, we are able to match 186 stocks with Refinitiv data. Table 2 gives an overview of the sample. We note that of the 189 excluded firms, 26 have had their exclusion revoked, and again been allowed to enter the GPFPG portfolio. Note that the 189 firms is a very small number compared to the fund's investment universe, where the fund had almost 10 thousand different companies in their portfolio at yearend 2021. Exclusion is truly an exceptional reaction.

In Figure 1 we give an overview of the exclusions over time. As the figure shows, the number of exclusions have been gradual, with an exception of 2014. That is the year when the Fund introduces production or use of coal as a separate product-based exclusion reason. This coal-based exclusion is a bit special, in the sense that it was decided upon by the Fund, and only later ratified by the Ministry as an official reason for exclusion.

The basis for our analysis is monthly returns. In addition to the returns we calculate market capitalizations as the product of shares outstanding and closing prices. All returns and market capitalizations are denominated in dollars (USD). From these returns we construct portfolio returns. For each portfolio, a stock enter the portfolio the month after the company has been excluded. If an exclusion is revoked, the stock leave the portfolio at the end of the

⁷See Appendix ?? for detailed breakdowns by country, industry and year.

Table 1: Reasons for exclusions

Overview of the reasons for exclusions in the period 2005–2022. The reasons are grouped into two major reasons, conduct and product based. Data from the Ethical Council and GPFG.

Exclusion reasons	Events
Conduct	66
Environmental damage	28
Individuals' rights in war or conflict	11
Violation of human rights	12
Environmental damage / Violation of human rights	4
Violation of ethical norms	5
Greenhouse gas emissions	4
Gross corruption	2
Product	123
Coal or coal-based energy	75
Weapons	27
Tobacco	21

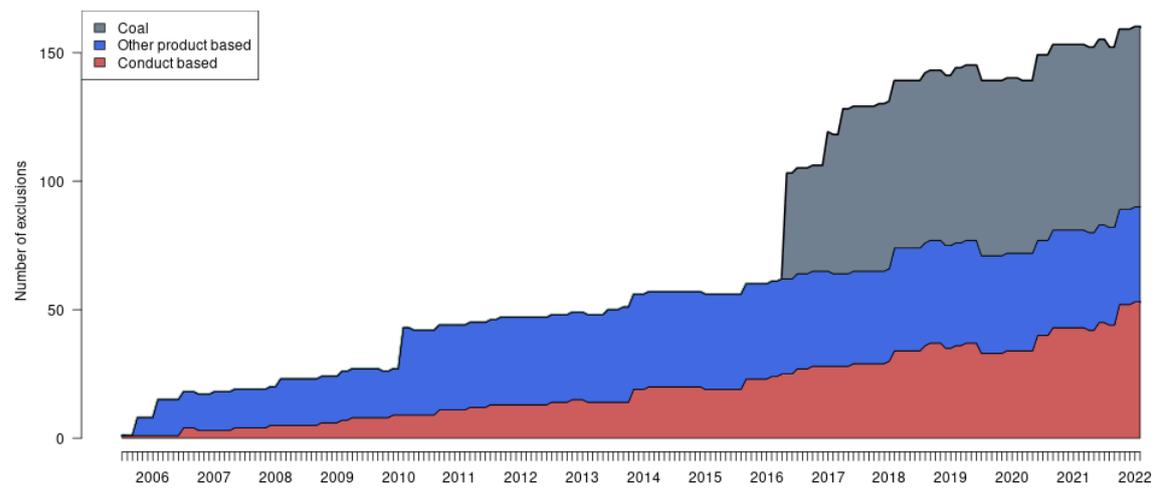
Table 2: Sample of stocks used in analysis

Overview of the exclusions, revocations and sample content. Data from the Ethical council, GPFG and Refinitiv.

Status	Events
Total exclusions	189
Exclusion revoked	26
Excluded again	1
Not matched with Refinitiv	3
Total sample	186
Conduct-based exclusions	66
Product-based exclusions	119

Figure 1: The time series evolution of the number of excluded shares

The figure shows the number of stock returns in the equally weighted exclusion portfolios, broken down by product-based and conduct-based. The product-based category is further broken down by coal-based and other product-based exclusions. Data from the Ethical council, GPF and Refinitiv.



month of exclusion. We use two methods to calculate portfolio returns: equally weighted and market weighted, using the market capitalization as weights.

3.1 Testing for performance

To measure portfolio performance we rely on the Fama-French international five factor model (Fama and French, 2017),⁸.

$$(r_{p,t} - r_{f,t}) = \alpha + \beta(r_{m,t} - r_{f,t}) + b^{SMB} SMB_t + b^{HML} HML_t + b^{RMW} RMW_t + b^{CMA} CMA_t + \varepsilon_{p,t}$$

where the factors are international versions of the corresponding US factors (Fama and French, 2015). To show robustness, we will also report a number of alternative formulations, including a one factor (CAPM), three and four factor specifications using the Ken French Global factors.⁹ For some of the analysis we alternatively employ Ken French's US factors.

4 Portfolio results

The exclusion portfolios represent the expected returns of stocks with low ESG ratings. We ask to what degree these returns can be explained by today's most common asset pricing model, the Fama French five factor specification.

4.1 Equally weighted portfolios

We first investigate the returns to the equally weighted portfolio. A simple, intuitive way to compare returns of two portfolios is to plot their cumulative returns, calculated as $CR_{p,T} = \prod_{t=1}^T (1 + r_{pt})$, where r_{pt} is the return for the portfolio p at time t .

In Panel A of Figure 2 we compare the evolution of the equally weighted exclusion portfolio with a global market portfolio. The exclusion portfolio clearly outperforms the market portfolio over the period. One observation is worth making from this picture. At the two large crises in this period, the '08 global financial crisis and the '20 Covid crisis, the fall in the

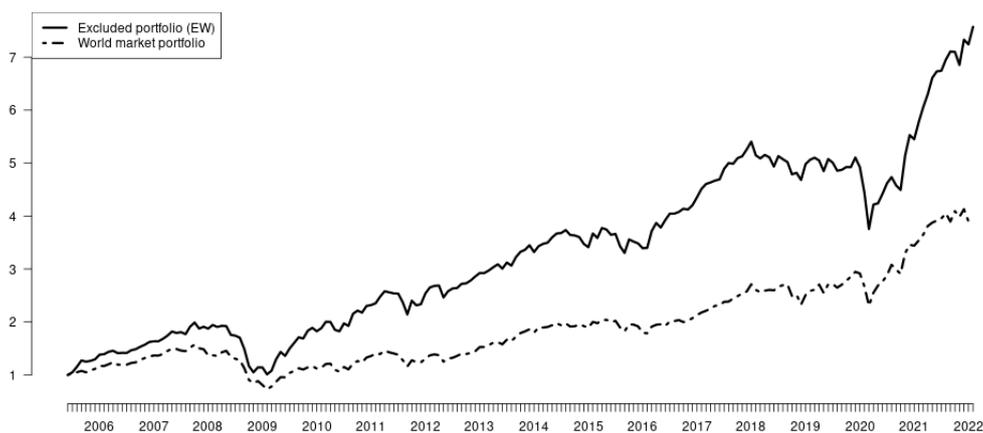
⁸See Dahlquist et al. (2015) and Dahlquist and Ødegaard (2018) for a discussion of relevant performance measurement for a fund like GPF.

⁹The factors are downloaded from Ken French's homepage. We are grateful for him for making the data available to the research community.

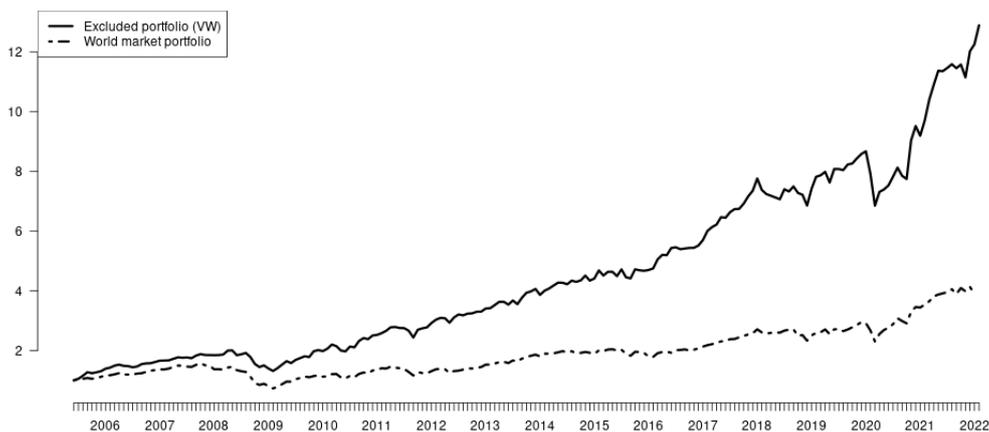
Figure 2: Cumulative returns of the exclusion portfolios and a global market portfolio

The figures shows the cumulative returns from two investments: The equally weighted exclusion portfolio (black line), and the world market portfolio provided by Ken French (broken line). Cumulative returns are calculated as $CR_{p,T} = \prod_{t=1}^T (1 + r_{p,t})$, where $r_{p,t}$ the monthly portfolio return.

Panel A: Equally weighted portfolio



Panel B: Value weighted portfolio



exclusion portfolio seem to be more prominent. This observation can be linked to some observations from the research literature. Lins et al. (2017) shows that high-quality ESG firms performed better during the '08 Financial Crisis, and Albuquerque et al. (2020) makes the similar observation at the onset of the Covid-19 crisis in March '20. As the exclusion portfolio contains low-quality ESG firms, from these results, we would expect them to perform worse, which is what we observe.

The comparison of cumulative returns of the equally weighted exclusion portfolio with the world market portfolio should, however, not be used to argue about expected return differences, it is merely an illustration.

To formally make a return comparison, one needs to account for risk differences through a performance estimation in the setting of an asset pricing model. Column (1) in Panel A of Table 3 reports estimates of the global five-factor Fama-French model. For our purposes, the key result is the alpha estimate, which is a positive, statistically significant alpha, in annualized terms 5.1%. Thus, the premium for “bad ESG” firms is more than 5%. The finding of a positive alpha is confirmed using the alternative asset pricing specifications in models (2)–(4) in the table, where the alphas vary between 4.2% and 5.6% in annual terms.

The table also reports estimates of the factor loadings. We note that the estimate of the “market beta” is below 1, the exclusion portfolio has lower systematic risk than the market. One obvious cause for this is a large number of coal companies in the exclusion portfolio. These companies are in the “Utilities” industry, with corresponding low betas.

4.2 Value-weighted portfolios

The equally-weighted portfolio above measures the expected return difference without regard to company size. Another approach is to think in terms of *economic importance*, a firm’s contributions to the economy. To measure this, we consider a value weighted exclusion portfolio, where the return of each excluded stock is weighted by the market capitalization of that equity. This portfolio is closer to measuring the economic importance of the exclusions.

Panel B of Figure 2 compares the evolution of the value weighted exclusion portfolio to a global market index. Comparing the value weighted with the equally weighted exclusion portfolio, the cumulative return on the value weighted exclusion portfolio is substantially higher. A similarity with the equally weighted portfolio is apparent during the market reverses in 2008 and 2020. The excluded stocks do worse than the general market, the flip side to the higher

Table 3: Asset pricing estimates of alpha for exclusion portfolio

Column (1) reports estimates of the regression $(r_{p,t} - r_{f,t}) = \alpha + \beta(r_{m,t} - r_{f,t}) + b^{SMB}SMB_t + b^{HML}HML_t + b^{RMW}RMW_t + b^{CMA}CMA_t + \varepsilon_{p,t}$, where $r_{p,t}$ is the return of the exclusion portfolio, $r_{f,t}$ the risk free rate, SMB , HML , RMW , CMA and WML the Ken French factors. Column (2) estimates the one factor CAPM $(r_{p,t} - r_{f,t}) = \alpha + \beta(r_{m,t} - r_{f,t}) + \varepsilon_{p,t}$, (3) estimates of the regression three-factor regression $(r_{p,t} - r_{f,t}) = \alpha + \beta(r_{m,t} - r_{f,t}) + b^{SMB}SMB_t + b^{HML}HML_t + \varepsilon_{p,t}$, and (4) the four-factor regression $(r_{p,t} - r_{f,t}) = \alpha + \beta(r_{m,t} - r_{f,t}) + b^{SMB}SMB_t + b^{HML}HML_t + b^{MOM}MOM_t + \varepsilon_{p,t}$. The equally weighted portfolio constructed from shares excluded from the GPFG. Data for 2005–2022. The international asset pricing factors are from Ken French's data page. Standard errors are Newey-West adjusted. Significance levels are indicated as: * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$.

Panel A: Equally weighted exclusion portfolio

	(1)	(2)	(3)	(4)
alpha	0.0043** (0.0017)	0.0036** (0.0017)	0.0042*** (0.0015)	0.0047*** (0.0017)
Rm-Rf	0.8990*** (0.0388)	0.9664*** (0.0492)	0.9407*** (0.0462)	0.9154*** (0.0459)
SMB	0.1507 (0.1045)		0.1656 (0.1153)	0.1651 (0.1130)
HML	0.4689*** (0.1272)		0.2791*** (0.0668)	0.2075** (0.0848)
RMW	0.1509 (0.1378)			
CMA	-0.3331 (0.2453)			
WML				-0.1144 (0.0830)
Observations	199	199	199	199
Adjusted R ²	0.8078	0.7863	0.8042	0.8077

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Panel B: Value weighted exclusion portfolio

	(1)	(2)	(3)	(4)
alpha	0.0056*** (0.0015)	0.0072*** (0.0016)	0.0071*** (0.0017)	0.0070*** (0.0016)
Rm-Rf	0.8427*** (0.0351)	0.7767*** (0.0326)	0.7862*** (0.0334)	0.7947*** (0.0337)
SMB	-0.3307*** (0.1077)		-0.4275*** (0.1126)	-0.4273*** (0.1102)
HML	0.1711* (0.0943)		0.2493*** (0.0740)	0.2736*** (0.0952)
RMW	0.2978** (0.1353)			
CMA	0.3437** (0.1373)			
WML				0.0387 (0.0620)
Observations	199	15 199	199	199
Adjusted R ²	0.7805	0.7325	0.7703	0.7699

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

resilience of high quality ESG stocks during these periods.

Again, to make a formal performance statement for the value weighted portfolio we perform performance regressions, reported in Panel B of Table 3. As one would suspect from the cumulative return figure, the alpha estimates are higher for the value weighted portfolio than the equally weighted one. In annual terms, the alpha in the five factor model is 6.9%, or almost 7%.

4.3 Return differences by exclusion reason?

The fund excludes companies for different reasons, with the main distinction being conduct and product based exclusions. To investigate differences in reasons for exclusion, we repeat the previous regressions separately for conduct and product based exclusion portfolios.

In Table 4 we report regression results for the two subsamples, using both equally and value weighted portfolios. In either case, we find that the alphas of the conduct based portfolios are double those of the alphas for the product based portfolios.¹⁰

4.4 The US portfolio

We finally look at a subsample using only stocks listed in the US. This because the US market is the most commonly studied single market, and we want to facilitate direct comparisons with studies on the US market. We therefore calculate exclusion portfolios (both equally weighted and value weighted) using only the 49 stocks with an US listing.

Figure 3 gives some descriptives for the US exclusion portfolios. Panel A shows the time series evolution of the number of shares in the portfolio. In the period 2006-2013 the portfolio contained between 10 and 20 stocks, a number that jumped to almost 40 in 2016, with a large number of coal-related exclusions. Panel B plots the cumulative returns for the US exclusions portfolio and compares them to a US index, the S&P 500 index (not the world index shown earlier). We are, however, observing the same pattern. The exclusion portfolios generally have superior returns to the market index but with marked larger falls during the '08 and '20 crises.

Table 5 shows the results of estimating a Fama French five factor model (Fama and French, 2015) for the US exclusion portfolios. We again find highly significant alpha estimates, which

¹⁰In the appendix we show cumulative return plots, where we show that it is particularly the last few years that seem to be driving the higher alpha estimates for the conduct based portfolio.

Table 4: Asset pricing estimates of alpha for exclusion portfolios by reason for exclusion

Estimates of the regression $(r_{p,t} - r_{f,t}) = \alpha + \beta(r_{m,t} - r_{f,t}) + b^{SMB}SMB_t + b^{HML}HML_t + b^{RMW}RMW_t + b^{CMA}CMA_t + \varepsilon_{p,t}$, where $r_{p,t}$ is the return on the exclusion portfolio. We consider two different samples of exclusion portfolios: The stocks excluded based on conduct, or based on product. For each of these samples we calculate equal or value weighted portfolios, for which we estimate the regression. The international factors are from Ken French's homepage. Standard errors are Newey-West adjusted. Significance levels are indicated as: * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$.

	ew conduct	ew product	vw conduct	vw product
	(1)	(2)	(3)	(4)
alpha	0.0068* (0.0038)	0.0030 (0.0031)	0.0089*** (0.0015)	0.0040*** (0.0015)
Rm-Rf	0.9927*** (0.1191)	0.8625*** (0.0757)	0.7872*** (0.0363)	0.8985*** (0.0353)
SMB	0.0877 (0.2849)	0.1429 (0.2539)	-0.2711** (0.1250)	-0.2997** (0.1232)
HML	0.9090*** (0.2200)	0.3099* (0.1640)	0.2905** (0.1178)	0.1958* (0.1064)
RMW	0.1089 (0.3398)	0.2053 (0.2820)	0.4066** (0.1711)	0.2966 (0.2109)
CMA	-1.2974*** (0.4191)	-0.0076 (0.2423)	0.2971* (0.1689)	0.2648* (0.1603)
Observations	199	196	199	196
Adjusted R ²	0.5712	0.7585	0.3699	0.7227

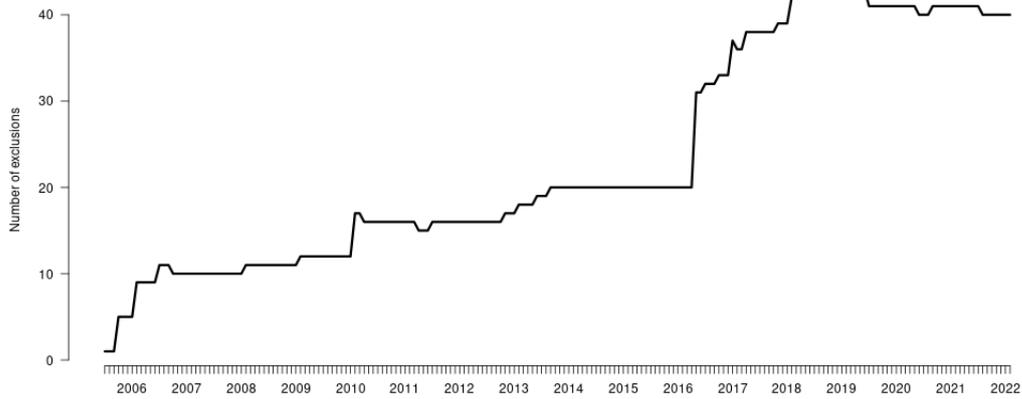
Note:

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Figure 3: The US exclusion portfolios

The figures summarize the US part of the exclusion portfolio. Panel A: The number of stocks in the US exclusion portfolio. Panel B: Comparison of cumulative returns, are calculated as $CR_T = \prod_{t=1}^T (1 + r_{pt})$, where r_{pt} is the monthly portfolio return.

Panel A: Number of exclusions



Panel B: Cumulative returns

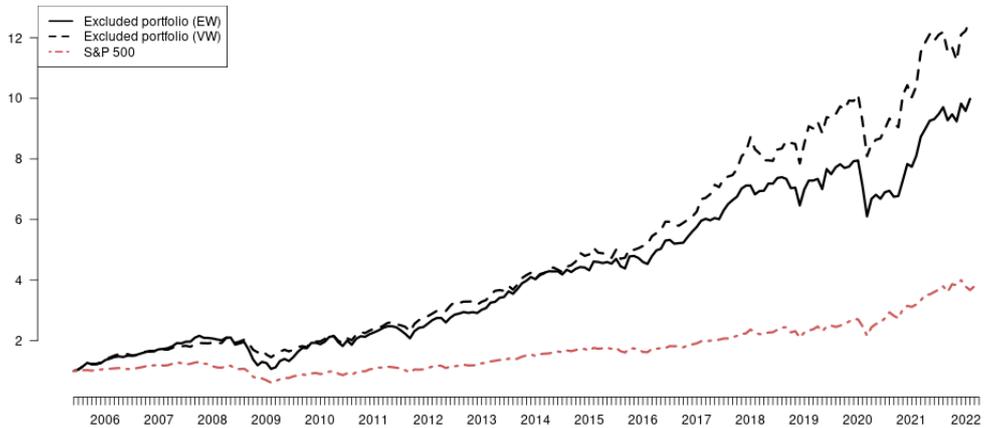


Table 5: Alpha estimation, US portfolio

Estimates of the regression $(r_{p,t} - r_{f,t}) = \alpha + \beta(r_{m,t} - r_{f,t}) + b^{SMB}SMB_t + b^{HML}HML_t + b^{RMW}RMW_t + b^{CMA}CMA_t + \varepsilon_{p,t}$, where $r_{p,t}$ is the return on the exclusion portfolio, and the others are the US factors provided by Ken French. We only consider stocks with a US primary listing, for which we calculate equal or value weighted portfolios. Standard errors are Newey-West adjusted. Significance levels are indicated as: * $p < 10\%$, ** $p < 5\%$, *** $p < 1\%$.

	Equally weighted (1)	Value Weighted (2)
alpha	0.0046** (0.0019)	0.0058*** (0.0019)
Rm-Rf	0.9108*** (0.0625)	0.7607*** (0.0421)
SMB	-0.0327 (0.1023)	-0.2819*** (0.0757)
HML	0.2018** (0.0785)	0.1528** (0.0764)
RMW	-0.0580 (0.1382)	0.2104* (0.1151)
CMA	0.1054 (0.1484)	0.1854* (0.1106)
Observations	200	200
Adjusted R ²	0.7095	0.6378

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

are actually even higher than for the whole portfolios, with annualized alpha estimates of 5.7% for the equally weighted and 7.2% for the value weighted exclusion portfolio.

4.5 Some perspectives on the findings

If we now take stock of the findings. We have shown clear evidence that the slightly less than 200 stocks excluded from the GPFG have superior returns (alpha). Comparing different reasons for exclusion, the stocks excluded for reasons of conduct have higher alphas than product-based exclusions. Also, the alpha estimates are higher for US stocks.

Let us now discuss this relative to our earlier theoretical discussion. We considered two potential reasons for the exclusion portfolio(s) to have higher returns: Short-term price pressure leading to temporary underpricing or a change in expected returns for low-quality ESG firms.

If we start by considering the first possibility, that exclusion events lead to downward pressure on stock prices, and returns going forward will correct this temporary mispricing. An argument against this being the complete story is the sheer magnitude of the estimated alphas. We are doing estimates over a 17-year period and find estimates of 5% (or more) extra annual return over the period. This can not be driven purely by correcting temporary mispricing. This argument can also be made with reference to the earlier mentioned event studies of Atta-Darkua (2020) and Eriksen et al. (2020). Both of these look at the announcement returns and find that stocks of excluded firms fall by about 1.4% (Atta-Darkua) or 0.3–0.5% (Eriksen et al.). It can be argued that these estimates are under-estimating the total effect of the GPFG's exclusion, as they only look at the public announcement of the exclusion, which happens after the GPFG has realized their holdings. A price pressure due to the Fund's sell-off will have occurred before the announcement, and the total price-pressure would be the sum of this pre-announcement effect and the post-announcement effect reported in these two studies. If we assume that the pre-announcement effect is of a similar magnitude as the post-announcement effect, the largest estimate of the total price pressure effect would be double the 1.4% Atta-Darkua estimate or 2.8%. The price-pressure effect is a one-time effect.

Thus, there is no way the price-pressure effect can be the only explanation for our alpha estimate of 5% per year, and we are left with concluding that much of the alpha must be due to changes in long term expected returns.

At this stage, it behooves us to come back to something we promised earlier, discussing

how our results tally with earlier research on the GPFG. We have already discussed the event study results, but we also need to consider the study that is closest to our, Hoepner and Schopohl (2018), which does a similar construction to ours. They also construct value weighted and equally weighted portfolios representing the exclusions by the GPFG, and estimate alpha. They find an estimate of the alpha lower than ours and not significant. There are, however, some differences. Their sample period is shorter, the sample period stopping in 2015. We have looked at this and done our estimations for the same subperiod.¹¹ Even for the same subperiod we are still estimating a significantly positive alpha. There are, however, some differences in research design that may cause the differences. In particular, they update portfolio compositions on an annual basis, unlike our use of the exact date when exclusions are announced. Their sample period thus includes the time of the negative return estimated in the event studies of Atta-Darkua (2020), and Eriksen et al. (2020), which may bias their estimates downward. We are also using the Fama-French five factor model in our asset pricing investigations, which is preferred to their use of the one-factor CAPM or four-factor models. In conclusion, we believe our results are conceptually more correct. Also, we examine various exclusions' portfolios, including the aggregate portfolio, the portfolios grouped by exclusion reasons, the US-based portfolios, using longer times series. Thus, we have a fuller picture of how the exclusions' portfolios perform based on longer time series and a larger sample of exclusions.

4.6 Robustness

Let us finally mention some of the analyses we have performed to investigate the robustness of our results.

First, we have looked at the robustness of the timing of when stocks enter the exclusion portfolio. Currently, we let stock enter the exclusion portfolios the month after the exclusion. We have also done the estimations including the month of the exclusion, without seeing any major changes in the alpha estimates.

We have also looked at the results when splitting the estimation period into two subperiods, 2005–2015 and 2016–2022 (See the appendix). We find that in the later period, the alpha estimates are still positive but lower and not always significant. We, however, note that this period only contains six years, which means the sample period is relatively short.

Finally, in the value weighted portfolio there is one company, Walmart, that has a very large

¹¹See Appendix Table ??.

weight in the portfolio in the early part of the period. We have therefore redone the analysis removing Walmart from the value weighted portfolio. This does not change our inferences.

5 Conclusion

We argued that current theoretical models of how ESG considerations affect equilibrium stock returns would lead to differences in expected returns linked to ESG ranking. We used the exclusions by the Norwegian Government Pension Fund Global, the world's largest SWF, to identify a set of firms with a low ESG ranking.

Applying a battery of performance tests to portfolios of these firms, we establish that these portfolios have a considerable excess return (alpha) relative to the predictions of standard asset pricing models, such as the global Fama-French five-factor model. The portfolios of these stocks have statistically significant positive excess returns (alpha) as high as 5% in annual terms. When we compare different reasons for exclusion, the stocks excluded for reasons of conduct have higher returns than product-based exclusions. Also, the alphas are even higher for the portfolio of only US-listed stocks.

We discussed two possible theoretical approaches. First, a short-term price pressure hypothesis, that exclusions lead to selling pressure and a short-term underpricing, which will be corrected going forward, and lead to higher returns in the short-term, or second, a shift in expected returns, low quality ESG firms have higher returns. We argued that the sheer magnitude of the return difference (5%) rules out short-term price pressure as a complete explanation and refers to the estimates of the one-time shock to stock prices at the time of exclusion announcement (1.5% or lower).

We are left to conclude that our results indicate that low quality ESG firms have a return premium. The question is how to incorporate this into models of asset pricing. If these ESG differences can be viewed as a source of systematic risk, an obvious approach would be to construct a "ESG factor" that could be added to the five factors of the Fama-French model. The stocks in the GPFG exclusion portfolio would presumably all load negatively on this factor, and we may find that our estimated positive alpha disappears. At the moment, we will not go in this direction, we just point to it as a promising venue for future research.

We will instead conclude by pointing out that the most important contribution of our research is to show the sheer magnitude of the return difference linked to ESG. Annual return differences as high as 5% (or 5.7% in the case of the US portfolio) will clearly be a challenge to

fit into a theoretical asset pricing framework.

References

- Rui Albuquerque, Yrjo Koskinen, Shuai Yang, and Chendi Zhang. Resiliency of Environmental and Social Stocks: An Analysis of the Exogenous COVID-19 Market Crash. *The Review of Corporate Finance Studies*, 9(3):593–621, 07 2020. doi:10.1093/rcfs/cfaa011.
- Andrew Ang, William N Goetzmann, and Stephen M Schafer. Evaluation of active management of the Norwegian government pension fund – global. Report available at the website of the Norwegian Ministry of Finance, December 2009.
- Andrew Ang, Michael W Brandt, and David F Denison. Review of active management of the Norwegian Government Pension Fund Global. Technical report, Report to the Norwegian Ministry of Finance, January 2014.
- Vaska Atta-Darkua. Corporate ethical behaviours and firm equity value and ownership: Evidence from the GPF’s ethical exclusions. Available at SSRN 3388868, 2020.
- Rob Bauer, Charlotte Christiansen, and Trond Døskeland. A review of the active management of Norway’s government pension fund global. SSRN working paper, January 2022.
- Roland Beck and Michael Fidora. The impact of sovereign wealth funds on global financial markets. *Intereconomics*, 43(6):349–358, 2008.
- David Blitz and Frank J Fabozzi. Sin stocks revisited: Resolving the sin stock anomaly. *The Journal of Portfolio Management*, 44(1):105–111, 2017.
- Patrick Bolton and Marcin Kacperczyk. Do investors care about carbon risk? *Journal of Financial Economics*, 142(2):517–549, 2021. doi:10.1016/j.jfineco.2021.05.008.
- Mark M Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82, March 1997.
- David Chambers, Elroy Dimson, and Antti Ilmanen. The Norway model. *The Journal of Portfolio Management*, 38(2):67–81, 2012.
- David Chambers, Elroy Dimson, and Antti Ilmanen. The Norway model in perspective. *The Journal of Portfolio Management*, 47(5):178–187, 2021. doi:10.3905/jpm.2021.1.230.
- Sudheer Chava. Environmental externalities and cost of capital. *Management Science*, 60(9):2223–2247, 2014.
- Guillaume Coqueret. Perspectives in ESG equity investing. SSRN, March 2021.
- Bradford Cornell. ESG preferences, risk and return. *European Financial Management*, 27(1):12–19, 2021. doi:10.1111/eufm.12295.

- Magnus Dahlquist and Bernt Arne Ødegaard. A Review of Norges Bank's active management of the Government Pension Fund Global. Technical report, Report to Norwegian Ministry of Finance, January 2018. URL <https://www.regjeringen.no/no/aktuelt/ekspertrapporter-om-spu/id2585465/>.
- Magnus Dahlquist, Christopher Polk, Richard Priestley, and Bernt Arne Ødegaard. Norges bank's expert group on principles for risk adjustment of performance figures - final report. Norges Bank (Central Bank of Norway) report, November 2015. URL <http://www.norges-bank.no/pages/104035/ExpertGroupFinalReportNov2015.pdf>.
- Kathryn L Dewenter, Xi Han, and Paul H Malatesta. Firm values and sovereign wealth fund investments. *Journal of Financial Economics*, 98(2):256–278, 2010.
- Robert B Durand, SzeKee Koh, and Paul Lijian Tan. The price of sin in the Pacific-Basin. *Pacific-Basin Finance Journal*, 21(1):899–913, 2013.
- Sondre Hansen Eriksen, Snorre Lindset, Quynh Trang Nguyen, and Marie Skara. Market reactions to ESG announcements: Evidence from a \$1 trillion fund. Available at SSRN 3640447, September 2020.
- Etikkrådet (Council of Ethics). Etikkrådet for Statens pensjonsfond – Utland – Årsmelding 2005, 2005.
- Etikkrådet (Council of Ethics). Etikkrådet for Statens pensjonsfond – Utland – Årsmelding 2006, 2006.
- Frank J Fabozzi, KC Ma, and Becky J Oliphant. Sin stock returns. *The Journal of Portfolio Management*, 35(1): 82–94, 2008.
- Eugene F Fama and Kenneth R French. Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51(1):55–85, 1996.
- Eugene F. Fama and Kenneth R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1 – 22, 2015.
- Eugene F. Fama and Kenneth R. French. International tests of a five-factor asset pricing model. *Journal of Financial Economics*, 123(3):441 – 463, 2017.
- Eugene F Fama and J MacBeth. Risk, return and equilibrium, empirical tests. *Journal of Political Economy*, 81: 607–636, 1973.
- Hans Petter Graver, Jarle Berge, Alexander Cappelen, and Ola Löhman. NOU 2003:22 Forvaltning for fremtiden: Forslag til etiske retningslinjer for Statens petroleumsfond, 2003.
- Andreas G F Hoepner and Lisa Schopohl. On the price of morals in market: An empirical study of the Swedish AP-funds and the Norwegian government pension fund. *Journal of Business Ethics*, 151:665–692, 2018.
- Harrison Hong and Marcin Kacperczyk. The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1):15–36, 2009. doi:10.1016/j.jfineco.2008.09.001.

- Jonathan Lewellen and Katharina Lewellen. Institutional investors and corporate governance: The incentive to be engaged. *The Journal of Finance*, 77(1):213–264, 2022. doi:<https://doi.org/10.1111/jofi.13085>.
- Hao Liang and Luc Renneboog. On the foundations of corporate social responsibility. *Journal of Finance*, 72(2):853–910, 2017. doi:[10.1111/jofi.12487](https://doi.org/10.1111/jofi.12487).
- Karl V. Lins, Henri Servaes, and Ane Tamayo. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, 72(4):1785–1824, 2017. doi:[10.1111/jofi.12505](https://doi.org/10.1111/jofi.12505).
- Ola Mestad, Johan Henrik Andresen, Egil Matsen, Astrid Undheim, Gro Nystuen, Olaug Svarva, Cecilie Hellestveit, and Svein Richard Brandtzæg. NOU 2020:7 Verdier og ansvar - Det etiske rammeverket for Statens pensjonsfond utland, 2020.
- Ministry of Finance. St.meld. nr. 2 (2001-2002) Revidert nasjonalbudsjett 2002, 2002.
- Ministry of Finance. Guidelines for Observation and Exclusion of companies from the Government Pension Fund Global (GPF), 2021.
- NBIM. Investing responsibly. The 20 year history. Norges Bank Investment Management, 2020.
- NBIM. Government pension fund global annual report 2021, 2021a.
- NBIM. The asset pricing effects of ESG investing, July 2021b. Discussion Note 01/2021, Norges Bank Investment Management.
- Norges Bank. Consultation response, NOU 2020:7 Values and Responsibility, 2020. URL <https://www.nbim.no/en/publications/submissions-to-ministry/2020/consultation-response-nou-20207-values-and-responsibility/>. Letter from Norges Bank to the Ministry of Finance.
- Lúboš Pástor, Robert F Stambaugh, and Lucian A Taylor. Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571, 2021. doi:[10.1016/j.jfineco.2020.12.011](https://doi.org/10.1016/j.jfineco.2020.12.011).
- Lasse Heje Pedersen, Shaun Fitzgibbons, and Lukasz Pomorski. Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2):572–597, 2021. doi:[10.1016/j.jfineco.2020.11.001](https://doi.org/10.1016/j.jfineco.2020.11.001).
- Michael L Ross. The political economy of the resource curse. *World politics*, 51(2):297–322, 1999.
- US SIF. 2020 report on US Sustainable, Responsible and Impact Investing Trends. https://www.ussif.org/store_product.asp?prodid=42, 2020.
- Tensie Whelan, Ulrich Atz, and Casey Clark. ESG and financial performance: Uncovering the relationship by aggregating evidence from 1,000 plus studies published between 2015 – 2020. Working Paper, NYU Stern Centre for Sustainable Business, 2021.