Essays on ESG Funds and Corporate Finance

by

Shodi Saliev

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2024

Doctoral Committee:

Professor Uday Rajan, Chair
Assistant Professor Shane Miller
Professor Gabriel Rauterberg
Assistant Professor Mirela Sandulescu
Assistant Professor Yongchen Zou

Shodi Saliev
ssaliev@umich.edu
ORCID iD: 0009-0002-8779-6584

In memory of my father, Tolib Shodievich Saliev.

ACKNOWLEDGEMENTS

First and foremost, I am deeply grateful to my advisor, Uday Rajan, for his infinite patience and support throughout the years. Uday has been extremely generous with his time, even when he was literally on the opposite side of the world. His immense knowledge and wisdom have helped me become a better researcher and human being. I am thankful to my finance committee members, Mirela Sandulescu and Shane Miller, for their invaluable support. Their feedback helped me tremendously to improve my work. I am also indebted to Gabriel Rauterberg, Yongchen Zou, Andrey Malenko, M.P. Narayanan, and Paolo Pasquariello for their invaluable advice and suggestions.

Last but not least, I would like to thank my fellow PhD students and friends for keeping me afloat and providing emotional support all these years, particularly Tamanna Singh Dubey, Thomas Flanagan, Emil Lakkis, Taeuk Seo, and Alex Wirth.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

ABSTRACT

This dissertation consists of three chapters. The first two chapters focus on ESG mutual funds, while the third chapter examines whistleblowers and fraud prevention in a corporate setting.

In Chapter 1, I examine the connection between fund fees, ESG performance, and risk-adjusted financial returns for active equity ESG mutual funds. In recent years, the growth of Environmental, Social, and Governance (ESG) mutual funds has been substantial, with a significant increase in both the number of funds and the amount of committed capital. Despite extensive research on the financial performance of ESG investments, the relationship between fund fees and performance within the ESG space remains underexplored. Our analysis reveals a counterintuitive finding: higher fund fees are associated with worse ESG performance, as measured by various ESG ratings. However, I find no negative relationship between fund fees and financial performance; on the contrary, higher fees are weakly correlated with better financial outcomes. Additionally, funds with higher ESG ratings tend to have superior risk-adjusted realized returns. The findings emphasize the complexities involved in ESG fund management and the need for investors to carefully evaluate ESG performance, financial returns and fee structures.

In Chapter 2, I examine the transition of traditional active mutual funds to sustainable or ESG-oriented funds. Using propensity matching and dynamic difference-in-difference methods, I find that converted funds indeed rebalance their portfolio towards firms with better ESG performance. Additionally I demonstrate that ESG fund conversions increase fund flows after about one and a half years after the treatment which might indicate the attractiveness of these funds to a growing base of environmentally and socially conscious

investors. In the subsample analysis, I demonstrate that increase in ESG ratings is primarily driven by low-cost funds which is in line with findings in Chapter 1. As for the fund flows, I find that there is a substantial heterogeneity between funds with high and low assets under management (AUM): high AUM non-ESG funds indeed have higher fund flows relative to the control group after conversion, while I do not observe significant effects for low-AUM funds. These findings contribute to the ongoing debate on the efficacy of ESG investing and provide insights for investors considering the transition to sustainability-focused strategies.

In Chapter 3, I examine how whistleblower compensation affects fraud disclosure and deterrence. My three-layered model has the firm with the CEO, the manager and the employee. The manager can steal from the firm, and the CEO, if informed, can expose him. Because of reputation concerns, she does not always disclose fraud, in which case the employee can blow the whistle. I show that, as the whistleblower reward increases, he is willing to disclose more fraud. This motivates the CEO to invest more in learning about fraud, however, the range of the CEO disclosed payoffs declines due to reputation costs. Because of that, the effect on ex-ante probability of fraud is ambiguous: depending on the distribution of cash flows, committed fraud might either increase or decline.

CHAPTER 1

Do Higher Fees Translate to Better ESG Outcomes? A Study of Active Equity ESG Funds

1.1 Introduction

In the past decade, the number of active equity Environmental, Social, and Governance (ESG) mutual funds has increased dramatically. Bloomberg (2022) reports that in 2021, the amount of capital in sustainable mutual funds and ESG-focused exchange-traded funds globally reached $2.7 trillion, representing a 53% growth from 2020. However, it remains unclear whether sustainable funds actually fulfill their promises regarding the sustainability of their investments. This uncertainty is particularly relevant given that ESG funds typically charge higher fees than their non-ESG counterparts, while the existing evidence on their financial performance is mixed.

The relationship between fees and ESG performance is especially relevant in light of widespread greenwashing concerns among investors and regulators. For instance, a report by the Security and Exchange Commission (2021) raised multiple concerns about the ESG products and services offered by investment advisers and mutual funds. The report cited *"...a lack of policies and procedures related to ESG investing; policies and procedures that did not appear to be reasonably designed to prevent violations of law, or that were not implemented;*

documentation of ESG-related investment decisions that was weak or unclear; and compliance programs that did not appear to be reasonably designed to guard against inaccurate ESG-related disclosures and marketing materials."

Numerous studies have examined whether investments aligned with ESG values deliver superior financial performance. This question has been explored at both the individual company level (Bolton and Kacperczyk, 2021, Baker et al., 2022, Sautner et al., 2023, among many others) and the fund level (Riedl and Smeets, 2017, Pástor et al., 2021, Van der Beck, 2021). For the latter, the existing literature indicates that ESG-conscious investors typically expect to earn lower returns, while the higher realized returns in recent years are primarily explained by fund flows. However, the relationship between fund fees and performance among ESG funds has received limited attention. In this chapter we examine the relation between fund fees, ESG performance and risk-adjusted before-fee returns. Unlike conventional or non-ESG funds, where the primary concern for investors is typically higher returns, for ESG-conscious investors fund performance consists of two different metrics: financial returns and ESG performance (e.g., Goldstein et al., 2022). While several studies have established a negative relationship between fees and risk-adjusted returns, indicating that higher fees often do not correspond to better performance (e.g., Gil-Bazo and Ruiz-Verdú, 2009, Ben-David et al., 2022), evidence on whether sustainable funds with higher fees actually deliver better ESG outcomes remain scarce.

In this chapter, we provide evidence on the relationship between fund fees, ESG performance, and risk-adjusted returns across the universe of all active equity ESG funds. First, we examine the relation between fund ESG performance and fund fees. Ex-ante, it is unclear whether higher fees are associated with better ESG performance. On the one hand, investors paying higher fees might expect superior outcomes, with ESG performance being a key component for funds explicitly marketed as sustainable. On the other hand, significant marketing efforts may target ESG-conscious, potentially less sophisticated investors, attracting them to higher-fee funds with mediocre ESG performance (Gil-Bazo et al., 2010).

Additionally, the limited attention of some investors could exacerbate this issue (e.g., Hirshleifer et al., 2011, Bailey et al., 2011, Kempf et al., 2017). Our findings indicate that higher fund expense ratios are associated with poorer ESG performance. Depending on the specification and measure of ESG performance, one standard deviation increase in fund annual expense ratio is associated with a decrease in ESG rating from 2 to 6 percent. To put this into perspective, one standard deviation increase in fund expense ratio is roughly equivalent to a jump from the 25th to the 75th percentile in our sample. The negative relationship between fees and ESG performance among active ESG funds highlights the potential inefficiencies and misalignments in the sustainable mutual fund market where higher fees do not necessarily translate to better ESG outcomes. This suggests that ESG-conscious investors need to scrutinize the fee structures and performance metrics of ESG funds carefully, as paying more does not guarantee superior ESG performance.

Second, we explore the relationship between financial performance and fund fees. In theory, since investors pay mutual fund fees for the services provided by the fund—one of which is portfolio management—fees should reflect the fund's risk-adjusted performance. However, previous literature suggests that mutual funds might face different demand curves, and those with less elastic demand curves may charge higher fees (e.g., Christoffersen and Musto, 2002, Gil-Bazo and Ruiz-Verdu, 2008). Contrary to these findings, our results do not indicate a negative relationship between fund fees and financial performance. We estimate monthly alphas based on the Fama-French three-factor and Carhart four-factor models. Across most specifications, we do not observe negative relationship between fees and alphas. One potential explanation is that, since most ESG funds have been launched over the last decade and are therefore significantly younger than conventional funds, investors have not fully adjusted to differences in fund performance as modeled in Berk and Green (2004), where investors are performance-sensitive and chase higher returns, even though the returns are not persistent (see also Sirri and Tufano, 1998). In other words, elasticity of demand curves for ESG funds might be less heterogeneous that the one for conventional funds as suggested in Christoffersen and Musto (2002).

Lastly, we examine the relationship between fund sustainability performance and risk-adjusted returns. Theoretically, imposing constraints on investment strategies could lead to poorer financial outcomes. However, this notion is influenced by changes in investor preferences (e.g., Pástor et al., 2021). If capital flows towards stocks more likely to be included in the portfolios of restricted funds, those funds may outperform their unrestricted counterparts. Our estimates indicate that a one standard deviation increase in the ESG rating is associated with an increase in monthly realized alpha by approximately 8 to 13 basis points, depending on the specification and risk-adjusted performance model used. However, the economic significance of this result is likely limited due to the high standard deviation of approximately 170 basis points in monthly alphas within our sample. Additionally, it is important to note the distinction between realized and expected returns. The estimates we present establish the association between realized returns and ESG performance, and the relationship between expected returns and sustainability is beyond the scope of this chapter.

To summarize, we find that funds with higher fees tend to have lower ESG ratings. This evidence underscores the importance of enhanced disclosure rules and other regulations in sustainable investing to ensure that mutual funds better adhere to their mandates. The apparent inability of sustainable fund competition to establish an adequate relationship between ESG performance and fees raises the question of whether improvements in fund governance could align fees more closely with the value that funds generate for investors. Additionally, we show that the established relationship in the finance literature between fund fees and risk-adjusted returns might not hold for the universe of ESG funds. One potential explanation is that, given ESG funds are significantly younger than traditional funds and that investors in ESG funds are likely less sensitive to poor financial performance, it might take more time for (sophisticated) investors to withdraw their money from underperforming funds and invest in better-performing ones. In other words, the explanation offered in Christoffersen and Musto (2002) might become relevant once performance-sensitive investors sort themselves into better-performing funds.

This chapter relates to three strands of literature. First, the chapter contributes to the rich literature on mutual fund fees. The determinants of mutual fund fees have been the subject of extensive investigation in the finance literature, reflecting the intricate interplay between market competition, fund performance, managerial skill, and investor behavior. For example, Khorana et al. (2009) find that fees are lower in more competitive markets, suggesting that market forces exert pressure on fund managers to reduce costs to attract and retain investors. This is corroborated by Elton et al. (2003), who demonstrate that superior past performance leads to lower fees due to economies of scale. As funds grow larger by attracting more assets under management, the average costs are reduced, allowing for lower fees while maintaining profitability. However, the relationship between performance and fees is not straightforward. Berk and Green (2004) propose a model where fund fees do not necessarily decrease with performance. Instead, they argue that in equilibrium, fund managers capture rents through higher fees, implying that high-performing funds can command higher fees due to their perceived value by investors. This perspective is supported by the idea that skilled managers are able to extract higher fees as a compensation for their superior performance, a notion further explored in Pástor et al. (2015). They show that managerial skill, proxied by the fund's alpha, is a significant determinant of fees, with more skilled managers charging higher fees. The role of investor behavior is another critical determinant of mutual fund fees. Barber et al. (2005) highlight that less informed investors are more likely to pay higher fees due to a lack of awareness or understanding of fee structures. This information asymmetry creates a market where fees can remain high despite poor performance. Christoffersen and Musto (2002) explore this by examining how mutual fund companies strategically set fees to maximize their revenue, often exploiting investor inertia and the tendency of investors to stick with their existing funds despite better alternatives. An important paper by Gil-Bazo and Ruiz-Verdú (2009) find, contrary to the traditional expectation that better-performing funds charge higher fees, a surprising negative relationship between fees and performance.

Specifically, they observe that funds with worse performance tend to charge higher fees, which is inconsistent with the notion that fees reflect the value provided by fund managers.

In recent years, the emergence of ESG funds has added another layer to the analysis of mutual fund fees. Gil-Bazo et al. (2010) find that ESG funds tend to have higher fees compared to conventional funds, which they attribute to the specialized screening and research processes required for ESG investments. Nofsinger and Varma (2014) find that investors are willing to pay a premium for funds that align with their ethical values, despite mixed evidence on performance relative to non-ESG funds. Finally, Riedl and Smeets (2017) provide evidence that socially responsible investors are less sensitive to fund fees and more focused on the ethical alignment of their investments, which can allow ESG funds to maintain higher fees. This chapter contributes to the literature by examining the relation between fund fees and fund performance. Specifically, we show that funds with higher fees tend to have lower ESG performance for multiple different measures of ESG ratings, and in contrast to some previous papers focusing on conventional funds, we do not find negative relation between funds financial performance and fund fees.

Second, this chapter relates to the literature on whether integrating ESG criteria into investment decisions influences financial performance. Friede et al. (2015) conducted a meta-analysis of over 2,000 empirical studies and found that the majority of studies reported a positive correlation between ESG criteria and corporate financial performance. In a similar vein, Edmans (2011) demonstrated that firms with high employee satisfaction, as an aspect of social responsibility, outperform their peers in terms of stock returns, suggesting that certain ESG factors can lead to superior financial performance. Moreover, Hong and Kacperczyk (2009) examined the so-called “sin stocks” and found that firms engaged in activities deemed socially irresponsible, such as tobacco and gambling, tend to yield higher returns due to investor aversion, implying a risk premium for such investments. In the context of ESG funds, Nofsinger and Varma (2014) found that socially responsible funds tend to perform better during periods of market turbulence, indicating that ESG funds may offer downside protection.

On the contrary, Revelli and Viviani (2015) performed a meta-analysis and reported mixed results regarding the risk-adjusted performance of socially responsible investments (SRI), highlighting that while some studies indicate a performance penalty, others show a neutral or positive impact. Additionally, Pástor et al. (2021) suggest that investor preferences for sustainability can lead to higher valuations of ESG-friendly firms, potentially resulting in lower expected returns as high valuations are usually associated with lower future return. In a related paper, Van der Beck (2021) show that in 2017-2022, ESG funds performance could be explained by higher flows, and in the absense of flows, ESG funds would not outperfrom the market. Albuquerque et al. (2019) propose that ESG practices can mitigate downside risk and create value for firms, particularly in high-stakes environments. This chapter looks at the relation between fund ESG performance and realized risk-adjusted alphas, and doesn't find negative relation between financial and ESG performance. If anything, we find that for some measures of ESG performance, funds with higher realized risk-adjusted alphas tend to have better ESG performance, while for other measures we do not observe statistically significant relation. However, it is crucial to remember that the estimates we present establish the association between realized returns and ESG performance, and the relationship between expected returns and sustainability is beyond the scope of this chapter.

Finally, there is a literature which discusses if ESG should be included as an additional factor. For example, Bolton and Kacperczyk (2021) provide evidence that ESG factors, particularly environmental performance, are priced into the market. Their study shows that firms with higher environmental ratings enjoy lower costs of capital, indicating that the market rewards sustainable practices. Berk and Van Binsbergen (2021) challenge the notion that ESG ratings significantly impact investment performance. Their empirical analysis reveals that while ESG ratings correlate with certain positive outcomes, the overall impact on investment returns is marginal when controlling for other firm characteristics. This skepticism highlights the complexity and potential limitations of using ESG ratings as a standalone factor. Our results at the fund level indicate that ESG performance of mutual funds might explain a part of funds excess returns, although, admittedly, our risk-adjusted alphas are derived using only Fama-French three-factor and Carhart four-factor models, and it remains to be seen if ESG factor has statistically significant explanatory power beyond the hundreds of factors proposed in the past (Feng et al., 2020).

The rest of the chapter is organized as follows. Section 1.2 describes the data and provides summary statistics. The empirical framework is explained in Section 1.3 while Section 1.4 presents the results. In Section 1.5, we discuss the findings. Finally, we conclude in Section 1.6.

1.2 Data

The data employed in this study are drawn from four primary sources: Morningstar, CRSP, MSCI, and Kenneth R. French's website. We use Morningstar's "Sustainable Funds U.S. Landscape Report 2022" in order to identify ESG equity mutual funds. The reports have been publishing by Morningstar since 2018, although the current Morningstar's website functionality allows to download only the latest edition of the report. The information published in the reports vary from year to year – for example, the 2022 report includes the information (names, tickers, AUM, returns) on all ESG funds active by the end of 2021, while the next edition of the report in 2023 reports only some aggregated statistics about performance of sustainable funds in general. In total there are 233 ESG funds by the end of 2021 in our sample. 58 out of those funds used to be non-ESG funds and had been repurposed as ESG funds by the end of 2021.

We also utilize data from the Center for Research in Security Prices (CRSP) pertaining to mutual fund holdings and characteristics. Specifically, for each mutual fund with holdings exceeding 100 million dollars, we have access to quarterly data encompassing management and expense ratios, assets under management (AUM), managing team, and numerous other attributes. CRSP also provides monthly data on fund holdings which we later use to calculate

ESG ratings for each fund. Additionally, the CRSP data includes monthly after-expense fund returns and the inception date of each fund, which, in combination with expense ratios, allows us to determine risk-adjusted performance and age of the fund, respectively. I use the CRSP data from January 2011 to September 2023.

Next, we use MSCI ratings for individual companies to estimate a fund's ESG rating from its holdings in a given month. Specifically, using fund holdings data from CRSP, I calculate the weighted average ESG rating of funds:

$$ESG\_rating_{it} = \sum_{j=1}^{N} ESG\_score_{jt} * weight_{ijt}, \tag{1.1}$$

where ESG_rating_{it} is the ESG rating of fund i at time t, ESG_score_{jt} is the MSCI ESG rating of company j at time t and $weight_{ijt}$ is the percent of assets of fund i invested in firm j at time t, i.e., firm j's weight in fund i's portfolio in month t.

The MSCI ESG rating is a weighted average of three primary components: environmental, social, and governance scores. Each score is derived from a range of underlying factors, and the specific factors and their respective weights can change over time. There are two different MSCI ESG ratings of individual firms: industry-adjusted and weighted-average ratings. I use both ratings for constructing ESG ratings of funds as well as individual scores for Environmental, Social, and Governance performance based on the weight of each company in a fund's portfolio in a given month. MSCI provides ESG ratings of individual firms starting from 1999. For fund ratings, I use data only from 2011 till March 2022 (the latest available data to the author) due to data quality issues before 2011.

MSCI is only one of several rating agencies, including Sustainalytics, S&P Global, Moody's ESG, KLD, and Refinitiv. Berg et al. (2022b) document the rating divergence, indicating that these different ratings often exhibit low correlations with each other. In this chapter, we use MSCI ESG ratings because they cover a significantly larger number of companies compared to other rating providers (e.g., in 2014, MSCI covered nearly twice as many companies as any other rating agency). This broader coverage allows us to include a higher fraction of fund holdings in our analysis.

The MSCI data does not contain *permno* or *permco* identifiers which are more stable than tickers. Consequently, we primarily rely on tickers, supplemented by company names, for matching MSCI company ratings with CRSP fund holdings. The matching process involves an initial step using tickers, followed by a verification step using company names in both the MSCI and CRSP datasets. For the latter, we employ the Damerau–Levenshtein distance comparison of company names, with a cutoff of 0.3 to determine the accuracy of the match. We exclude observations where we were able to match less than 40 percent of fund holdings, corresponding to the first percentile of matched holdings. On average we were able to match about 86 percent of fund holdings, although this number varies a lot by time: for example, for 2018 about 90 percent of holdings were matched which is a significant improvement from the 67 percent match rate in 2011.

Finally, from the website of Kenneth R. French, we use the data on factor returns (market, size, book-to-market, and momentum) to calculate funds' risk-adjusted returns based on Fama-French three-factor and Carhart four-factor models.

Table 1.1 presents the summary statistics of the sample. After excluding index funds and ETFs, our sample consists of 233 active equity ESG mutual funds. On average, these funds manage about $390 million, charge approximately 0.9 percent in annual expense fees, and deliver 65 basis points in monthly after-fee returns. However, the Fama-French three-factor and Carhart four-factor alphas are negative, at around minus 15 basis points monthly. The average ESG rating ranges from 3.58 to 4.51, depending on the ESG measure used.

1.3 Empirical Strategy

Mutual Fund Performance Estimation

We use Fama-French three-factor model (Fama and French, 1993) to estimate before-fee risk-adjusted performance:

$$r_{it} = \alpha_i + \beta_{rm,i} rm_t + \beta_{smb,i} smb_t + \beta_{hml,i} hml_t + \epsilon_{it}, \tag{1.2}$$

where r_{it} is fund i's before-expense return in month t in excess of the 30-day risk-free interest rate – proxied by 1-month Treasury bill rate; rm_t is the market portfolio return in excess of the risk-free rate; and hml_t and smb_t denote the return on portfolios that proxy for book-to-market and size risk factors, respectively. Note that CRSP data on mutual funds report after-expense returns and in order to retrieve monthly before-expense return, we add back the annual expense divided by 12 to reported returns.[1]

Following Carhart (1997), we implement a two-stage estimation procedure to obtain a panel of monthly risk-adjusted performance estimates. In the first stage, for every month t in years 2011 to 2023, we regress funds' before-fee excess returns on the risk factors over the previous 5 years. In case when less than 5 years of data is available for a specific fund-month, we require the fund to be in the sample at least for 36 months in the previous 5 years, and then run the regression with available data. In the second stage, we estimate a fund's risk-adjusted performance in month t as the difference between the fund's before-expense excess return and the excess return predicted by the Fama-French three-factor model.

The Relation between Fees and Performance

To examine the relation between fund fees and fund performance – either financial in the form before-fee risk-adjusted alphas or ESG in the form of ESG ratings, we estimate the

[1]This gives only approximation of before-expense returns since we ignore the compounding effect and also the fact that annual expenses are not necessarily evenly distributed across different months.

following pooled ordinary least squares (OLS) regression equation:

$$y_{it} = \lambda_t + \beta f_{it} + X_{it} + \varepsilon_{it}, \tag{1.3}$$

where f_{it} is the expense ratio of fund i in month t; X_{it} is fund characteristics, including age, fund flows and assets under management. The regression also includes time fixed effects λ_t to ensure that the estimated coefficient β captures cross-section relation between funds' expense ratio and the outcome (i.e., funds' alphas or ESG ratings), and not the potentially correlated time trends in those variables. The standard errors ε_{it} are clustered at the fund level in order to incorporate potential correlation of the errors for the same fund across time. y_{it} represent the outcome of interest: fund ESG or financial performance.

We are also interested in the relation between fund ESG performance and financial returns. In theory a fund with stricter ESG constraints should be more restricted in its investment strategies than a fund with less ESG constraints, resulting in lower returns for the former. To test this prediction, we implement the following simple estimation strategy:

$$\alpha_{it} = \gamma_t + \delta \times rating_{it} + X_{it} + \varepsilon_{it}, \tag{1.4}$$

where α_{it} is either Fama-French three-factor or Carhart four-factor alpha of fund i in month t; $rating_{it}$ is the ESG rating of fund i in month t; X_{it} is fund characteristics, including age, fund flows and assets under management; and γ_t represent month fixed effects. As in Equation (1.3), the standard errors ε_{it} are clustered at the fund level.

On the other hand, stricter ESG constraints might be correlated with quality of fund governance which, in turn, is correlated with better financial outcomes. Although we do not have data on fund governance quality, we split the sample based on ESG ratings to three subsamples – low ESG rating, medium ESG rating, and high ESG rating – and estimate Equation (1.4) separately for each of the subsamples.

1.4 Results

In this section we report the results from the estimation of Equations (1.3) and (1.4).

Fund ESG Ratings and Expense Ratio

We estimate Equation (1.3) to understand the relationship between fund expense ratio and ESG performance. The results are presented in Table 1.2 and Table 1.3.

In column (1) of Table 1.2, we estimate the relationship between industry-adjusted ESG score and annual expense ratio. The results indicate that a one standard deviation increase in the annual expense ratio is associated with a decrease in the industry-adjusted ESG score by about 0.16 points, or 3.6 percent. To put this into perspective, a one standard deviation increase in expense ratio is roughly equivalent to a jump from the 25th percentile to the 75th percentile in expense ratio. The other coefficients indicate that the fund ESG score is positively correlated with monthly returns, log of total net assets and the percentage of the portfolio invested in common stock. Although the corresponding coefficients (0.014, 0.061, 0.050) are statistically significant, they represent a very limited economic association with the fund ESG score. The coefficient for log of fund age is statistically indistinguishable from zero.

In columns (2)-(5), we estimate the same regression, but instead of industry-adjusted ESG rating, we consider weighted-average ESG rating, Environmental Fund Score, Social Fund Score, and Governance Fund Score, respectively. The results are qualitatively similar (although not statistically significant for Governance Fund Score) with the main coefficient of interest, β, ranging from about -0.460 to -0.141. Given that the standard deviation of the fund annual expense ratio is about 0.4, and the means of the ESG ratings in Table 1.1, we find that a one standard deviation increase in the expense ratio is associated with approximately 2, 4.1, and 1.8 percent decreases in weighted average ESG rating, Environmental, and Social fund scores, respectively. The effect on Governance fund score is statistically insignificant, although the point estimates indicate a decrease of about 1.5 percent in Governance Score associated with a one standard deviation increase in expense ratio.

In Table 1.3, we repeat the same analysis without including any controls except for month fixed effects. Across all specifications, the estimates remain qualitatively similar and become somewhat larger in absolute terms than those from Table 1.2. The results indicate that a one standard deviation increase in annual expense ratio is associated with about 4.7, 3.1, 5.4, and 3 percent decreases in industry-adjusted ESG rating, weighted average ESG rating, Environmental, and Social scores, respectively.

We also repeat the same analysis, but this time including squared terms for all control variables (see Appendix A). The estimates remain similar, although they cease to be statistically significant when we include the squared expense ratio.

Overall, the estimates of the expense ratio coefficient, β, indicate that, surprisingly, ESG funds with higher expense ratios tend to have worse ESG performance, as measured by various fund ESG ratings.

Financial Performance and Expense Ratio

Next, we explore the relationship between financial performance and fund fees for ESG funds. Tables 1.4 and 1.5 present the results for Fama-French three-factor alphas, while Tables 1.6 and 1.7 use Carhart alphas.

In column (1) of Table 1.4, we run regression 1.3 without controlling for any fund characteristics or fixed effects and find that the coefficient of interest, β, is statistically indistinguishable from zero. The corresponding R-squared is very low, consistent with the absence of fixed effects. In column (2), we include month fixed effects, and in column (3), we further control for fund fixed effects. Both coefficients remain statistically insignificant.

In Table 1.5, we include fund characteristics such as turnover ratio, total net assets, fund age, and the percentage of the portfolio invested in equities. The estimates in columns (1) and (2) are statistically insignificant. The last column shows a statistically positive coefficient, indicating that a one standard deviation increase in monthly expense ratio is associated with an increase in monthly Fama-French three-factor alpha by about 6 basis points. Given that the standard deviation of the Fama-French alpha is about 170 basis points, this estimate has little economic significance. When we include fund fixed effects instead of fund type fixed effects, the coefficient ceases to be statistically significant, although it remains positive.

Tables 1.6 and 1.7 report the results for alphas calculated based on the Carhart four-factor model. The only statistically positive coefficients (0.948 and 1.733) are in the last columns of both tables when we include both month and fund type fixed effects. The economic significance of these coefficients seems limited, as a one standard deviation increase in expense ratio is associated with a 3 to 5.8 basis point increase in monthly Carhart alpha, while the latter has a standard deviation of about 172 basis points. Moreover, the statistical significance of the coefficients does not survive the inclusion of fund fixed effects.

One important consideration is that, ideally, since alphas are estimated from a regression, the standard errors in the regressions should be adjusted using the Shanken (1992) correction (see also Chapter 13 in Cochrane, 2009). Given that this adjustment typically increases the standard errors, it is even less likely that our previously insignificant estimates will become significant.

Overall, in contrast to the existing literature on fund fees and performance for conventional funds, we do not find a negative relationship between fund fees and financial performance. If anything, we observe that funds with higher fees might have better financial outcomes. We discuss the potential mechanisms and implications of this finding in the next section.

ESG Performance and Financial Performance

In theory, an investment strategy with more constraints is likely to generate inferior returns compared to a less constrained one. This subsection tests this hypothesis and reports the results.

In Table 1.8, we analyze the relationship between industry-adjusted fund ESG ratings and alphas. Across all specifications, we find that ESG funds with higher ratings tend to perform better financially. Column (1) includes controls for monthly expense ratio, total net assets, age, and the percentage of the portfolio in common stock but no fixed effects. Column (2) adds month fixed effects, and column (3) further includes fund type fixed effects. The results consistently show that a one standard deviation increase in industry-adjusted fund rating is associated with an increase in Fama-French three-factor alpha by 11.7, 14, and 8.7 basis points, respectively. Columns (4)-(6) present estimates for Carhart four-factor alphas, yielding similar results: 11.2, 13.7, and 8.3 basis points for the specifications without fixed effects, with month fixed effects, and with both month and fund type fixed effects, respectively.

Tables 1.9 - 1.12 extend this analysis to other measures of ESG performance, including weighted-average ESG rating, Environmental Score, Social Score, and Governance Score. The results are largely consistent, with the exception that the estimates for Social and Governance Scores lose statistical significance when both month and fund type fixed effects are included (columns (4) and (6) in Tables 1.11 and 1.12). This may be due to the greater complexity and variability in measuring social and governance performance compared to environmental performance, which often relies on more objective metrics, such as emissions data reported by the Environmental Protection Agency. Consequently, the Environmental Score may contain less noise and serve as a more accurate measure of ESG performance.

Overall, the hypothesis that incorporating ESG objectives into investment strategies inevitably leads to poorer financial performance is not supported by our findings. If anything, the funds with higher ESG ratings seem to have better realized financial outcomes. However, as we mentioned ealier, it is crucial to remember the difference between realized and expected returns. Our estimates establish the relationship between realized returns and ESG outcomes, but they do not provide insights into the relationship between expected returns and sustainability.

1.5 Discussion

In this section we discuss the results described earlier. Contrary to intuitive expectations, our results indicate that higher fund fees are associated with worse ESG performance, yet they do not negatively impact financial returns. One possible mechanism behind the negative relationship between fund fees and ESG performance is the varying levels of investor sophistication. Funds with higher fees might attract less sophisticated investors who may not scrutinize the actual ESG impact of their investments rigorously. These investors might be more influenced by marketing efforts and the perception of higher fees equating to higher quality. As a result, fund managers could be capitalizing on these marketing dynamics without necessarily delivering superior ESG outcomes. Another explanation could involve the prevalence of greenwashing within the ESG fund sector. Greenwashing refers to the practice of funds exaggerating their ESG credentials to attract investment. The SEC's 2021 report on ESG practices highlighted significant deficiencies in the policies, procedures, and documentation related to ESG investing. These regulatory gaps might allow funds to charge higher fees without genuinely enhancing their ESG performance, contributing to the observed negative relationship.

An important caveat to the finding of a negative relationship between ESG performance and fund fees is the variety of approaches to ESG investing. For example, Broccardo, Hart and Zingales (2022) explore conditions under which voice is a more efficient strategy than exit. Similarly, Berk and Van Binsbergen (2021) argue that socially conscious investors should invest and use their shareholder rights to influence and change corporate policy, rather than divesting. High-cost ESG funds with mediocre ESG performance might engage in costly corporate activism to improve the ESG performance of a firm, rather than divesting from firms with low ESG ratings. This approach can incur significant costs, which might explain higher fees without immediate reflection in ESG performance metrics. To test this hypothesis, one could examine shareholder voting records to determine if funds with higher fees and lower ESG ratings vote more pro-ESG than other funds, i.e., whether they are actively using their influence to push for better ESG practices. Such an analysis would shed light on whether the higher fees are justified by the fund's engagement activities, which might not be immediately visible in the fund's ESG ratings but could lead to long-term improvements in ESG performance.

The existing finance literature (e.g., Gil-Bazo and Ruiz-Verdú, 2009, Ben-David et al., 2022) established that funds with higher fees tend to have poorer financial performance. In our sample we do not observe this pattern. The weak positive correlation between higher fees and financial performance could be attributed to the age and demand dynamics of ESG funds. ESG funds are relatively younger than their conventional counterparts, and investors in these funds might exhibit different sensitivities to performance. According to the Sirri and Tufano (1998), investors tend to chase higher returns. However, in the ESG fund universe, this chasing behavior might be less pronounced, given that these investors also value non-pecuniary benefits. This dynamic could allow younger ESG funds to charge higher fees while delivering slightly better financial outcomes without significant investor backlash.

Although imposing more restrictions on the investment strategies mechanically worsens[2] the outcomes, it is not clear if that holds in practice. For example, **Pástor et al. (2021)** show that ESG funds have higher realized returns. We augment their findings by showing that among ESG funds better ESG ratings are associated with higher realized returns. One caveat in this analysis is that ESG funds with higher ESG ratings might be just better funds: for example, in terms of ability of the portfolio manager/managing team or governance structure. Improved governance structures might enhance the fund's ability to navigate ESG criteria effectively while maintaining strong financial performance. Another caveat is that one should be cautious interpreting this relation between ESG performance and returns as a factor for choosing an ESG fund. Indeed, if there is a temporary shock in investors preferences which leads to higher capital flows to higher ESG companies, then funds with better ESG ratings mechanically have higher returns. In other words, the observed relation between

[2]To be mathematically precise, we should say "does not improve".

ESG and realized financial performance might be the result of shift in investors preferences and not the better stock-picking ability of the funds. Hopefully, future research will provide insights into the relationship between expected financial returns and the incorporation of ESG values.

1.6 Conclusion

In this chapter we examine the relationship between fund fees, ESG performance, and risk-adjusted financial returns in the universe of active equity ESG mutual funds. Contrary to the intuitive expectation that higher fees should correspond to better ESG performance due to the specialized screening and research processes required for ESG investments, our findings reveal that higher fund fees are associated with worse ESG performance. Depending on the specification and measure of ESG performance, one standard deviation increase in fund annual expense ratio is associated with a decrease in ESG rating from 2 to 6 percent. This result is robust across different measures of ESG performance, including industry-adjusted and weighted-average ESG ratings, as well as environmental, social, and governance scores.

Our analysis suggests that this negative relationship between fees and ESG performance may be influenced by several factors. One possible explanation is that funds with higher fees may attract less sophisticated investors who are less vigilant about the actual ESG impact of their investments, thereby allowing these funds to charge higher fees without delivering superior ESG performance. Additionally, marketing efforts may play a role in attracting ESG-conscious investors to higher-fee funds that do not necessarily provide better ESG outcomes.

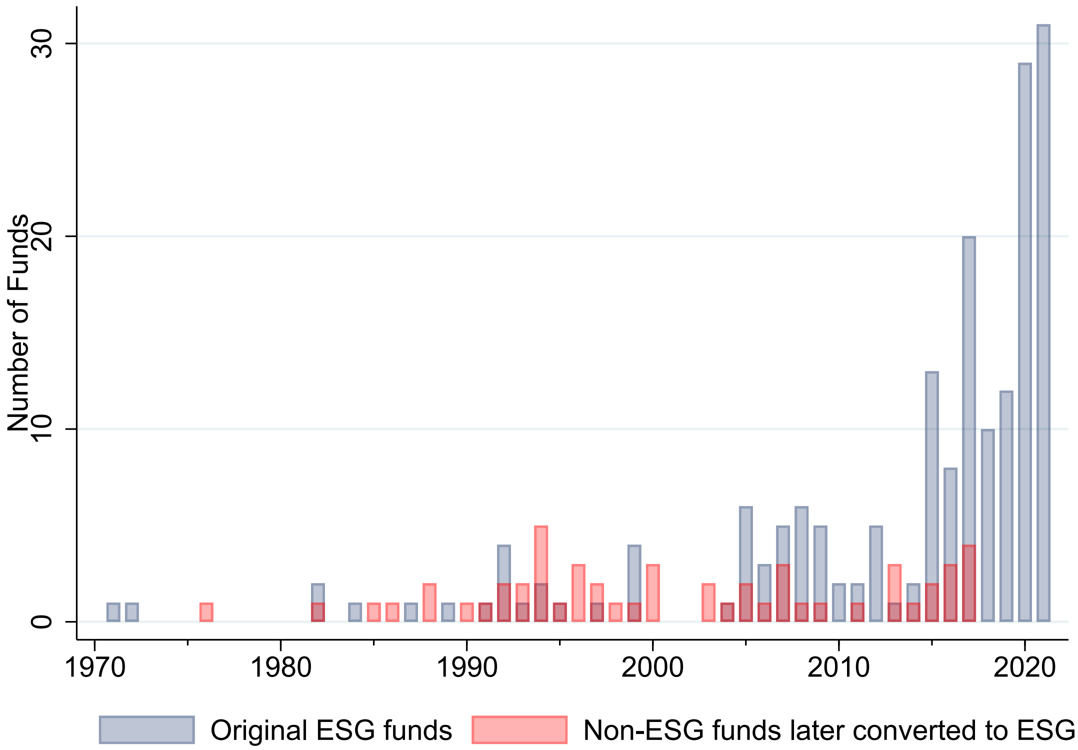
Despite the negative correlation between fees and ESG performance, our study does not find a corresponding negative relationship between fund fees and financial performance. In fact, higher fees are weakly associated with better risk-adjusted financial returns. This finding contrasts with much of the existing literature on conventional funds, which often reports a negative relationship between fees and performance. One potential explanation is that investors in ESG funds may prioritize different attributes compared to conventional fund investors, potentially valuing the ethical alignment of their investments over purely financial returns. Furthermore, the younger age and different demand dynamics of ESG funds may contribute to this observed pattern.

We also explore the relationship between ESG ratings and financial performance, finding that higher ESG ratings are associated with better realized financial outcomes. This indicates that ESG performance and financial performance are not mutually exclusive and that funds with strong ESG ratings can achieve superior risk-adjusted realized returns. This finding aligns with the broader literature suggesting that integrating ESG criteria can enhance financial performance by mitigating risks and capitalizing on sustainable opportunities. However, it is important to note that the estimates we present establish the association between realized returns and ESG performance. The relationship between expected returns and sustainability is beyond the scope of this chapter.

Overall, our findings have implications for both investors and regulators. For investors, the results highlight the need for due diligence when selecting ESG funds, as higher fees do not necessarily translate into better ESG performance. Investors should critically evaluate the ESG credentials of funds and not rely solely on expense ratios as a proxy for quality. For regulators, the evidence of a disconnect between fees and ESG performance underscores the importance of enhanced disclosure requirements and governance standards for ESG funds. Improved transparency and accountability can help ensure that funds truly adhere to their sustainable investing mandates and deliver value to investors.

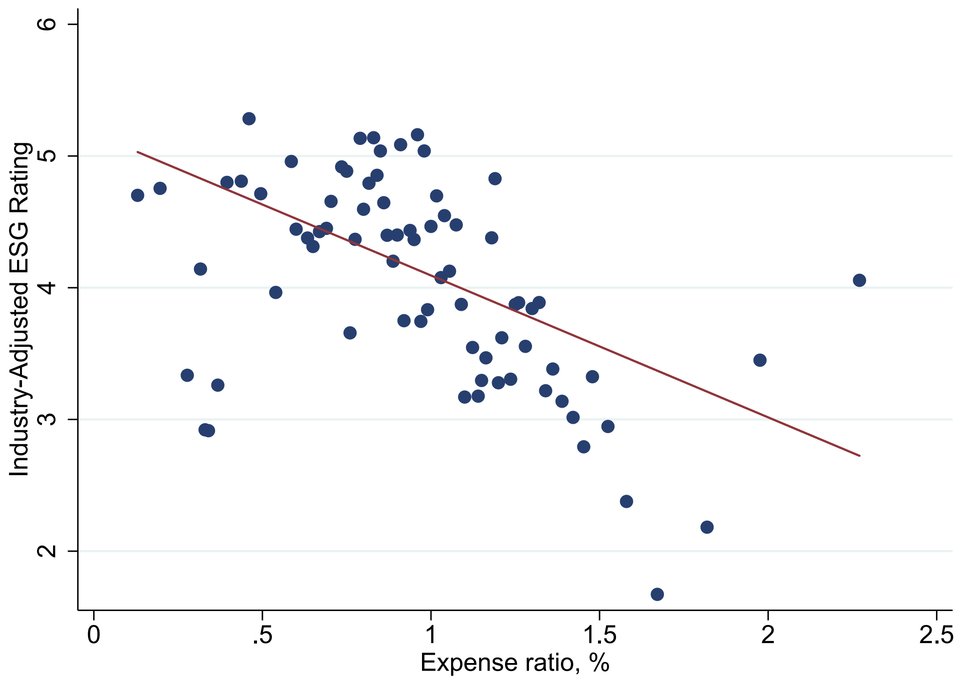
Tables and Figures

Figure 1.1: Inception of ESG equity funds



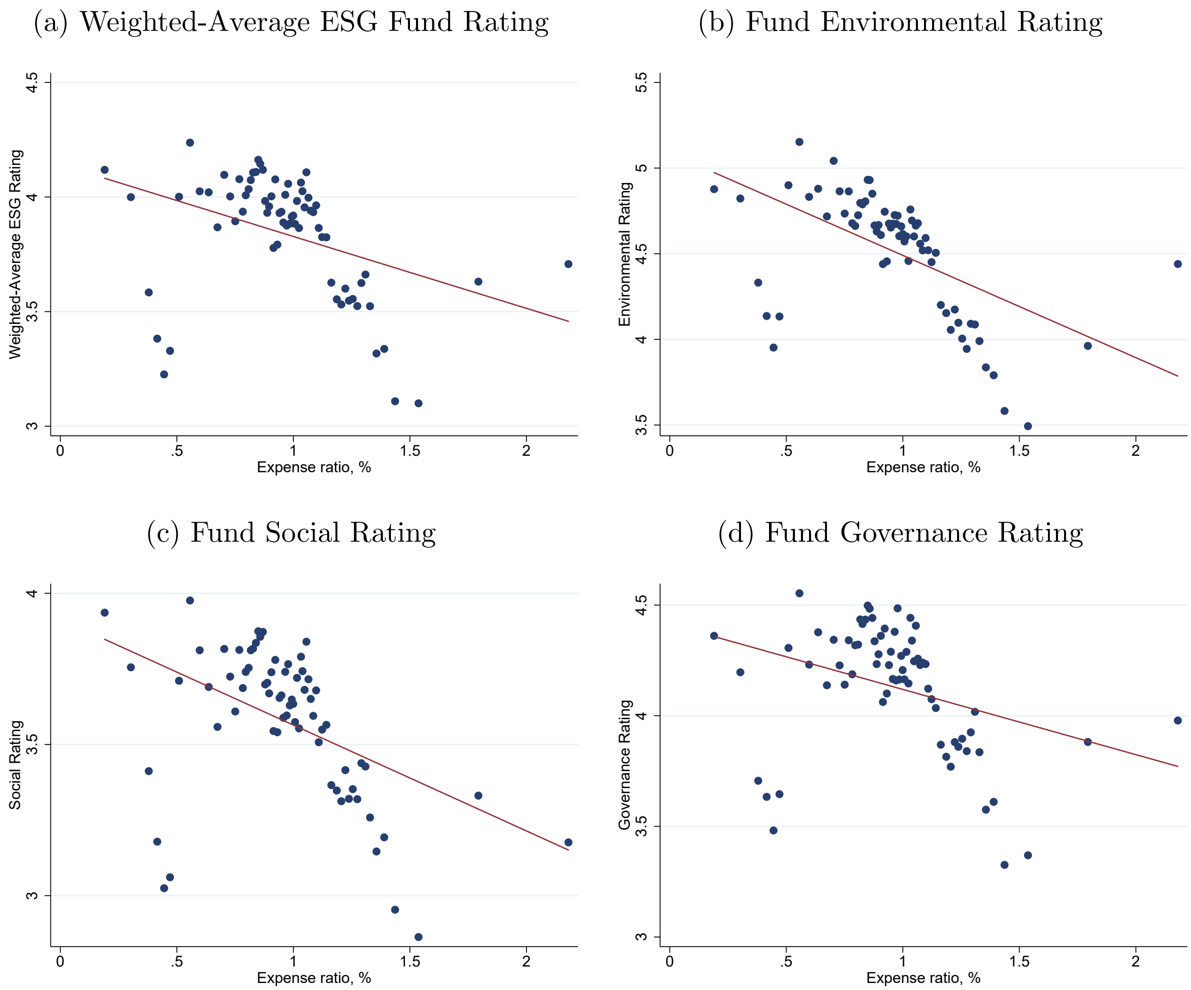
Note: The graph shows the inception dates of ESG equity funds up to the end of 2021. Non-ESG funds that were later converted to ESG are highlighted in red. For these converted funds, the date of ESG conversion is used as the inception date. The total number of funds in this sample is 233.

Figure 1.2: Industy-Adjusted ESG Fund Rating and Fund Expense Ratio



Note: The scatterplot illustrates the relationship between the industry-adjusted ESG fund ratings and the fund expense ratios. A total of 6,351 observations are uniformly distributed across 70 bins. The scatterplot includes month fixed effects but does not control for any fund characteristics.

Figure 1.3: Fund Rating and Fund Expense Ratio



(a) Weighted-Average ESG Fund Rating

(b) Fund Environmental Rating

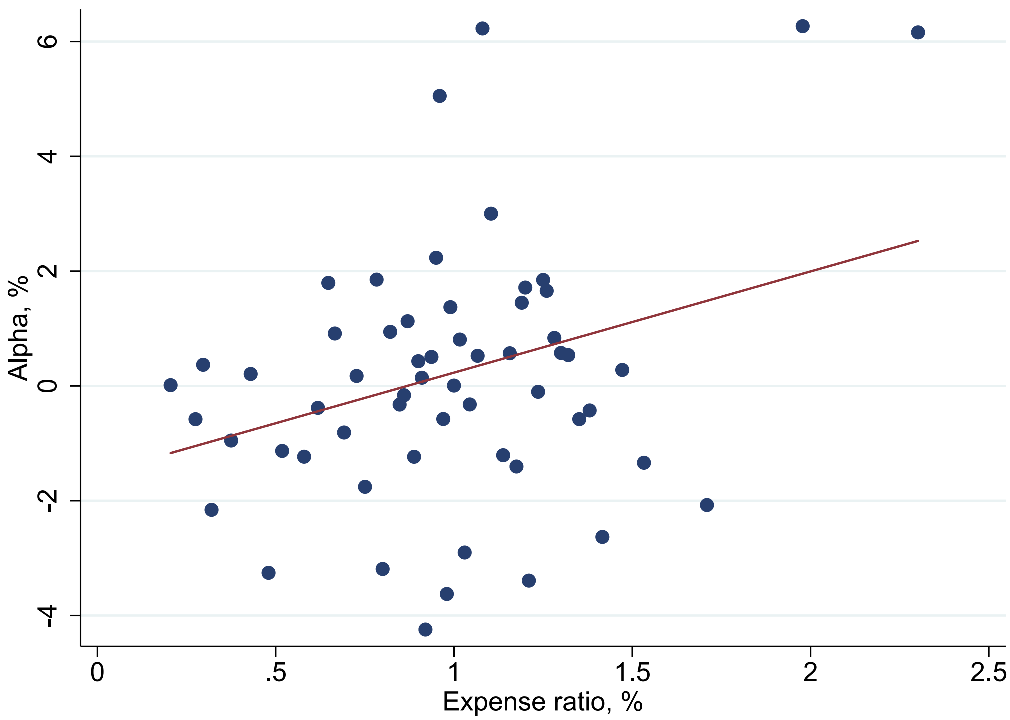
(c) Fund Social Rating

(d) Fund Governance Rating

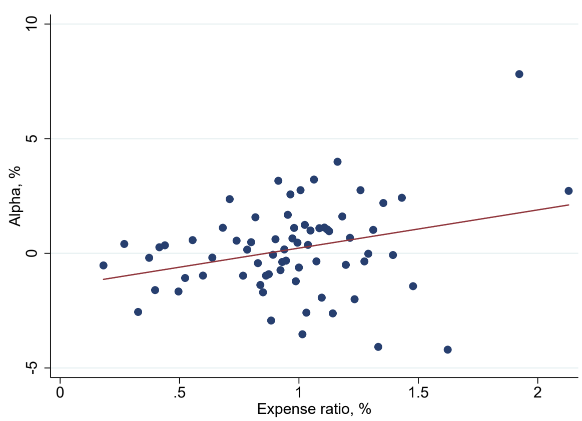
Note: The scatterplots illustrates the relationship between fund ratings and the fund expense ratios. A total of 6,351 observations are uniformly distributed across 70 bins. The scatterplot includes month fixed effects but does not control for any fund characteristics.

Figure 1.4: Fama-French Three-Factor Alpha and Expense Ratio

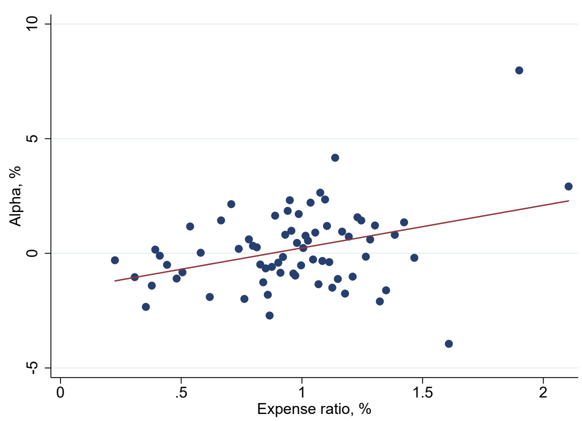
(a) No Controls or Fixed-Effects



(b) With Time Fixed Effects



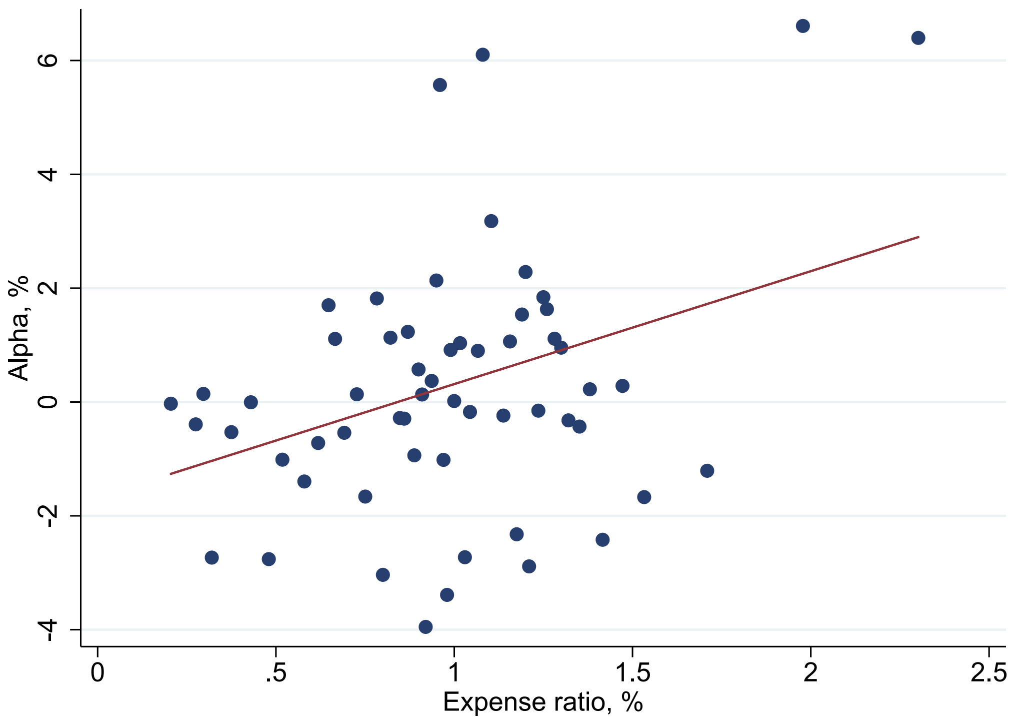
(c) With Controls and Time Fixed Effects



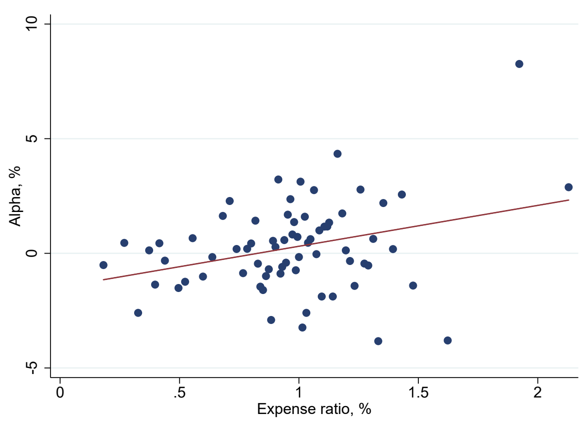
Note: The scatter plots illustrate the relationship between Fama-French three-factor alphas and fund expense ratios. 9,414 observations are uniformly distributed across 70 bins. Scatter plot (a) does not account for any fund characteristics. In contrast, scatter plot (b) incorporates time fixed effects, and scatter plot (c) further controls for fund age and assets under management (AUM).

Figure 1.5: Carhart Four-Factor Alpha and Expense Ratio

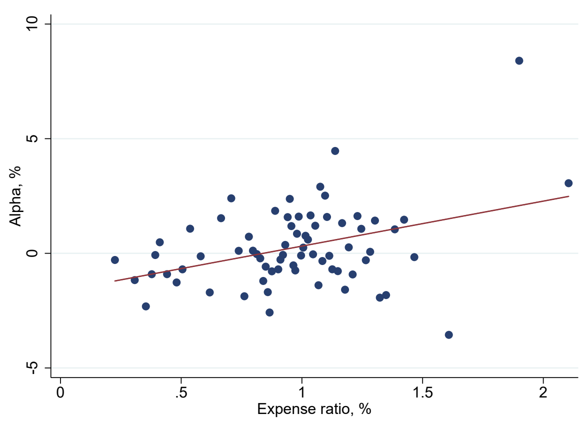
(a) No Controls or Fixed-Effects



(b) With Time Fixed Effects

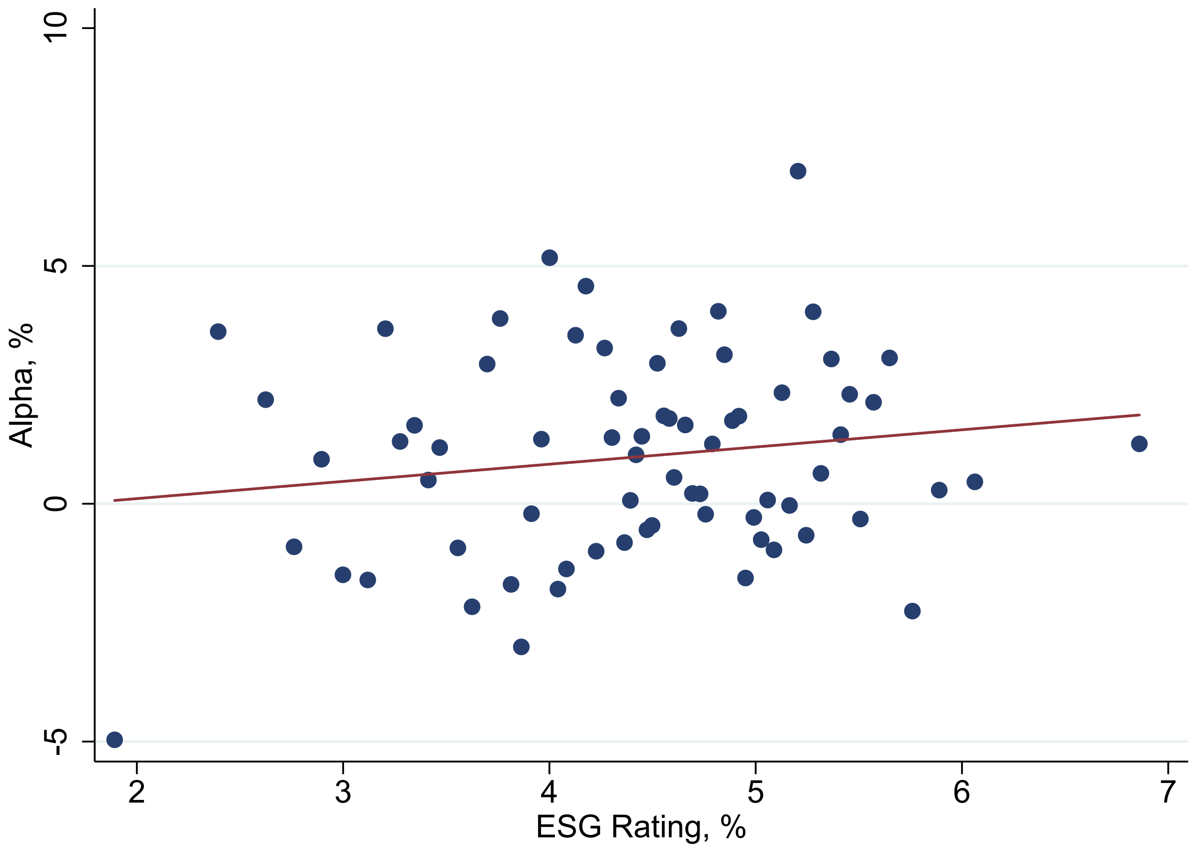


(c) With Controls and Time Fixed Effects



Note: The scatter plots illustrate the relationship between Carhart four-factor alphas and fund expense ratios. 9,414 observations are uniformly distributed across 70 bins. Scatter plot (a) does not account for any fund characteristics. In contrast, scatter plot (b) incorporates time fixed effects, and scatter plot (c) further controls for fund age and assets under management (AUM).

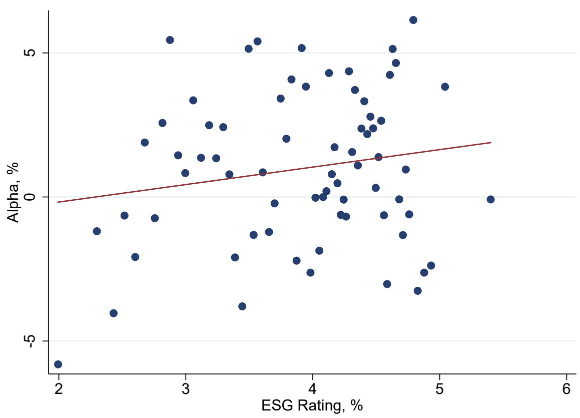
Figure 1.6: Industry-Adjusted ESG Rating and Fama-French Three-Factor Alpha



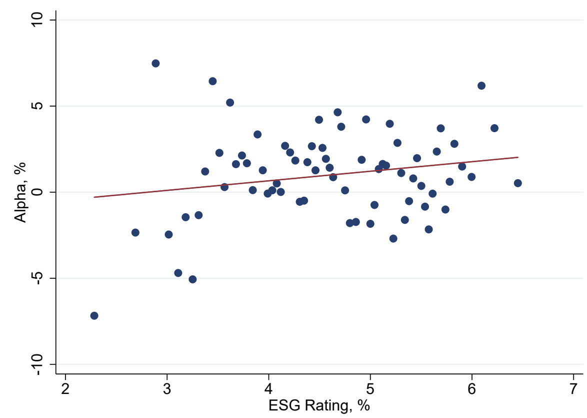
Note: The scatter plot illustrates the relationship between monthly Fama-French three-factor alpha and monthly industry-adjusted ESG rating. The alphas are annualized. Observations are uniformly distributed across 70 bins. The scatter plot includes time fixed effects.

Figure 1.7: ESG Ratings and Fama-French Three-Factor Alpha

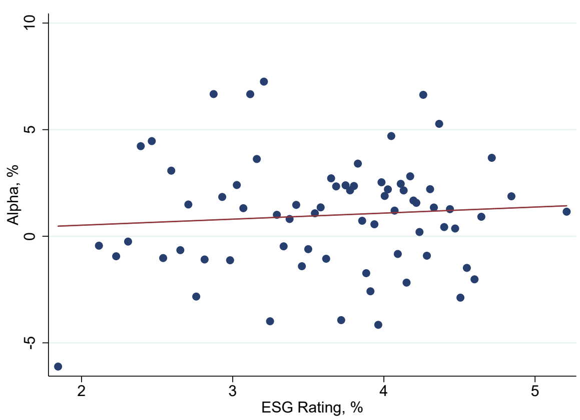
(a) Weighted-Average ESG Rating



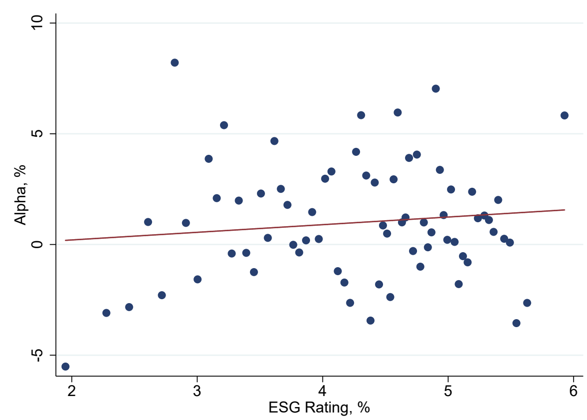
(b) Environmental Rating



(c) Social Rating

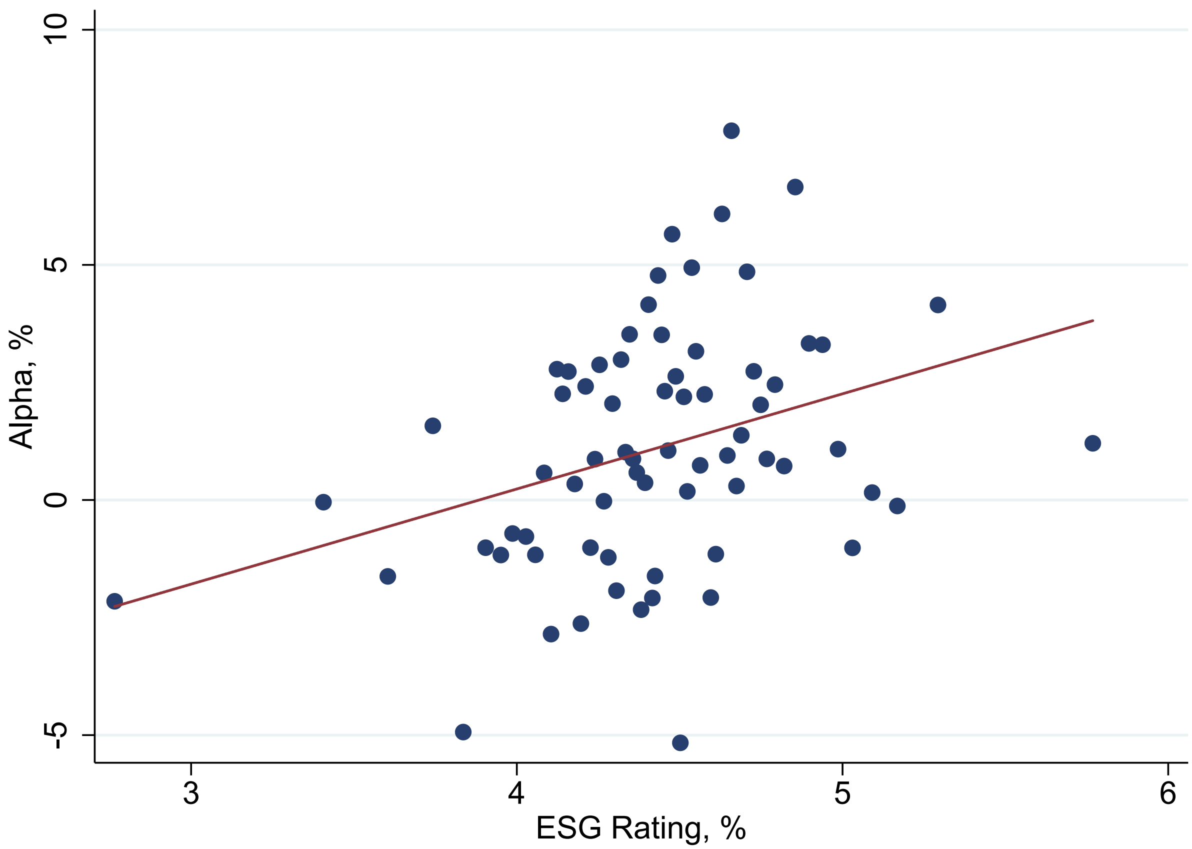


(d) Governance Rating



Note: The scatter plots illustrate the relationship between monthly Fama-French three-factor alphas and different ESG ratings. The alphas are annualized. Observations are uniformly distributed across 70 bins. All scatter plots include time fixed effects.

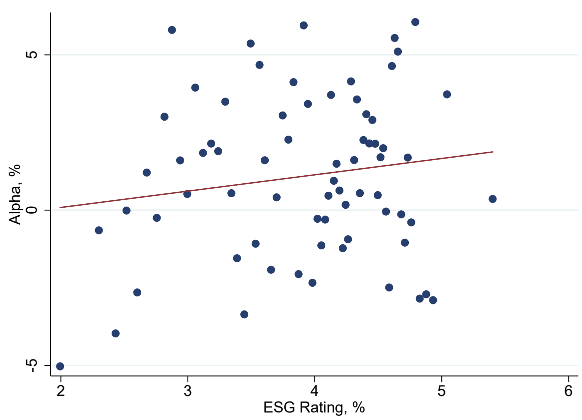
Figure 1.8: Industry-Adjusted ESG Rating and Carhart Four-Factor Alpha



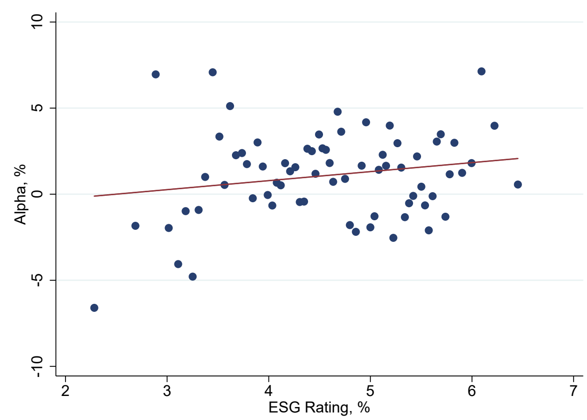
Note: The scatter plot illustrates the relationship between monthly Carhart four-factor alpha and monthly industry-adjusted ESG rating. The alphas are annualized. Observations are uniformly distributed across 70 bins. The scatter plot includes time fixed effects.

Figure 1.9: ESG Ratings and Carhart Four-Factor Alpha

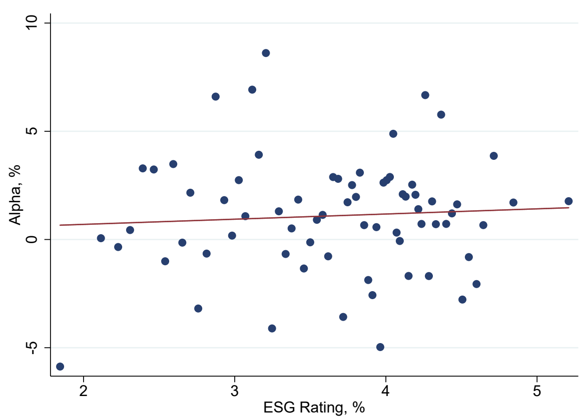
(a) Weighted-Average ESG Rating



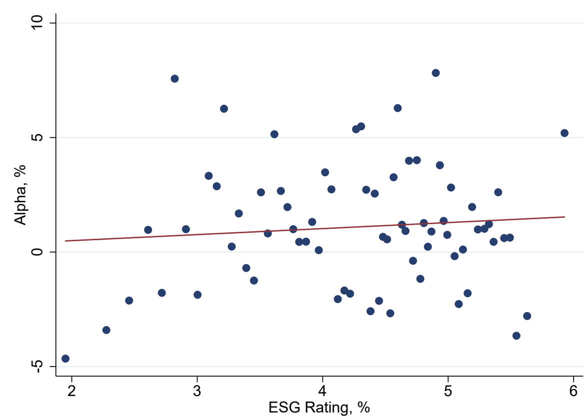
(b) Environmental Rating



(c) Social Rating



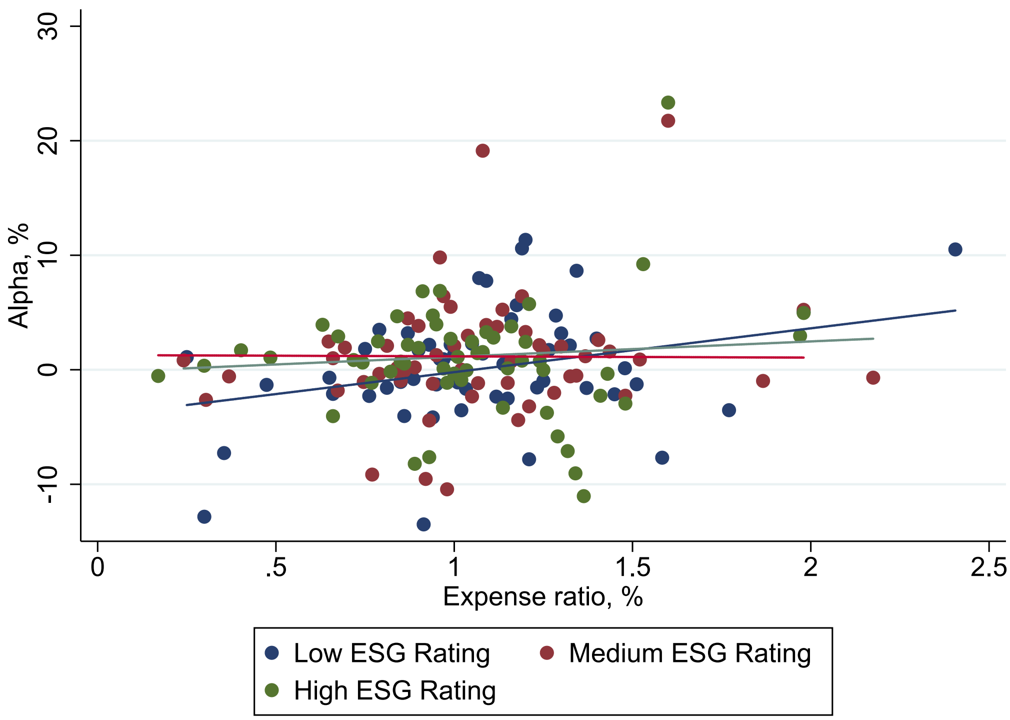
(d) Governance Rating



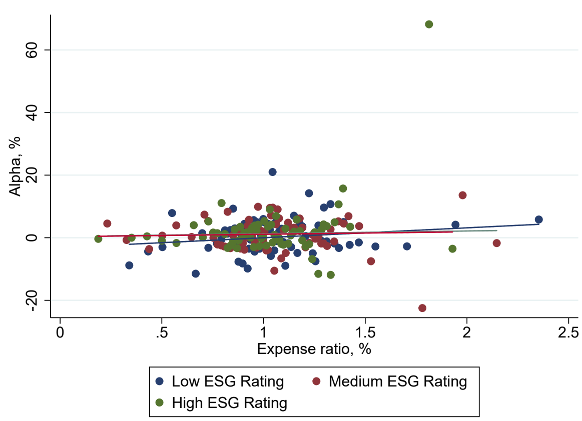
Note: The scatter plots illustrate the relationship between monthly Carhart four-factor alphas and different ESG ratings. The alphas are annualized. Observations are uniformly distributed across 70 bins. All scatter plots include time fixed effects.

Figure 1.10: Fama-French Three-Factor Alpha and Expense Ratio

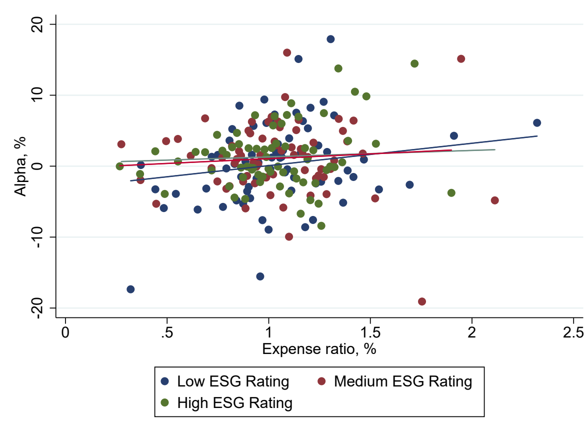
(a) No Controls or Fixed-Effects



(b) With Time Fixed Effects



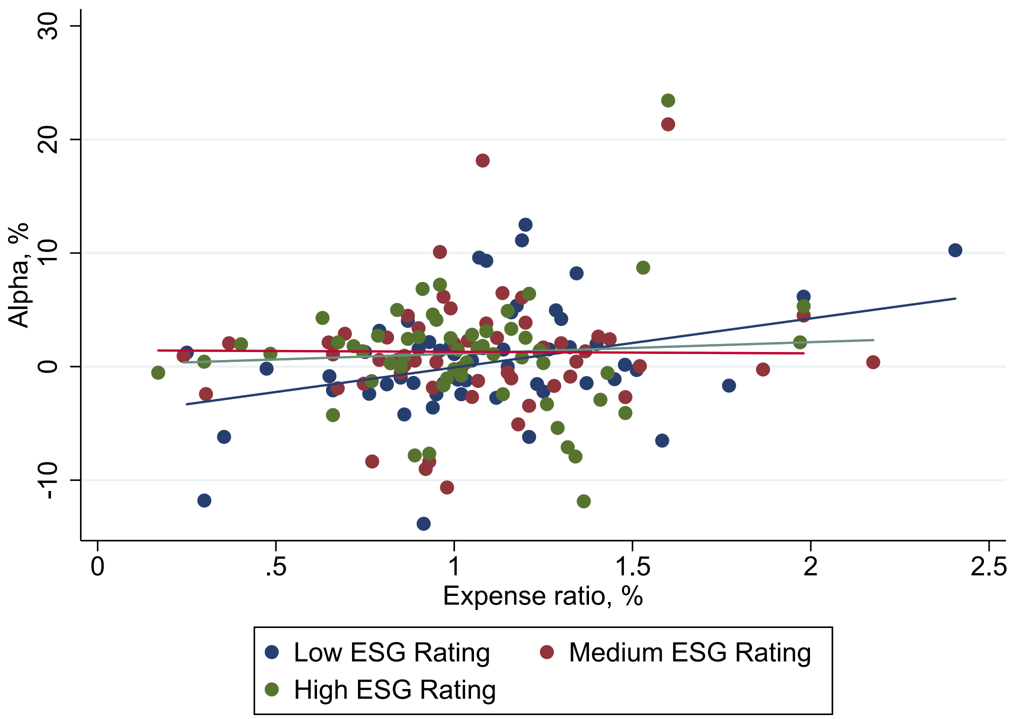
(c) With Controls and Time Fixed Effects



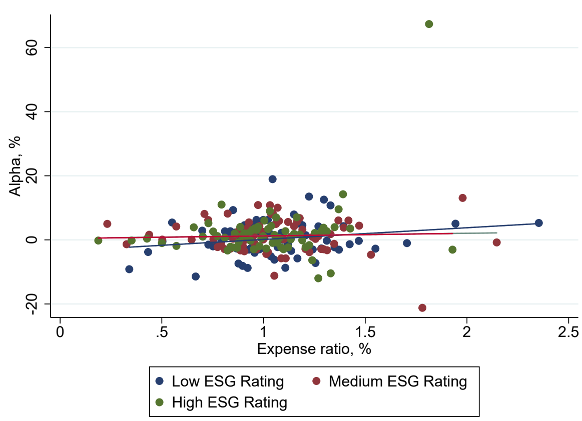
Note: The scatter plots illustrate the relationship between Fama-French three-factor alphas and fund expense ratios. Observations are uniformly distributed across 70 bins. Scatter plot (a) does not account for any fund characteristics. In contrast, scatter plot (b) incorporates time fixed effects, and scatter plot (c) further controls for fund age and assets under management (AUM).

Figure 1.11: Carhart Four-Factor Alpha and Expense Ratio

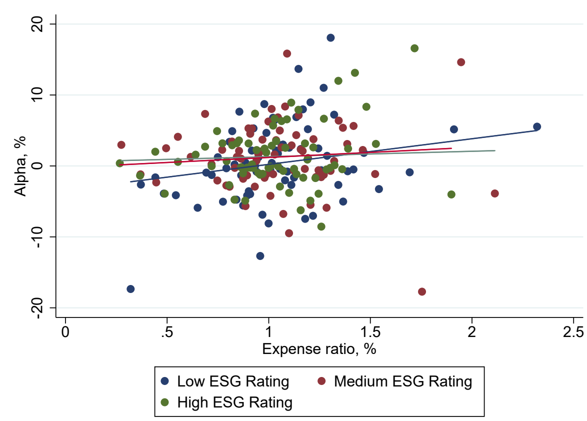
(a) No Controls or Fixed-Effects



(b) With Time Fixed Effects



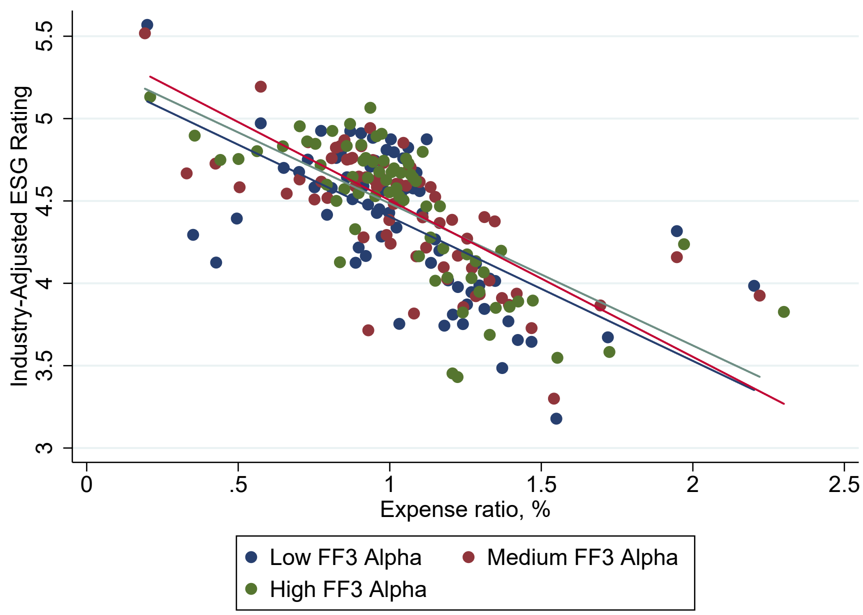
(c) With Controls and Time Fixed Effects



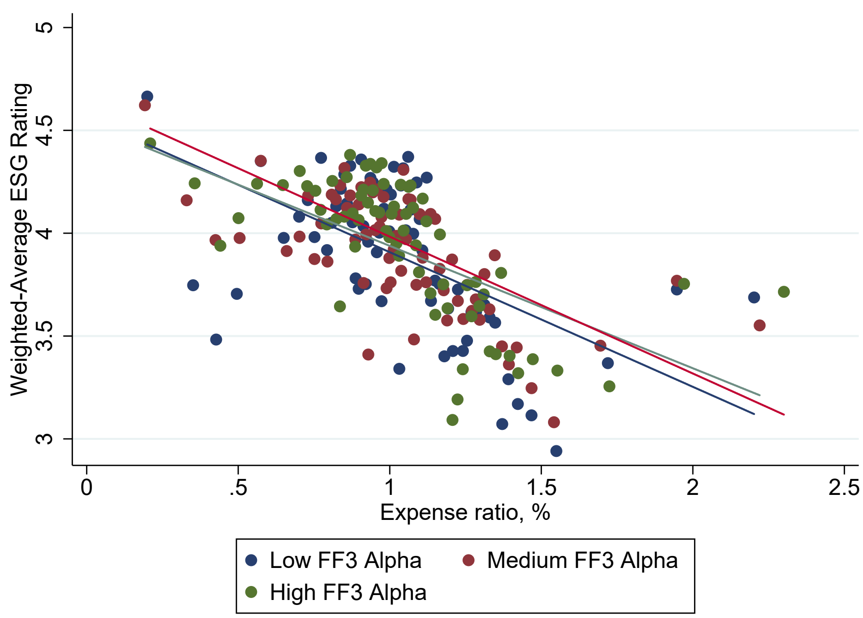
Note: The scatter plots illustrate the relationship between Carhart four-factor alphas and fund expense ratios. Observations are uniformly distributed across 70 bins. Scatter plot (a) does not account for any fund characteristics. In contrast, scatter plot (b) incorporates time fixed effects, and scatter plot (c) further controls for fund age and assets under management (AUM).

Figure 1.12: ESG Ratings vs Expense Ratio by terciles of FF3 Alphas

(a) Industry-Adjusted ESG Rating



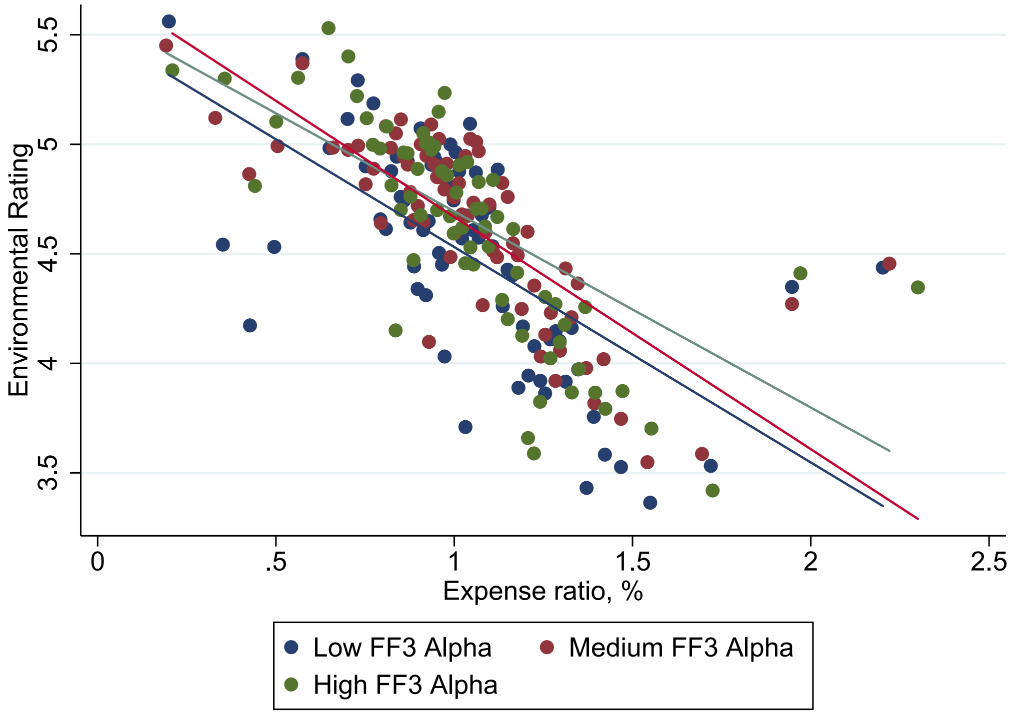
(b) Weighted-Average ESG Rating



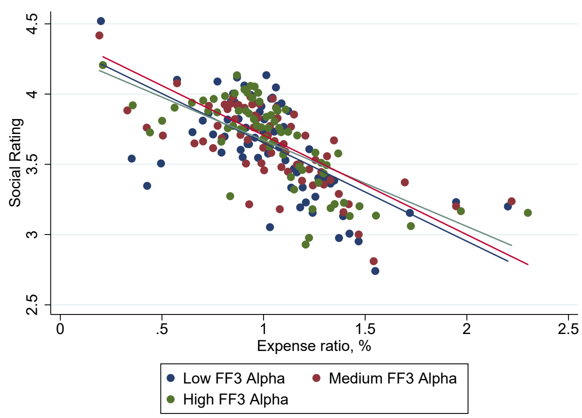
Note: The scatter plots illustrate the relationship between fund expense ratios and (a) industry-adjusted (b) weighted average fund ESG ratings for low, medium and high terciles of Fama-French Three-Factor alphas. Observations are uniformly distributed across 70 bins. All scatter plots include month fixed effects.

Figure 1.13: ESG Scores vs Expense Ratio by terciles of FF3 Alphas

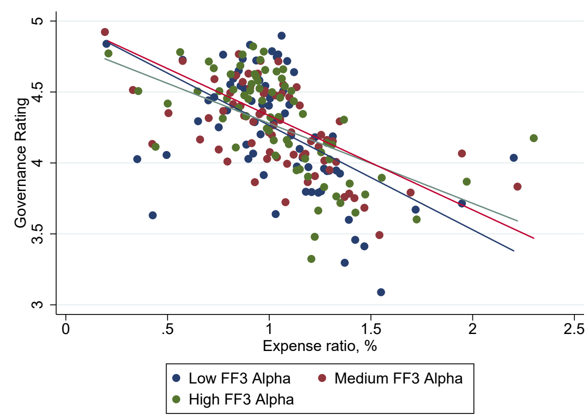
(a) Environmental Rating



(b) Social Rating



(c) Governance Rating



Note: The scatter plots illustrate the relationship between fund expense ratios and (a) Environmental (b) Social (c) Governance Scores for low, medium and high terciles of Fama-French Three-Factor alphas. Observations are uniformly distributed across 70 bins. All scatter plots include month fixed effects.

Table 1.1: Summary Statistics

	Mean	25th	50th	75th	95th	99th	Max	N
Total Net Assets	391.73	10.30	63.30	265.30	1,707.20	5,587.10	15,407.70	15,824
Expense Ratio, percent	0.89	0.68	0.92	1.08	1.45	1.98	2.65	15,125
Management Fee	-0.26	0.05	0.56	0.75	1.08	1.58	1.98	13,658
Turnover Ratio	0.49	0.19	0.34	0.61	1.28	2.25	13.82	13,604
Monthly returns (after fee)	0.65	-2.19	1.04	3.48	8.25	12.63	27.54	15,711
FF-3 Alpha	-0.15	-0.97	-0.08	0.71	2.64	5.04	5.04	9,417
Carhart Alpha	-0.14	-0.98	-0.07	0.73	2.62	4.99	4.99	9,417
ESG rating (ind-adj)	4.36	3.72	4.36	5.08	5.94	6.48	7.44	6,477
ESG rating (w.-av.)	3.84	3.28	3.96	4.48	4.92	5.17	5.80	6,477
Environmental Fund Score	4.51	3.78	4.51	5.30	6.07	6.57	7.11	6,477
Social Fund Score	3.58	3.02	3.63	4.18	4.72	5.09	5.60	6,477
Governance Fund Score	4.12	3.35	4.17	4.86	5.67	6.39	7.81	6,477
Observations	15830							

Note: This table reports summary statistics for the sample, which comprises 233 active equity ESG funds. This includes 58 funds that were initially non-ESG but have since been converted to ESG. For these converted funds, the date of ESG conversion is used as the inception date.

Table 1.2: Fund ESG ratings and Expense Ratio

	(1) ESG rating (ind-adj)	(2) ESG rating (w.-av.)	(3) Environmental Fund Score	(4) Social Fund Score	(5) Governance Fund Score
Expense Ratio, percent	-0.400***	-0.190	-0.460***	-0.225**	-0.141
	(0.135)	(0.117)	(0.168)	(0.105)	(0.159)
Percent in equities	0.050***	0.046***	0.049***	0.045***	0.053***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
Monthly returns (after fee)	0.014**	0.014***	0.021***	0.011**	0.014**
	(0.006)	(0.005)	(0.006)	(0.004)	(0.006)
Log(Total Net Assets)	0.061*	0.047*	0.064**	0.062**	0.050*
	(0.033)	(0.025)	(0.029)	(0.025)	(0.030)
Log(Age in quarters)	0.058	0.054	0.072	0.014	0.061
	(0.053)	(0.039)	(0.046)	(0.039)	(0.044)
Constant	-0.219	-0.456	0.061	-0.543**	-0.865**
	(0.410)	(0.287)	(0.434)	(0.269)	(0.382)
Month FE	Yes	Yes	Yes	Yes	Yes
R-Squared	0.514	0.515	0.464	0.546	0.531
Observations	6281	6281	6281	6281	6281

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Note: The table shows estimated slope coefficients for the OLS regression of funds' ESG performance on expense ratios in the period from January 2011 to March 2022. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. In columns (3)-(5) the dependent variables are Environmental, Social, and Governance Scores, respectively. All regressions include time fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.3: Fund ESG ratings and Expense Ratio: No Controls

	(1) ESG rating (ind-adj)	(2) ESG rating (w.-av.)	(3) Environmental Fund Score	(4) Social Fund Score	(5) Governance Fund Score
Expense Ratio, percent	-0.521***	-0.314**	-0.597***	-0.350**	-0.295
	(0.183)	(0.154)	(0.200)	(0.146)	(0.189)
Constant	4.866***	4.142***	5.088***	3.914***	4.414***
	(0.203)	(0.168)	(0.217)	(0.160)	(0.193)
Month FE	Yes	Yes	Yes	Yes	Yes
R-Squared	0.173	0.068	0.135	0.096	0.176
Observations	6351	6351	6351	6351	6351

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Note: The table shows estimated slope coefficients for the OLS regression of funds' ESG performance on expense ratios in the period from January 2011 to March 2022. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. In columns (3)-(5) the dependent variables are Environmental, Social, and Governance Scores, respectively. All regressions include time fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.4: Fama-French Three-Factor Alpha and Expense Ratio: No Controls

	(1) FF-3 Alpha	(2) FF-3 Alpha	(3) FF-3 Alpha
Monthly Expense Ratio	-0.174	-0.691	0.920
	(0.743)	(0.737)	(0.560)
Constant	-0.134**	-0.092	-0.222***
	(0.060)	(0.060)	(0.048)
Month FE	No	Yes	Yes
Fund Type FE	No	No	Yes
R-Squared	0.000	0.216	0.227
Observations	9417	9417	9417

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly alphas on monthly expense ratios in the period from January 2011 to September 2023. The dependent variable in all columns is the fund's Fama-French three-factor monthly alpha. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.5: Fama-French Three-Factor Alpha and Expense Ratio

	(1) FF-3 Alpha	(2) FF-3 Alpha	(3) FF-3 Alpha
Monthly Expense Ratio	0.559	0.119	1.842***
	(0.858)	(0.865)	(0.572)
Turnover Ratio	-0.149**	-0.141**	-0.153***
	(0.066)	(0.065)	(0.047)
Log(Total Net Assets)	0.014	0.009	0.012
	(0.013)	(0.013)	(0.009)
Log(Age in quarters)	0.030	0.036*	-0.010
	(0.021)	(0.021)	(0.016)
Percent in equities	0.000	0.001	0.002
	(0.001)	(0.001)	(0.002)
Constant	-0.320**	-0.304**	-0.456***
	(0.140)	(0.137)	(0.161)
Month FE	No	Yes	Yes
Fund Type FE	No	No	Yes
R-Squared	0.002	0.197	0.213
Observations	8302	8302	8302

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly alphas on monthly expense ratios in the period from January 2011 to September 2023. The dependent variable in all columns is the fund's Fama-French three-factor monthly alpha. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.
t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.6: Carhart Four-Factor Alpha and Expense Ratio: No Controls

	(1) Carhart Alpha	(2) Carhart Alpha	(3) Carhart Alpha
Monthly Expense Ratio	-0.020	-0.623	0.948*
	(0.682)	(0.678)	(0.550)
Constant	-0.140**	-0.091	-0.218***
	(0.056)	(0.056)	(0.047)
Month FE	No	Yes	Yes
Fund Type FE	No	No	Yes
R-Squared	0.000	0.214	0.224
Observations	9417	9417	9417

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly alphas on monthly expense ratios in the period from January 2011 to September 2023. The dependent variable in all columns is the fund's Carhart four-factor monthly alpha. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.7: Carhart Four-Factor Alpha and Expense Ratio

	(1) Carhart Alpha	(2) Carhart Alpha	(3) Carhart Alpha
Monthly Expense Ratio	0.465	0.031	1.733***
	(0.791)	(0.792)	(0.570)
Turnover Ratio	-0.153**	-0.150**	-0.163***
	(0.063)	(0.062)	(0.046)
Log(Total Net Assets)	0.012	0.005	0.008
	(0.013)	(0.013)	(0.010)
Log(Age in quarters)	0.032	0.037*	-0.006
	(0.021)	(0.020)	(0.016)
Percent in equities	0.001	0.001	0.002
	(0.001)	(0.001)	(0.002)
Constant	-0.298**	-0.275**	-0.415***
	(0.133)	(0.129)	(0.153)
Month FE	No	Yes	Yes
Fund Type FE	No	No	Yes
R-Squared	0.002	0.196	0.210
Observations	8302	8302	8302

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly alphas on monthly expense ratios in the period from January 2011 to September 2023. The dependent variable in all columns is the fund's Carhart four-factor monthly alpha. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.8: Industry-Adjusted ESG Ratings and Fund Financial Performance

	(1) FF-3 Alpha	(2) FF-3 Alpha	(3) FF-3 Alpha	(4) Carhart Alpha	(5) Carhart Alpha	(6) Carhart Alpha
ESG rating (ind-adj)	0.124***	0.140***	0.083*	0.118***	0.137***	0.083*
	(0.040)	(0.048)	(0.044)	(0.039)	(0.047)	(0.042)
Monthly Expense Ratio	1.212	0.581	1.540	1.198	0.624	1.597
	(1.189)	(1.096)	(1.114)	(1.100)	(0.992)	(1.048)
Log(Total Net Assets)	0.005	0.002	0.008	0.006	0.000	0.006
	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)	(0.015)
Log(Age in quarters)	0.018	0.023	0.005	0.018	0.024	0.006
	(0.020)	(0.020)	(0.022)	(0.020)	(0.020)	(0.022)
Percent in equities	-0.005*	-0.006**	0.002	-0.005*	-0.006**	0.002
	(0.003)	(0.003)	(0.005)	(0.002)	(0.003)	(0.005)
Constant	-0.384	-0.277	-0.836*	-0.376	-0.276	-0.801*
	(0.258)	(0.241)	(0.479)	(0.244)	(0.231)	(0.439)
Month FE	No	Yes	Yes	No	Yes	Yes
Fund Type FE	No	No	Yes	No	No	Yes
R-Squared	0.003	0.191	0.197	0.003	0.193	0.198
Observations	4261	4260	4259	4261	4260	4259

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly financial performance on monthly industry-adjusted funds' ESG ratings in the period from January 2011 to March 2022. The dependent variable in columns (1)-(3) is the fund's monthly Fama-French three-factor alpha, while in columns (4)-(6), the dependent variable in the fund's monthly Carhart four-factor alpha. All regressions include time and fund-type fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.9: Weighted-Average ESG Ratings and Fund Financial Performance

	(1) FF-3 Alpha	(2) FF-3 Alpha	(3) FF-3 Alpha	(4) Carhart Alpha	(5) Carhart Alpha	(6) Carhart Alpha
ESG rating (w.-av.)	0.192***	0.201***	0.105	0.176***	0.190***	0.104*
	(0.056)	(0.060)	(0.065)	(0.056)	(0.059)	(0.062)
Monthly Expense Ratio	0.973	0.440	1.379	0.930	0.458	1.432
	(1.088)	(0.987)	(1.097)	(1.009)	(0.899)	(1.039)
Log(Total Net Assets)	0.004	0.001	0.008	0.005	0.000	0.007
	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)	(0.015)
Log(Age in quarters)	0.010	0.021	0.005	0.011	0.022	0.006
	(0.019)	(0.020)	(0.022)	(0.019)	(0.021)	(0.022)
Percent in equities	-0.007**	-0.008***	0.002	-0.007**	-0.008***	0.001
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)
Constant	-0.314	-0.237	-0.810*	-0.309	-0.237	-0.773*
	(0.254)	(0.240)	(0.484)	(0.240)	(0.230)	(0.444)
Month FE	No	Yes	Yes	No	Yes	Yes
Fund Type FE	No	No	Yes	No	No	Yes
R-Squared	0.004	0.192	0.197	0.004	0.193	0.198
Observations	4261	4260	4259	4261	4260	4259

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly financial performance on monthly industry-adjusted funds' ESG ratings in the period from January 2011 to March 2022. The dependent variable in columns (1)-(3) is the fund's monthly Fama-French three-factor alpha, while in columns (4)-(6), the dependent variable in the fund's monthly Carhart four-factor alpha. All regressions include time and fund-type fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.10: Environmental Scores and Fund Financial Performance

	(1) FF-3 Alpha	(2) FF-3 Alpha	(3) FF-3 Alpha	(4) Carhart Alpha	(5) Carhart Alpha	(6) Carhart Alpha
Environmental Fund Score	0.155***	0.169***	0.138***	0.151***	0.169***	0.145***
	(0.034)	(0.036)	(0.044)	(0.033)	(0.035)	(0.043)
Monthly Expense Ratio	1.530	0.931	1.576	1.530	0.995	1.649*
	(1.078)	(0.984)	(1.034)	(0.991)	(0.887)	(0.966)
Log(Total Net Assets)	0.003	0.000	0.008	0.004	-0.001	0.007
	(0.016)	(0.015)	(0.016)	(0.015)	(0.015)	(0.016)
Log(Age in quarters)	0.012	0.019	0.001	0.012	0.019	0.001
	(0.020)	(0.021)	(0.022)	(0.020)	(0.021)	(0.023)
Percent in equities	-0.006**	-0.007***	-0.001	-0.006**	-0.007***	-0.002
	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)	(0.006)
Constant	-0.441*	-0.351	-0.771	-0.434*	-0.351	-0.736
	(0.246)	(0.235)	(0.519)	(0.233)	(0.226)	(0.478)
Month FE	No	Yes	Yes	No	Yes	Yes
Fund Type FE	No	No	Yes	No	No	Yes
R-Squared	0.005	0.193	0.198	0.005	0.195	0.199
Observations	4261	4260	4259	4261	4260	4259

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly financial performance on monthly Environmental Scores in the period from January 2011 to March 2022. The dependent variable in columns (1)-(3) is the fund's monthly Fama-French three-factor alpha, while in columns (4)-(6), the dependent variable in the fund's monthly Carhart four-factor alpha. All regressions include time and fund-type fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.11: Social Scores and Fund Financial Performance

	(1) FF-3 Alpha	(2) FF-3 Alpha	(3) FF-3 Alpha	(4) Carhart Alpha	(5) Carhart Alpha	(6) Carhart Alpha
Social Fund Score	0.143**	0.179***	0.107	0.138**	0.174***	0.108
	(0.058)	(0.062)	(0.068)	(0.057)	(0.062)	(0.066)
Monthly Expense Ratio	0.850	0.431	1.465	0.859	0.472	1.524
	(1.168)	(1.052)	(1.138)	(1.080)	(0.957)	(1.083)
Log(Total Net Assets)	0.003	-0.001	0.007	0.004	-0.002	0.006
	(0.017)	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)
Log(Age in quarters)	0.020	0.031	0.008	0.020	0.032	0.009
	(0.020)	(0.020)	(0.022)	(0.020)	(0.020)	(0.022)
Percent in equities	-0.005	-0.007**	0.002	-0.005*	-0.007**	0.001
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)
Constant	-0.309	-0.225	-0.805*	-0.304	-0.225	-0.770*
	(0.256)	(0.238)	(0.475)	(0.239)	(0.225)	(0.435)
Month FE	No	Yes	Yes	No	Yes	Yes
Fund Type FE	No	No	Yes	No	No	Yes
R-Squared	0.002	0.191	0.197	0.002	0.192	0.198
Observations	4261	4260	4259	4261	4260	4259

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly financial performance on monthly funds Social SCores in the period from January 2011 to March 2022. The dependent variable in columns (1)-(3) is the fund's monthly Fama-French three-factor alpha, while in columns (4)-(6), the dependent variable in the fund's monthly Carhart four-factor alpha. All regressions include time and fund-type fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table 1.12: Governance Scores and Fund Financial Performance

	(1) FF-3 Alpha	(2) FF-3 Alpha	(3) FF-3 Alpha	(4) Carhart Alpha	(5) Carhart Alpha	(6) Carhart Alpha
Governance Fund Score	0.096**	0.132***	0.050	0.089**	0.122**	0.046
	(0.038)	(0.049)	(0.053)	(0.038)	(0.050)	(0.050)
Monthly Expense Ratio	0.140	0.082	1.262	0.168	0.107	1.316
	(1.127)	(0.958)	(1.089)	(1.019)	(0.875)	(1.030)
Log(Total Net Assets)	0.007	0.004	0.010	0.008	0.002	0.009
	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Log(Age in quarters)	0.007	0.025	0.005	0.007	0.026	0.006
	(0.021)	(0.022)	(0.023)	(0.021)	(0.022)	(0.023)
Percent in equities	-0.003	-0.006**	0.003	-0.003	-0.005**	0.003
	(0.002)	(0.003)	(0.006)	(0.002)	(0.003)	(0.006)
Constant	-0.244	-0.201	-0.773	-0.243	-0.205	-0.737
	(0.257)	(0.242)	(0.487)	(0.239)	(0.231)	(0.448)
Month FE	No	Yes	Yes	No	Yes	Yes
Fund Type FE	No	No	Yes	No	No	Yes
R-Squared	0.002	0.191	0.197	0.002	0.192	0.197
Observations	4261	4260	4259	4261	4260	4259

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows estimated slope coefficients for the OLS regression of funds' monthly financial performance on monthly Governance Scores in the period from January 2011 to March 2022. The dependent variable in columns (1)-(3) is the fund's monthly Fama-French three-factor alpha, while in columns (4)-(6), the dependent variable in the fund's monthly Carhart four-factor alpha. All regressions include time and fund-type fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

CHAPTER 2

From Conventional to Conscious: Exploring the Transformation of Non-ESG Funds into ESG Funds

2.1 Introduction

Sustainable investing has gained significant traction in recent years. Over the past two decades, both the number and variety of funds offering sustainable investment options have experienced a remarkable surge. A 2020 report by the US SIF Foundation reveals that the amount of capital committed to sustainable funds in the United States reached $17.1 trillion in 2020, representing a 42% increase from 2018. This growth highlights the increasing importance of non-pecuniary characteristics in investment decisions for a substantial number of investors (Bauer et al., 2021).

Despite tremendous popularity of ESG funds, the evidence on whether such funds "walk the talk" remains scarce and appears to depend on the sustainability metrics, time horizon or asset class. One contributing factor is the lack of a clear definition of what constitutes sustainable or ESG investing, which often allows fund managers and fund families to commit to ESG strategies without incurring significant costs. This is further complicated by the proliferation of ESG ratings, which frequently exhibit low correlations with one another (Berg et al., 2022b, Berg et al., 2022a).

The surge in interest in ESG investing has resulted in diverse interpretations of ESG criteria and varied implementation strategies among fund managers. For example, some funds adopt an exclusionary approach, avoiding investments in sectors with high environmental impacts, such as oil and gas. In contrast, other funds may include investments in oil and gas companies, provided these firms demonstrate a commitment to decarbonization initiatives. However, the inherent ambiguity of ESG metrics may lead certain funds to engage in practices commonly referred to as "greenwashing" (Albuquerque et al., 2024). This term denotes the deliberate exaggeration of the degree to which products or services incorporate environmental and sustainability considerations (SEC 2021).

In this chapter we examine the decision of non-ESG (or "conventional") equity mutual funds to switch and become ESG funds. Between 2013 and 2021, 58 active non-ESG mutual funds and exchange-traded funds (ETFs) (hereafter "funds") declared themselves as ESG funds. The repurposing process typically involved alterations in fund prospectuses and/or changes in fund names. The research question we are interested in is twofold. First, we would like to find out the effect of fund conversion on the degree to which the funds incorporate ESG criteria in ther investment decision. Utilizing company-specific ESG scores from MSCI and fund holdings data from CRSP, we construct ESG scores at the fund level and investigate the effect of fund conversion on ESG fund ratings. Second, we examine the effects of fund repurposing on various fund characteristics, including fund flows, returns, risk-adjusted alphas, and expense ratios.

Addressing these questions is not always straightforward, as fund repurposing is not an exogenous event. The decision is usually made by fund managers or fund families, rather than being imposed by regulators or other third parties. Moreover, the decision to convert a fund to ESG is likely to be correlated with fund characteristics such as age, returns, or fund flows. Therefore, conventional estimation strategies like difference-in-differences (DiD) cannot be directly applied to disentangle the treatment effect of fund repurposing on the outcomes. Instead, we employ two different approaches using dynamic DiD models in order to estimate causal effect of fund conversion. First, we augment the sample of 58 repurposed funds with conventional non-ESG funds that were not converted by the end of the sample period (i.e., these funds serve as "never treated" control units). We use K-nearest neighbor matching based on fund characteristics such as age, assets under management, and returns.

In the second approach, as a robustness check, we apply the event study approach only to the initial sample of 58 repurposed funds. In this context, fund repurposings were spread over time rather than happening at a single point. We compare conversion funds that have already been converted to those that have not been converted yet, employing an event study methodology with staggered treatment timing (e.g., Athey and Imbens, 2022; Baker et al., 2022). In this case, there are no "never treated" funds, and as time passes, the control group shrinks while the treatment group increases.

First, we explore the effect of conversion on ESG ratings. If ratings do not change following fund conversion, it might indicate that the fund's decision to change labels was primarily a marketing strategy rather than a genuine shift in investment approach. However, our findings indicate that, on average, converted funds experienced a significant increase in their ESG ratings. This improvement was particularly evident in the industry-adjusted ESG ratings (0.4 standard deviations) and, to a lesser extent, in the weighted average ESG ratings (0.2 standard deviations). Notably, the enhancement in ESG scores was most pronounced in the social component of the ESG ratings. This suggests that repurposed funds actively rebalanced their portfolios towards firms with higher ESG performance, particularly in social metrics, alleviating concerns about greenwashing to some extent among investors and regulators.

Next, we look at fund flows to determine if repurposing influenced investor behavior. Before conversion, repurposed funds experienced declining flows, suggesting that the decision to convert may have been partially driven by the need to attract more capital. Post-conversion, there was an observable increase in fund flows, but this uptick was not immediate. Annual flows began to rise approximately five quarters after conversion, while quarterly flows showed improvement after two quarters. This delayed response may be due to informational frictions or investors taking time to evaluate the impact of the conversion on ESG performance. We also consider an alternative measure of fund flows: the number of quarters with negative fund flows over the last two years.[1] The event study results show that after the conversion, the converted funds are more likely to have negative fund flows in the first two quarters. However, this relationship changes after that period, with repurposed funds becoming more likely to have positive fund flows compared to the control group.

We also examined the impact on financial performance, including annual returns, Fama-French three-factor alpha, expense ratios, and profits. The results indicated no consistent patterns of change in these financial metrics post-conversion. Most estimates were not statistically significant, suggesting that ESG conversion did not lead to a substantial change in these areas. This implies that while ESG repurposing may enhance ESG ratings and attract fund flows, it does not necessarily impact other performance metrics.

Finally, we explore the heterogeneity of the effects of ESG conversion. We conduct subsample analyses based on fund size (AUM), expense ratios, and risk-adjusted performance defined by Fama-French three-factor alpha. We find that the observed increase in fund flows was primarily driven by smaller funds (low AUM funds) with no significant change for larger funds. At the same time, the increase in ESG rating post-conversion is similar for low- and high-AUM funds, suggesting that fund flows do not necessarily reflect changes in ESG or financial performance. The improvement in ESG ratings was mainly explained by low-fee funds rebalancing their portfolios towards more ESG-friendly firms. Lastly, the results for high- and low-alpha funds are similar. Overall, the subsample analysis highlighted that different types of funds experienced varying impacts from ESG conversion, reflecting diverse strategic responses to the ESG repurposing.

[1] The rationale behind this measure is twofold: first, to use a more aggregated measure of fund flows, and second, to incorporate the possibility that a binary outcome for fund flows (i.e., positive or negative) might be factored into the decision to repurpose a fund. A recent article in Barron's points out that approximately two-thirds of repurposed funds from 2019 to 2022 experienced outflows in the year leading up to their conversion.

The findings refute the concern that ESG funds simply change the label without introducing any changes. We find that ESG rating substantially increases after ESG repurposing, suggesting a causal effect of conversion. In addition, the results suggest that one of the reasons of ESG conversion might be a decline in fund flows and a correct expectation that fund flows would increase as a result of ESG conversion, attracting ESG conscious investors. We also show that conversion funds are able to increase their ESG rating without hurting financial performance: annual returns and Fama-French three-factor alpha do not significantly change either before or after the conversion. Moreover, conversion funds do not increase their expense ratio. These findings alleviate concerns about the lack of transparency or misleading labeling. On average, ESG funds seem to attract investors by increasing their ESG rating in line with the announced repurposing without increasing fees or exhibiting a decline in financial performance. However, heterogeneity analysis suggests that at least some conversion funds might mislead consumers into investing. For instance, high-fees conversion funds do not substantially increase their ESG ratings, while still experiencing an increase in fund flows (although most dynamic effects are not statistically significant). Funds with high fees might be able to attract unsophisticated consumers by aggressive marketing or by other means. These consumers could benefit from simple and accessible information on ESG funds. In an effort to facilitate informed decision-making among investors, the Securities and Exchange Commission (SEC) introduced a set of proposed rules in 2022 aimed at enhancing transparency concerning the ESG metrics of funds. These proposed regulations encompass: (1) the establishment of new disclosure and reporting obligations pertaining to ESG investments, and (2) the refinement of fund naming conventions related to ESG to prevent potential investor confusion. Our findings point to potential benefits of these regulation for certain types of consumers.

The rest of the chapter is organized as follows. Section 2.2 describes the background on ESG funds and discusses the related literature. Section 2.3 describes the data and provides summary statistics. The empirical framework is explained in Section 2.4. Section 2.5 presents the results, and Section 2.6 concludes.

2.2 Background

Regulatory bodies and investors are increasingly scrutinizing the methodologies and outcomes associated with numerous ESG funds. In the absence of comprehensive reporting standards, it becomes challenging for investors to ascertain the manner in which a fund incorporates ESG considerations into its investment decisions and the tangible effects of those investments on the purported ESG objectives. This ambiguity raises concerns about the transparency and effectiveness of ESG funds in contributing to sustainable development goals.

The regulation agencies, especially the Securities and Exchange Commission (SEC), have been actively working to enhance transparency and standardization in ESG disclosures and ESG funds. The SEC's Division of Examinations published a Risk Alert[2] on in April 2021 highlighting observations from recent examinations of investment advisers, registered investment companies, and private funds offering ESG products and services. The alert aims to inform market participants about focus areas related to ESG investing and assist firms in developing and enhancing their compliance practices. Key findings from the examinations include:

- *Inconsistent ESG Disclosures:* Firms were found to have inconsistencies between their ESG-related disclosures and actual practices. Some disclosures lacked specificity, leading to potential confusion among investors.

- *Lack of Policies and Procedures:* Some firms did not have adequate written policies and procedures to ensure that ESG-related disclosures and marketing materials were consistent with the firm's practices.

[2]https://www.sec.gov/files/esg-risk-alert.pdf

- *Proxy Voting Inconsistencies*: Firms' proxy voting may not have aligned with their stated ESG approaches, raising concerns about whether they were following their own ESG guidelines.

- *Compliance Programs:* The SEC observed that some firms lacked sufficient compliance oversight regarding their ESG investment processes and disclosures.

The Risk Alert emphasizes the importance of clear, consistent, and transparent ESG disclosures and practices. It also highlights the need for robust compliance programs to ensure that firms adhere to their stated ESG investment approaches and meet regulatory obligations.

In the last few years, the SEC came up with several proposals to address the aforementioned issues. For example, in March 2022 the SEC issued the Climate-Related Disclosures Proposal. The SEC proposed rules that would require public companies to disclose detailed information about their climate-related risks and greenhouse gas emissions. The proposed rules aim to provide investors with consistent, comparable, and decision-useful information for assessing climate-related risks. Later that year, the SEC issued another proposal: ESG Fund Naming and Marketing which proposed amendments to the "Names Rule" (Rule 35d-1 under the Investment Company Act of 1940) to address concerns about potential greenwashing in ESG funds. The proposed amendments would require funds with ESG-related names to ensure that at least 80% of their assets are invested in accordance with their ESG focus. Additionally, the SEC proposed enhancements to the disclosure requirements for ESG funds, including more detailed information about their ESG strategies and how ESG factors are integrated into investment decisions.

The SEC has also taken enforcement actions against entities that provided misleading or inadequate ESG disclosures. For instance, in 2021, the SEC settled charges with BNY Mellon Investment Adviser for misrepresentations and omissions regarding ESG considerations in decision-making for specific mutual funds. BNY Mellon was penalized $1.5 million and ordered to cease-and-desist from future violations. In 2022, Goldman Sachs Asset Management faced charges for failing to adhere to its ESG investment policies and procedures for certain mutual funds and separately managed accounts, resulting in a $4 million penalty and a cease-and-desist order. Furthermore, Wahed Invest, LLC, a New York-based robo-adviser, was charged with making misleading statements, breaching fiduciary duty, and compliance failures in its Shari'ah advisory business. Despite marketing itself as providing advisory services compliant with Islamic, or Shari'ah law, and emphasizing the importance of its income purification process, the SEC discovered that Wahed Invest failed to establish and implement written policies and procedures to ensure ongoing Shari'ah compliance.

Given the time period when the non-ESG funds in our sample converted to ESG, they were subject to limited regulatory or certification requirements to convert to an ESG label. Significant attention to ESG disclosures increased at the end 2020 when the SEC's Asset Management Advisory Committee emphasized the importance of ESG investment implementation and recommending standardized disclosure practices. Combined with the aforementioned "Climate-Related Disclosures Proposal" and "ESG Fund Naming and Marketing Proposal" issued in 2022, nowadays ESG funds (as well as former non-ESG funds converted to ESG) are subject to greater scrutiny from the regulators, especially concerning disclosure requirements. However, for the funds in our sample, which converted before 2022, we do not expect these regulations to have had much impact.

Related Literature

This chapter relates to three strands of the literature. First, the chapter contributes to the growing literature on the ESG and socially responsible investing. There is mixed evidence on whether mutual funds that label themselves as ESG-oriented actually invest in firms with superior ESG performance. Gibson Brandon et al. (2022) find that only international US-domiciled institutions that publicly commit to responsible investing by signing the United Nations Principles for Responsible Investment (PRI) are more likely to pick firms with better

ESG scores, while there is no such evidence for US PRI signatories. Similarly, Kim and Yoon (2023) show that PRI signatories do not improve fund-level weighted average ESG scores and exhibit a decrease in returns after signing the PRI. Raghunandan and Rajgopal (2022) investigate whether ESG funds make investments that align with stakeholder interests. They analyze employee satisfaction ratings, environmental regulatory violations, and federal contract allocation data of U.S. portfolio companies held by ESG funds compared to non-ESG funds, and find no significant differences in employee satisfaction or environmental regulatory violations between companies held by ESG and non-ESG funds. In contrast, Dikolli et al. (2022) find that ESG funds are more likely to vote in favor of ESG proposals compared to non-ESG funds, suggesting that ESG funds do "walk the talk" and align their voting behavior with their marketed ESG orientation. This chapter contributes to the literature by examining the behavior of non-ESG mutual funds that were subsequently repurposed as ESG funds. We find that these funds' ESG ratings improved significantly following their conversion.

Second, there is a strand of literature that examines performance of ESG funds. Research on the performance of sustainable, ESG, and impact mutual funds has yielded mixed results (see Gillan et al. (2021) and Starks (2023) for a more detailed review of the existing literature). Bauer et al. (2005) found no evidence of significant differences in risk-adjusted returns between ethical and conventional funds for the 1990-2001 period. Domestic US and UK ethical funds exhibited lower risk-adjusted returns than conventional funds, but the difference was not statistically significant. Their results suggest that investors can pursue ethical objectives without compromising financial returns. Renneboog et al. (2008) found that while SRI funds in the US, UK, and many European and Asia-Pacific countries underperformed their benchmarks, risk-adjusted returns were not statistically different from conventional funds in most countries. They suggest investors may be willing to accept lower financial returns for social or moral considerations. Similarly, Liang et al. (2022) document that hedge funds managed by PRI signatories underperform other hedge funds by 2.45% per annum after adjusting for risk factors. Barber et al. (2021) examined impact investing in venture capital and found impact funds earned 4.7 percentage points lower IRRs compared to traditional VC funds. Using a willingness-to-pay model, they estimated investors accept 2.5-3.7 percentage points lower expected returns for impact funds. Hartzmark and Sussman (2019) exploit a quasi-natural experiment launched by Morningstar in 2016 and provide evidence that investors value sustainability (see also Bauer et al., 2021). The findings reveal that funds categorized as low in sustainability experienced net outflows exceeding $12 billion, while those categorized as high in sustainability saw net inflows over $24 billion. However, the study found no evidence that high-sustainability funds outperformed low-sustainability funds in financial terms. This chapter contributes to the literature by examining the performance of non-ESG funds that were later repurposed as ESG funds. We show that the conversion is associated with higher fund flows and higher annual flows approximately two years after the conversion than the control funds.

Third, there is an growing literature on the choice of firms to engage in ESG activities. Firms might decide to switch from "brown" to clean technologies when there are strong incentives for innovation and adoption of clean technologies. Acemoglu et al. (2012) develop a theoretical model to analyze this transition, demonstrating that government policies creating incentives for clean technology innovation, such as carbon taxes or clean research subsidies, can drive firms to switch from dirty to clean technologies, especially when clean and dirty inputs are highly substitutable. Empirical studies support this theoretical framework. Flammer (2021) finds that firms issue green bonds to finance environmentally friendly projects when facing greater environmental risks, leading to improvements in their environmental performance, as evidenced by an increase in environmental ratings. Krueger et al. (2020) survey institutional investors, revealing that they believe climate risks have financial implications for their portfolio firms and that these risks, particularly regulatory risks, have already begun to materialize. Consequently, many institutional investors actively engage with firms on climate-related issues and incorporate climate risks into their investment processes, with long-term investors more likely to consider these risks compared to short-term investors. Gillan et al. (2022) examine how different types of owners affect a firm's likelihood of adopting environmentally friendly practices and find that one standard deviation increase in the interaction of investor demand and the largest shareholder's voting rights is associated with an about 8.5% higher environmental score. Our study contributes to this literature by exploring the decision to switch to ESG investing in the context of mutual funds. We shed light on the characteristics that may drive fund managers or fund families to convert a fund from conventional to ESG, extending the understanding of the factors influencing the adoption of ESG practices.

2.3 Data

The data we utilize comprise three primary datasets: Morningstar data to identify non-ESG funds that have been repurposed as ESG funds; CRSP data for mutual fund holdings and characteristics; and finally, MSCI ESG ratings on individual firms. The detailed description of each dataset could be fund in Chapter 1, Section 1.2.

Identifying Repurposed ESG funds

We use Morningstar's "Sustainable Funds U.S. Landscape Report" in order to identify non-ESG funds which were repurposed as ESG funds. In total there are 58 non-ESG equity funds which had been repurposed as ESG funds by the end of 2021 in our sample.

MSCI ESG ratings

There exist many different ESG ratings provided by multiple financial analytic firms. In this chapter, we will use ESG ratings of individual companies provided by MSCI. MSCI has been providing ESG ratings since at least 1999. However, because MSCI uses tickers as firm identifiers and given the poor quality of mutual fund holdings data in the 2000s, we will consider only the period from 2011 to the end of 2021. This should not have any implications for this project since the first repurposing of a non-ESG fund happened back in 2013.

CRSP

Lastly, we use data from the Center for Research in Security Prices (CRSP) on mutual fund holdings and characteristics. Specifically, for each mutual fund with holdings exceeding 100 million dollars, we have access to quarterly data on fund holdings, management and expense ratios, returns, assets under management (AUM), managing team, and numerous other attributes. Importantly, we also observe the inception date of each fund, enabling us to determine fund age.

From CRSP data on assets under management, we can calculate fund flows according to the following formula:

$$flow_t = \frac{AUM_t - AUM_{t-1} * (1 + r_t)}{AUM_{t-1}} = \frac{AUM_t}{AUM_{t-1}} - r_t - 1, \tag{2.1}$$

where AUM_t is the fund AUM at t, and r_t is the fund returns in period t (i.e., between periods $t-1$ and t).

Construction of ESG ratings

Construction of ESG ratings at the fund level could be found in Chapter 1, Section 1.2 (specifically, Equation 1.1).

Construction of Alphas

We use Fama-French Three factor model (Fama and French, 1993) to estimate before-fee risk-adjusted fund-performance:

$$r_{it} = \alpha_i + \beta_{rm,i} rm_t + \beta_{smb,i} smb_t + \beta_{hml,i} hml_t + \epsilon_{it}, \tag{2.2}$$

where r_{it} is fund i's before-expense return in month t in excess of the 30-day risk-free interest rate – proxied by 1-month Treasury bill rate; rm_t is the market portfolio return in excess of the risk-free rate; and hml_t and smb_t denote the return on portfolios that proxy for book-to-market and size risk factors, respectively. Note that CRSP data on mutual funds report after-expense returns. To retrieve monthly before-expense returns, we add back the annual expense ratio divided by 12 to the reported returns.

Following Carhart (1997), we use a two-stage estimation process to derive monthly risk-adjusted performance estimates. In the first stage, for each month t, we regress the before-fee excess returns of the funds on the relevant risk factors using data from the previous five years. If a specific fund-month has less than five years of available data, we include funds that have at least 36 months of data within the previous five years and conduct the regression with the available data. In the second stage, we calculate the fund's risk-adjusted performance for month t by calculating the difference between the fund's before-expense excess return and the excess return predicted by the Fama-French three-factor model. Once we obtain monthly alphas, we convert them to quarterly alphas to match the quarterly period of our data sample.

Matching

We use propensity matching to define a set of non-converted funds that are similar to converted funds, so that we have a comparable comparison group. We construct a sample of non-ESG funds which were never repurposed in the sample period. For this, we use the K-nearest neighbor method for $K = 5$.[3] For the matching procedure, we use propensity score matching based on a set of fund characteristics: average AUM over the last year, annual returns, annual fund flows, the number of quarters with negative flows over the last two years, and fund age. Since all the repurposed funds were converted at different times,

[3] The results are qualitatively similar for other values of K, including 3 and 10. Note, however, larger K results in more funds in the sample which is especially beneficial given the limited number of converted funds.

in the matching procedure we cannot use cross-sectional data for matching. Therefore, for each converted fund, we match that fund with non-ESG funds one year before the fund's conversion.

Summary statistics

Table 2.1 presents summary statistics of the control and treated groups in the sample with matched funds, and in Table 2.2, we provide the results of balance test for matching procedure. Overall we have 257 funds from 2011 to 2022. This number is lower than $K \times 58$ (the size of the treated group) because some of the never-repurposed funds are used as matches for more than one treated fund.

Figure 2.2 presents the yearly number of non-ESG funds converted to ESG. It illustrates that fund repurposing is a relatively new phenomenon: the first repurposing occurred in 2013, and it gained popularity after 2017, with almost half of the repurposings taking place in 2020 and 2021. Figure 2.3 demonstrates the average age at conversion for non-ESG funds repurposed by the end of 2021. Remarkably, approximately half of the converted funds were at least 20 years old at the time of conversion.

2.4 Empirical Strategy

In order to study the effect of ESG conversion on fund ESG rating and its financial outcomes as well as to explore whether some changes precede the conversion, we employ two empirical approaches. In the first one, we study the converted funds complemented with never converted matched funds as a comparison for the converted funds. In the second approach, we focus only on the converted funds.

The first empirical approach includes the funds that never converted during the sample period as a comparison group. Due to the fact that conversion of the funds does not happen at the same time but rather spread across multiple periods, the main empirical framework we employ is the dynamic difference-in-difference approach. Using a sample of ESG conversion funds and matched non-ESG funds, we implement the following empirical strategy:

$$y_{it} = \lambda_t + X_{it} + Conversion_i \times \left[\sum_{j=-h, j \neq -1}^{H} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it}, \tag{2.3}$$

where y_{it} is the outcome of interest for fund i and quarter t. We include quarter fixed effects λ_t to account for temporal variation and trends in the data. Fund characteristics X_{it} include fund age, fund type and the share the holdings in common stock. An indicator $I\{t - t_j^* = j\}$ captures the relative time to conversion and is equal to one if a fund is j quarters away from being repurposed. Parameters h and H are the number of quarters before and after the conversion that are considered in the analysis. A dummy variable *Conversion_i* is equal to one if fund i is a conversion fund, and zero otherwise. The coefficients of interest are β_j that measure the evolution in outcome y_{it} before and after the conversion relative to a quarter immediately preceding the ESG conversion. Standard errors are clustered on the industry level.

Additionally, we also use the static difference-in-differences approach to estimate the average effect of a fund repurposing on ESG rating and other outcomes:

$$y_{it} = \beta D_{it} + \lambda_t + X_{it} + \varepsilon_{it}, \tag{2.4}$$

where D_{it} is a dummy variable which is equal to 1 if a fund has been already repurposed by time t, and 0 otherwise. As in the dynamic version of this regression, we include quarter fixed effects λ_t to account for temporal variation and trends in the data, while X_{it} contains fund characteristics. The parameter of interest, β captures treatment effect of fund conversion on the outcome of interest. The implicit assumption in the difference-in-difference analysis is that treatment and control groups evolve along parallel trends in the absence of the treatment. We evaluate the plausibility of parallel trends assumption by assessing the pre-conversion coefficients in the dynamic DID estimation.

In our setting the treatment (i.e., the event when a non-ESG fund converted to an ESG fund) is not exogenous – there is no government agency or regulator which randomly would force a fund to convert. Instead, it is natural to assume that there are fund characteristics which would incentivize the fund manager or fund family to convert a particular non-ESG fund to ESG. Apart from fund characteristics, there might be other considerations for a fund family to convert a fund: for example, the fund family might want to appeal to a broader investor base, specifically to investors with preference for sustainable investments, or the fund family might use the converted fund for marketing purposes. Predictably, the above arguments undermine our ability to draw causality between conversion and fund outcomes as the treatment is not assigned randomly.

The regressions above could be safely interpreted as associations between funds' ESG conversions and the outcomes. However, once we are able to control for different fund characteristics and match the initial sample with never-converted funds, we might hope to alleviate the selection concern at least to some degree. For example, if the pre-treatment coefficients are statistically indistinguishable from zero, then one might cautiously consider the post-treatment coefficients as causal treatment effect.

The intuition here is similar to the one in synthetic control studies (e.g., see Abadie and Gardeazabal, 2003, Abadie et al., 2010, Abadie, 2021, Doudchenko and Imbens, 2016). When treatment happens at the country level and the number of treated units is low (typically there is only one treated unit), one method is to create a "synthetic" version of the treated unit. The "synthetic" treated unit then used as a counterfactual in order to estimate the treatment effect. For example, in Abadie et al. (2015) the authors estimate the effect of German unification on the economic development of West Germany (Federal Republic of Germany). The idea there is to assign weights to other European countries in order to create a synthetic West Germany which would have very similar economic outcomes before the unification in 1990. Then the authors – after conducting a number of robustness/placebo checks – would attribute any post-reunification difference in GDP between the synthetic and real West Germany to the treatment effect, i.e., to German reunification in 1990. In our context, the number of treated units is high, 58, which does not make the synthetic control estimation a very viable option. However, the intuition is about the same: if we have statistically indistinguishable from zero coefficients in the pre-treatment periods this might give us some confidence in interpreting the post-treatment coefficients in a causal way.

This logic is especially applicable to the analysis with the matched sample. Indeed, for each treated fund we find K regular "never treated" non-ESG funds and then we run a balance test to make sure that the treated funds and the matched funds are comparable in the pre-treatment periods (recall that since different funds might be repurposed/converted at different times the pre-treatment periods are also different).

Second, we focus only on conversion funds, comparing converted with yet-to-be converted funds. We use an empirical strategy very similar to (2.3) to evaluate the evolution of outcomes in converted funds before and after the conversion:

$$y_{it} = \lambda_t + X_{it} + \sum_{j=-h, j\neq -1}^{H} \beta_j I\{t - t_i^* = j\} + \varepsilon_{it}. \tag{2.5}$$

Lastly, we conduct heterogeneity analysis. It is probable that funds with different characteristics would have different reasons for conversion to ESG as well as varying effects of conversion on fund outcomes. We explore fund heterogeneity in terms of size defined by AUM, expense ratio and financial performance defined by Fama-French three-factor alpha. For each characteristic, we classify funds into high and low category as above and below median in the period from 2011 to 2017, when the overwhelming majority of funds have not yet converted to ESG status. Then, we estimate equations 2.3 and 2.4 separately for funds in high and low category.

2.5 Results

In this section, we report the results of the estimation of Equations (2.3) and (2.4) for the whole sample of 257 funds, including conversion funds and matched never-treated funds. . The results for conversion funds only, along with the corresponding graphs, are presented in Appendix B.

ESG Ratings

We start the analysis by examining how fund ESG ratings change with fund repurposing. The primary purpose of this exercise is twofold. First, we want to examine whether to-be-repurposed funds engage in any sort of preparations, e.g., if they reorganize their portfolio with more weight on firms with higher ESG ratings. Indeed, since fund repurposing is not an exogenous event and it is up to fund managers or fund families to make such a decision, these funds might be tilting their portfolios towards ESG-friendlier individual firms. Another possibility is that the decision to repurpose a fund might be made due to poor ESG rating of the fund – for example, if the fund manager believes that this will please ESG-conscious investors and will positively impact the fund flows. Second, we want to explore if fund repurposing is associated with higher fund ESG rating. One potential possibility is that fund repurposing is entirely about changing labels and other marketing activities. On the other hand, since investors have (imperfect) information about fund ESG rating, this might motivate the funds actually "walk the talk", i.e., rebalance their portfolio towards firms with higher ESG ratings. Figures 2.4, 2.5, and B.1 illustrate the evolution of ESG ratings over time.

In Figures 2.4 and 2.5, the sample includes both treated funds, i.e. those that had been converted to ESG by the end of 2021, and control (matched) funds. In Figure 2.4 shows that treated funds had slightly lower industry-adjusted ESG scores about two years before the repurposing but the difference almost entirely disappears about nine months before the treatment. After the repurposing, the ESG scores of the treated funds increases (relative to the control group of funds) and the difference seems to widen about two and a half years after conversion. By five years after the conversion, industry-adjusted ESG rating increases by 0.8 standard deviations relative to the control group. Figure B.1 (a) has a very similar qualitative pattern but, as expected, the estimates are a little noisier due to smaller sample size of funds that eventually convert to ESG. Figure 2.5 illustrates the effect on other ESG ratings, specifically, weighted average ESG rating, and individual ratings for environmental, social and governance performance. The pattern of changes for the weighted average ESG rating is qualitatively similar, although the estimates are less precise and in some cases are not statistically significant. Comparing environmental, social and governance ESG ratings in panels (b)-(d), the increase is the highest for social rating and is not statistically significant for other ratings. This finding suggests that funds that convert to ESG status mostly focus on social component of ESG. Overall, ESG performance of the treated funds improves in the post-treatment period, and treated funds start rebalancing their portfolios toward high ESG firms about 2 years before conversion to ESG.

Table 2.3 reports the results of estimation of equation 2.4 to examine the average post-treatment effect on ESG ratings. Column (1) reports the effect on industry-adjusted ESG ratings, in column (2), the outcome is weighted average ESG rating, and in columns (3)-(5) we consider Environmental, Social, and Governance ratings, respectively. In all specifications, there is a positive effect of ESG repurposing on ESG ratings, although estimates are of different magnitude and statistical significance. Column (1) indicates that fund repurposing is associated with increase in industry-adjusted ESG rating by about 0.47 standard deviations on average in the five years after repusposing. Column (2) shows that the effect on the weighted average ESG rating is smaller and around 0.2 standard deviations. Results from Columns (3)-(5) indicate that social fund rating increases the most: by around 2.7 standard deviations in the post-period, while the effects on environmental and governance ratings are smaller and are less statistically significant. Table B.1 presents the results of the event study regression from equation 2.3.

Fund Flows

Previous research suggests that fund flows are extremely important for fund managers being the key aspect of their performance evaluation and compensation (Cen et al., 2023, Ma et al., 2019). Naturally one would wonder whether decisions to repurpose a fund at least partially were driven by the managers' desire to attract investors capital to the funds. Again, as with the other outcomes we have looked at, there are two periods we are interested in: pre-treatment and post-treatment. Before treatment, if we observe that, for example, treated funds have lower fund flows then it might indicate that funds were converted at least partially due to poor fund flows with the hopes to attract additional capital from investors. In the absense of pre-treatment differences, we could attribute any post-treatment difference in fund flows to the treatment. In other words, if fund flows diverge after treatment one could interpret this as the treatment effect of fund repurposing.

Figure 2.6 shows the evolution of the logarithms of quarterly and annual fund flows for treated relative to control funds over time. In Figure 2.6 (a), we observe that, relative to the control group, the annual fund flows of repurposed funds were decreasing before conversion, although most estimates are not statistically significant. Post-treatment, there is no significant difference between treated and control funds in the first several quarters. Naturally, it takes time to see any difference in annual flows unless the change in quarterly flows is very large. However, after about one and a half years, there is an increase in annual flows in repurposed funds compared to the control group. Panel (b) of Figure 2.6 shows similar pattern for quarterly flows, with the difference between treatment and control funds that starts to emerge after two quarters from treatment and incerases over time, although many of the effects are noisy.

Another metric of flow performance is how often a fund experiences outflows. Panel (c) of Figure 2.6 explores the effect on the number of quarters with fund outflows over the last two years. For example, if a fund experienced only fund inflows over the last two years, the metric is equal to zero. Conversely, if a fund was bleeding money for the last two years, the metric is equal to 8.[4] Panel (c) of Figure 2.6 shows that before repurposing, treated funds experienced an increase in the number of quarters with negative fund flows over the last two years, with statistically significant effects for around two years before the conversion. The trend continues in the first three post-treatment periods, as the number of quarters with outflows continues to increase in treated funds relative to the control funds. This pattern reverses after three quarters, and after about one and a half years repurposed funds have lower occurrence of outflows than the control funds. The point estimates suggest further decline in subsequent quarters, although most estimates are not statistically significant. Overall, the dynamic effects on fund flows suggest that increase in occurrence of fund outflows and the corresponding decline in flows could be one of the reasons for fund conversion: the management might hope to attract ESG conscious investors. After the conversion, flows increase but not immediately, which might happen if there are information frictions or herding behavior, in line with Nofsinger and Sias (1999). Also note that ESG rating increases immediately after the conversion but also substantially accumulates over time, which could explain investors waiting for substantial change in ESG rating to invest in a given ESG fund. In addition, cautious investors might wait to see whether ESG conversion would lead to a decline in financial performance before investing, since ESG is a novel phenomenon and investors might not have enough historical information to form reliable expectations.

The corresponding DID results are presented in Columns (4)-(6) of Table 2.4 with estimates of equation 2.4. For all three measures of fund flows, the effect is positive but not statistically significant. Motivated by the patterns of dynamic effects, we examine the effect on six quarters leads of fund flows. Table 2.5 reports the results of the difference-in-differences estimation on outcome leads. Column (1) indicates that from six to twenty quarters after conversion, the treated funds have 15.1 percent higher annual fund flows (equivalent to 0.46 standard deviations) than the control funds relative to pre-period and five quarters after the treatment. Column (2) shows that fund repurposing is associated with about 3.3 percent

[4] The findings are robust for other time windows, e.g., two and a half or three years.

higher quarterly fund flows (equivalent to 0.4 standard deviations) than the control funds in the period from six to twenty quarters after the conversion. The effect on the number of quarters with outflows in Column (3) is not statistically significant, however, DID does not take into account the differential pre-trends for this outcome.

Other fund outcomes

Other outcomes that might be affected by the fund conversion to ESG status are measures of financial performance, expense ratio or profits. Figure 2.7 presents estimates from equation 2.3 for fund financial performance measured as annual returns and Fama-French three-factor alpha, as well as expense ratio or profits. There are no consistent patterns of change for any of the outcomes and almost all of the estimates are not statistically significant. The corresponding DID estimates in Columns (1)-(3) and (7)-(8) of Table 2.4 are small and not statistically significant. These results suggest that fund financial performance does not change as a result of ESG conversion, and ESG conversion does not lead fund managers to adjust fund expense ratio.

Subsample analysis

To explore potential explanations for the discussed results, we consider heterogeneity in effect by different subsamples of funds.

First, explore heterogeneity by fund size: we consider the funds with above and below median AUM in 2011-2017. Figure 2.8 presents the findings. Panel (a) illustrates that low- and high-AUM funds had similar trends in industry-adjusted ESG ratings both pre- and post-treatment. About three years after the conversion, low-AUM funds had higher increase in ESG rating than high-AUM funds, although the estimates are noisy. Panel (b) shows the effect on log of annual flows. Point estimates suggest that, in the pre-treatment period, the flows into high-AUM and low-AUM funds exhibited opposite trends relative to their corresponding control groups. Low-AUM funds experienced decreasing flows, while high-

AUM funds experienced an increase in flows. However, many effects are not statistically significant. After the treatment, low-AUM funds exhibited an increase in flows compared to their control group, with effects being statistically significant from about a year and a half after the treatment. In contrast, high-AUM funds initially saw a decrease in flows (not statistically significant), which zero effect starting from around a year after the conversion. Panel (c) indicates that high-AUM funds increased their expense ratios in the post-treatment period relative to their control group, while there was no effect for low-AUM funds. Lastly, Panel (d) shows that there is no difference between low- and high-AUM funds in terms of the effect on the before-fee risk-adjusted returns. Overall, Figure 2.8 indicates that the observed results on fund flows for the entire sample are primarily driven by low-AUM funds. Moreover, low-AUM funds not only exhibit increased ESG rating and fund flows after the conversion, but also increase their expense ratio.

Second, we divide the sample into two groups based on the fund expense ratio: funds with above-median and below-median expense ratios from 2011 to 2017. Figure 2.9 (a) demonstrates that there was a pre-trend for low-fee funds that experienced a small increase in industry-adjusted ESG ratings before the conversion in the two years before the conversion. After the conversion, low-fee conversion funds had significantly higher ESG ratings than the corresponding control group, with statistically significant effects in each post-treatment quarter. In contrast, high-fee conversion funds had similar patterns to their control group before the treatment, experienced an increase in ESG rating after the treatment, although not statistically significant, with dynamic effects returning to zero after about two years after the conversion. Panel (b) illustrates similar effects of ESG conversion on annual fund flows across the two groups. Panel (c) shows the effect on expense ratio. Although the estimates are not statistically significant, the point estimates suggest that funds with high expense ratios tend to decrease their fees in the post-treatment period, while low-fee funds tend to increase their fees. Finally, Panel (d) shows the estimates for the Fama-French three-factor alphas. While most of the estimates are not statistically significant, some in the first few post-treatment periods indicate that high-fee funds tend to perform poorer than their control group. To sum up, the observed patterns for two groups suggests that the overall result for ESG ratings in the entire sample is driven by low-fee funds.

Third, we divide the sample into two groups based on risk-adjusted fund performance: funds with above-median and below-median alphas from 2011 to 2017. Figure 2.10 shows that effects of conversion on ESG ratings, annual flows, or FF3 alphas are similar for the two groups. Patterns for expense ratio in Panel (c) are somewhat different but most of the estimates are not statistically significant.

Overall, the subsample analysis highlights the heterogeneity of funds and their behavior before and after conversion: (i) the results on annual fund flows are mainly driven by low-AUM funds, and (ii) the changes in ESG ratings are primarily explained by low-fee funds rebalancing their portfolios toward more ESG-friendly firms.

2.6 Discussion and Conclusion

The question of whether decision makers in mutual funds and individual firms do good with investors' capital has always been of interest for researchers. For example, Cheng et al. (2023) explore whether firm managers may overinvest in CSR to pursue their own social preferences instead of maximizing shareholder value. In the context of mutual funds, an important concern is the possibility of fund managers engaging in greenwashing to attract more capital and boost their own compensation. In this chapter we address this question in the context of non-ESG funds which were repurposed as ESG funds. The findings tend to indicate that at least in terms of ESG ratings of portfolio holdings repurposed funds improve their ESG performance.

As depicted in Figure 2.2, the inaugural instance of fund repurposing was observed in 2013. This observation raises a pertinent question: What factors contributed to the delayed emergence of this practice? One potential driver of the observed trend could be investors' growing preference for "clean" stocks, a term that encompasses not only environmental considerations such as emissions and pollution but also broader aspects of corporate governance and social responsibility. This shift in investor preferences is reflected in the rising inflows into socially responsible investment (SRI) funds. According to the US SIF Foundation's 2020 Report on US Sustainable and Impact Investing Trends, sustainable investing assets reached $17.1 trillion at the start of 2020, marking a 42% increase from $12 trillion in 2018. This growth indicates a significant shift towards investments that consider ESG criteria.

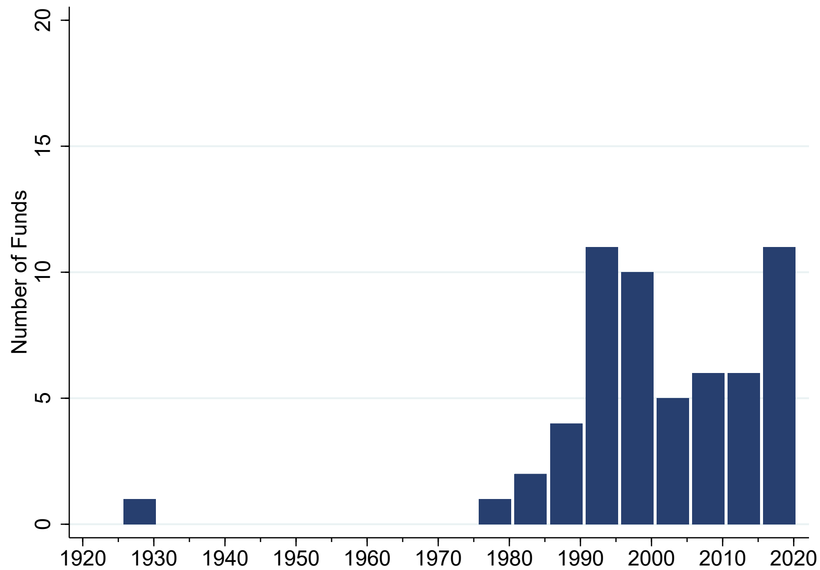
Furthermore, a growing body of research suggests that social issues are becoming increasingly important to investors. For example, a study by Edmans (2011) found that companies with high employee satisfaction, a key social factor, outperformed their peers in terms of stock returns. Additionally, the COVID-19 pandemic has further highlighted the importance of social considerations, with investors paying closer attention to companies' treatment of employees, supply chain practices, and community engagement.[5]

In this chapter, we examine the decision of non-ESG equity mutual funds to transition into ESG funds. Our analysis, employing propensity score matching and dynamic difference-in-difference methods, yields several important findings. First, we observe that the conversion of funds into ESG funds results in a significant enhancement of the funds' ESG ratings. This improvement suggests a genuine shift toward sustainable investing and provides evidence against the involvement of these funds in greenwashing. Second, we demonstrate that fund repurposing is associated with higher fund flows in the quarters following the transition. This finding aligns with the literature on investors' preference for sustainable investing. Additionally, we present evidence suggesting that repurposed funds may exhibit higher returns, although this result appears to be primarily driven by the outperformance of high-ESG stocks during our sample period.

[5]For example, see these articles:
(i) https://corpgov.law.harvard.edu/2020/10/20/survey-analysis-esg-investing-pre-and-post-pandemic/
(ii) https://www.jpmorgan.com/insights/global-research/esg/covid-19-esg-investing

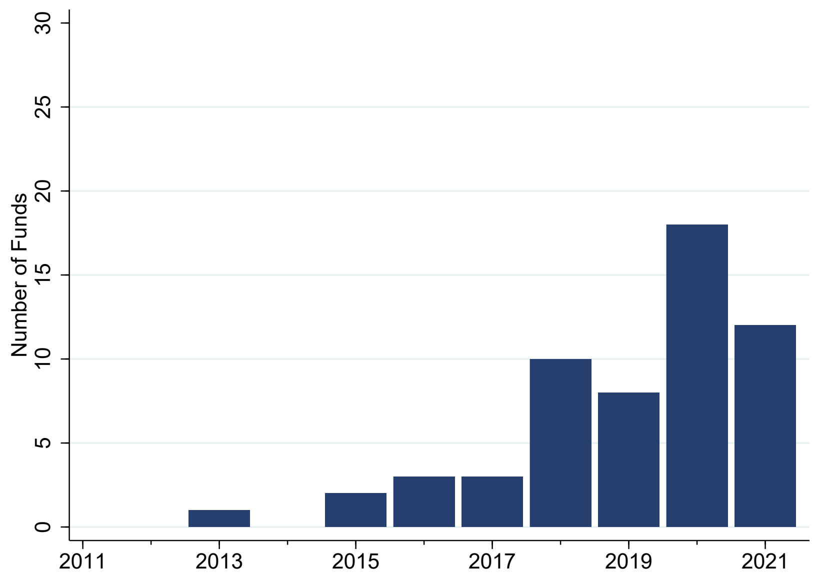
Figures and Tables

Figure 2.1: Inception of equity funds that eventually become ESG.



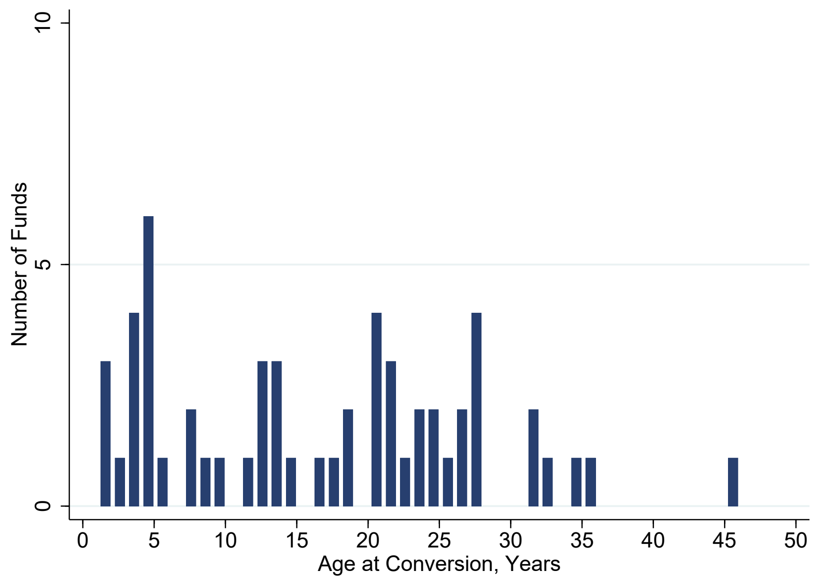
Note: The graph shows inception dates of equity non-ESG funds that were subsequently converted into ESG funds by the end of 2021.

Figure 2.2: ESG conversion of equity funds.



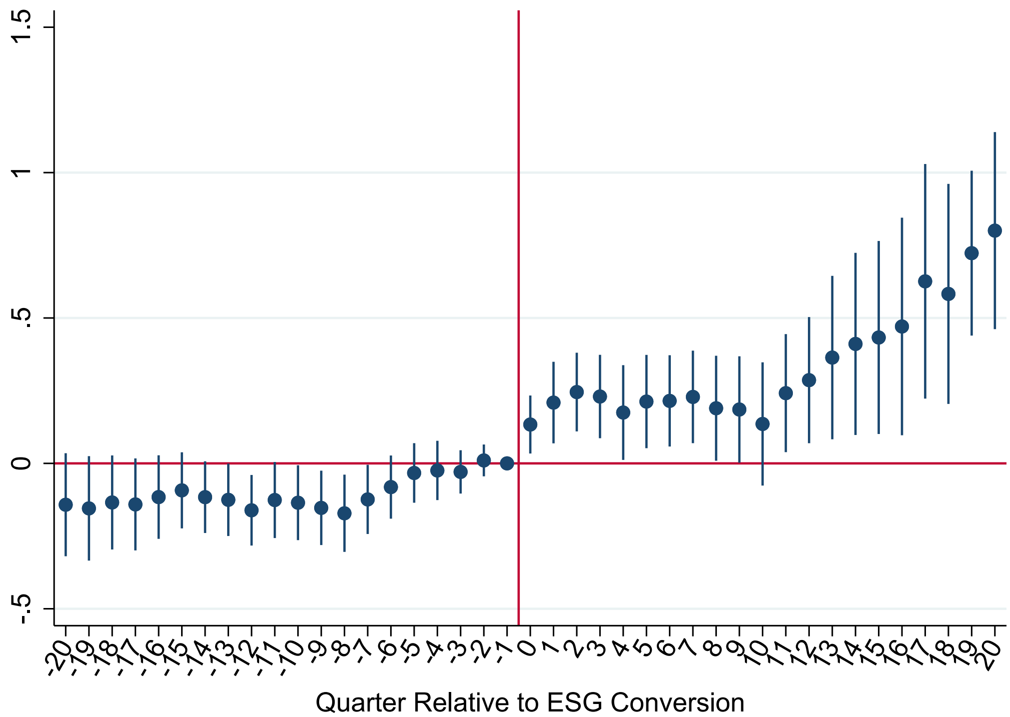
Note: The graph shows the yearly number of non-ESG funds that were transformed into ESG funds by the end of 2021.

Figure 2.3: Fund's Age at the Time of ESG Conversion



Note: The plot shows the age (in years) of equity non-ESG funds at the time of their conversion into ESG funds, considering only those funds that were converted by the end of 2021. The Pioneer Fund, established in 1928 and repurposed in 2021, is not included in this figure.

Figure 2.4: Treatment Effect of non-ESG Fund Repurposing on ESG Ratings



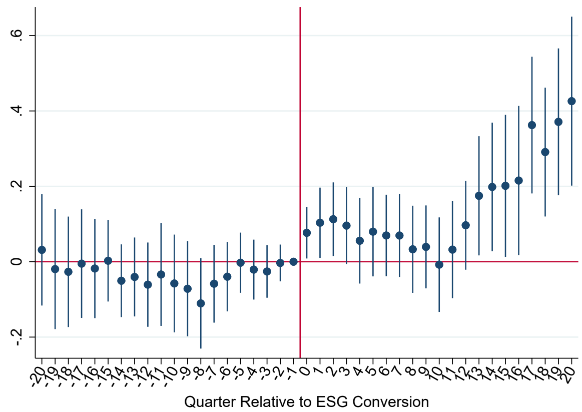
Note: The figure shows the dynamic treatment effect of non-ESG funds conversion to ESG on funds ESG industry-adjusted rating estimated from the following specification:

$$Rating_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + Conversion_i \times \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

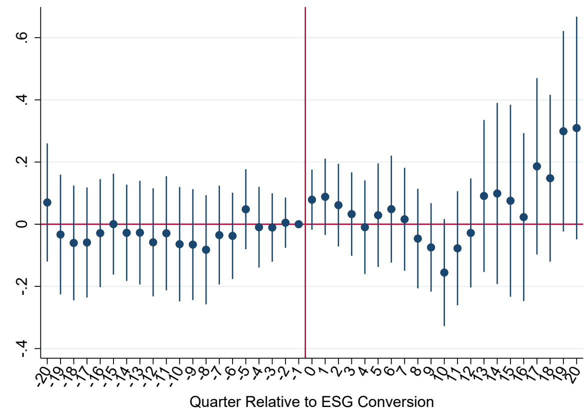
where $Rating_{it}$ is fund's i industry-adjusted ESG rating, λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. $Conversion_i$ is a dummy if fund i were repurposed by the end of 2021, and X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Figure 2.5: Treatment Effect of non-ESG Fund Repurposing on ESG Ratings

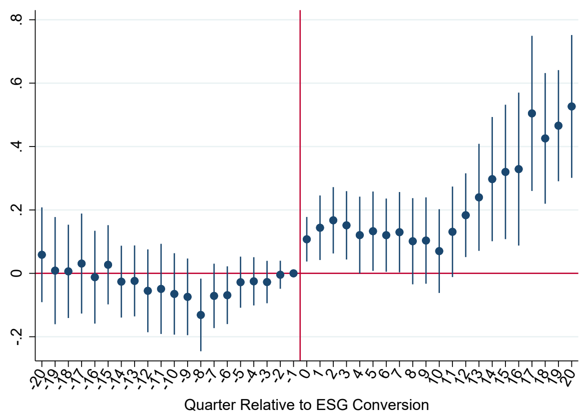
(a) Weighted Average ESG Fund Rating



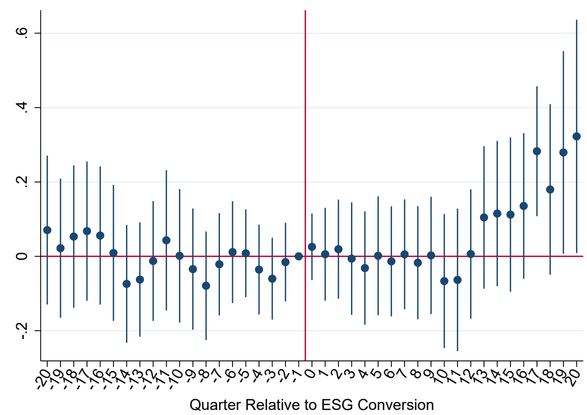
(b) Environmental Fund Rating



(c) Social Fund Rating



(d) Governance Fund Rating



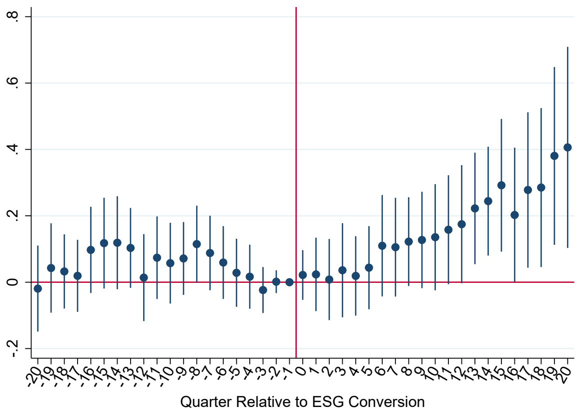
Note: The figures show the dynamic treatment effect of non-ESG funds conversion to ESG on funds ESG ratings estimated from the following specification:

$$Rating_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + Conversion_i \times \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

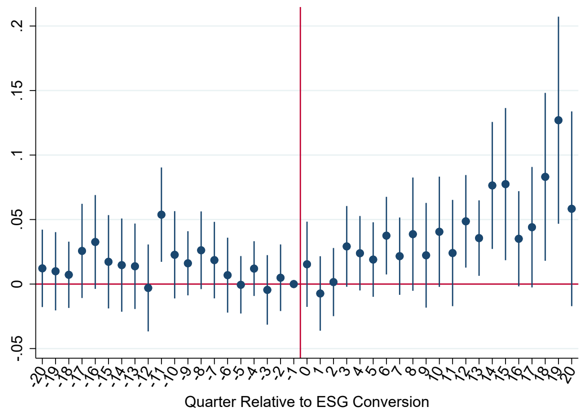
where $Rating_{it}$ is fund's i ESG rating, λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. $Conversion_i$ is a dummy if fund i were repurposed by the end of 2021, and X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Figure 2.6: Treatment Effect of non-ESG Fund Repurposing on Fund Flows

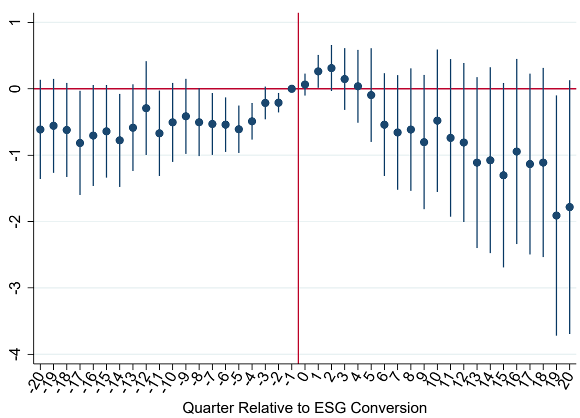
(a) Log of Annual Flows



(b) Log of Quarterly Flows



(c) Number of Quarters with Negative Fund Flows over the Last 2 Years



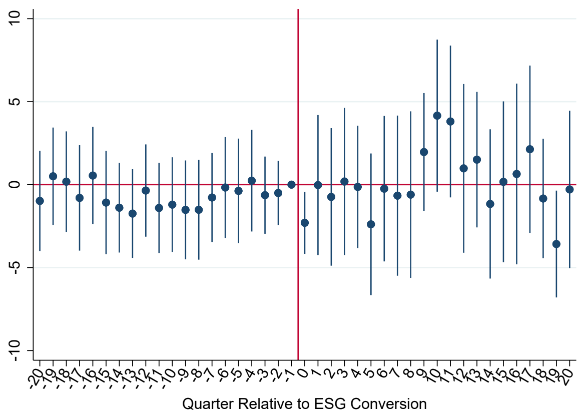
Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on fund flows estimated from the following specification:

$$Log(1 + Flow_{it}) = \lambda_t + \alpha_{j(i)} + X_{it} + Conversion_i \times \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

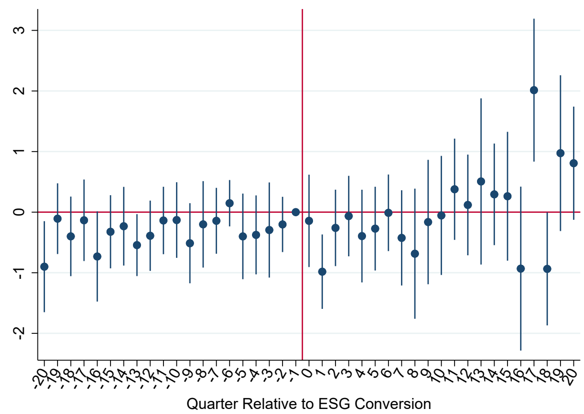
where $Flow_{it}$ is i fund's flows in period t, λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. $Conversion_i$ is a dummy if fund i were repurposed by the end of 2021, and X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Figure 2.7: Treatment Effect of non-ESG Fund Repurposing on Select Outcomes

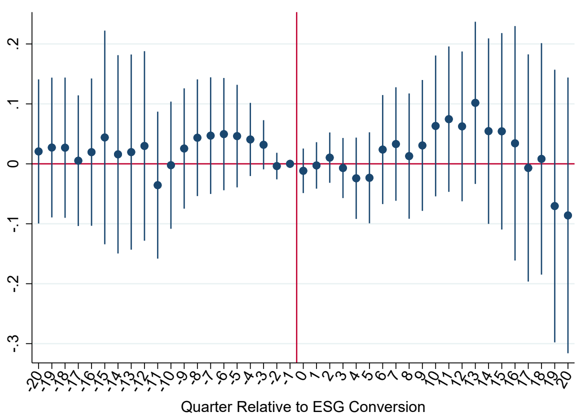
(a) Annual Returns



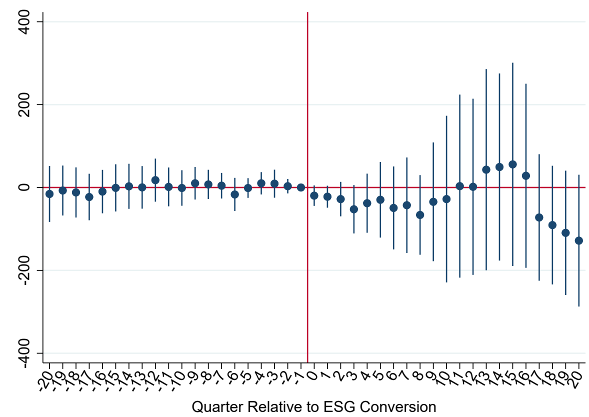
(b) Fama-French Three-Factor Alpha



(c) Expense Ratio



(d) Profits

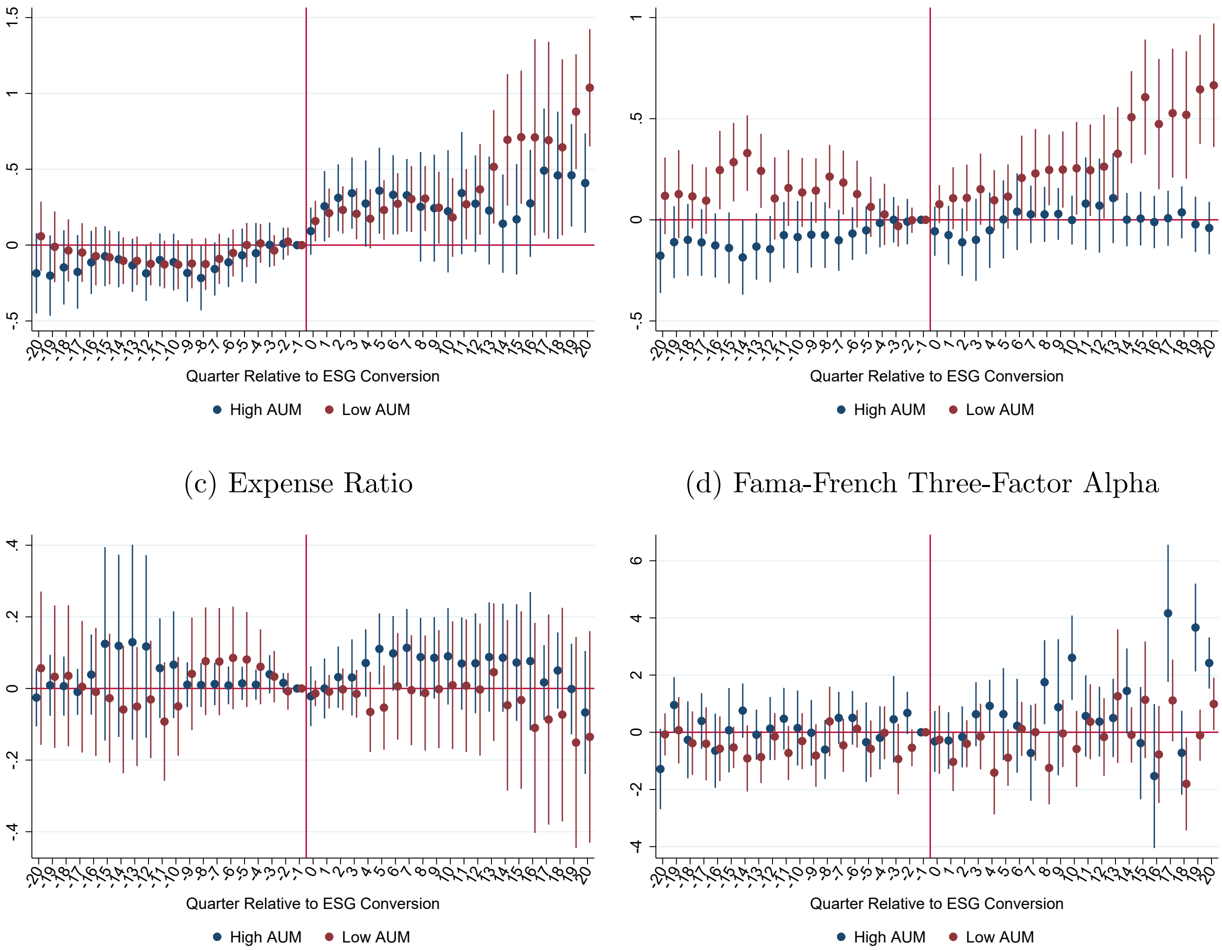


Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on different outcomes estimated from the following specification:

$$Outcome_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + Conversion_i \times \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

where $Outcome_{it}$ is (a) Annual Returns; (b) Quarterly Returns; (c) Expense Ratio; (d) Profits of fund i in period t. λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. $Conversion_i$ is a dummy if fund i were repurposed by the end of 2021, and X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Figure 2.8: Effect on Select Outcomes, by Fund Size



(a) Industry-Adjusted ESG Ratings

(b) Log of Annual Flows

(c) Expense Ratio

(d) Fama-French Three-Factor Alpha

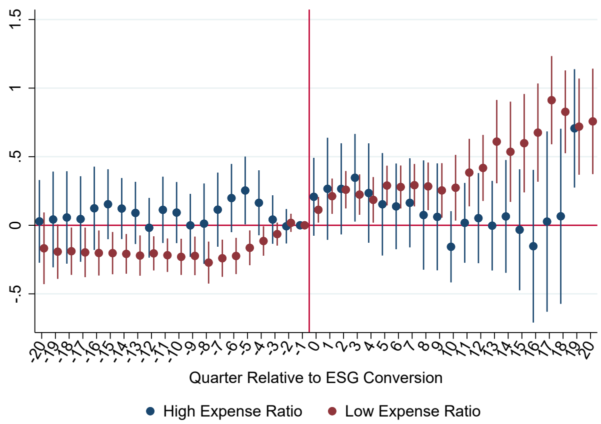
Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on different outcomes estimated from the following specification:

$$Outcome_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + Conversion_i \times \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

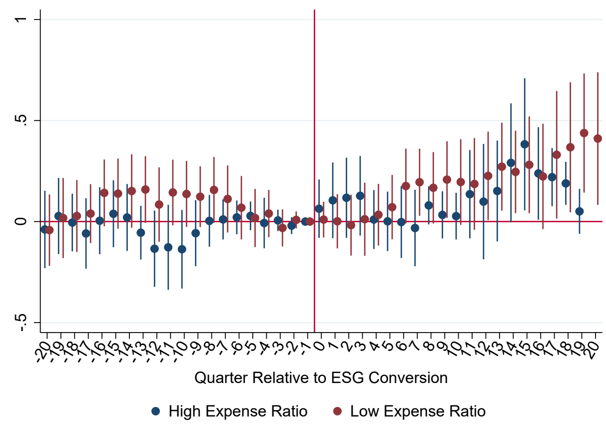
where $Outcome_{it}$ is (a) Industry-Adjusted ESG Ratings; (b) Log of Annual Flows; (c) Expense Ratio; (d) Quarterly Fama-French Three-Factor Alpha. λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. $Conversion_i$ is a dummy if fund i were repurposed by the end of 2021, and X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. The sample is divided into two groups: funds with above and below median AUM in 2011-2017. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Figure 2.9: Effect on Select Outcomes, by Fund Expense Ratio

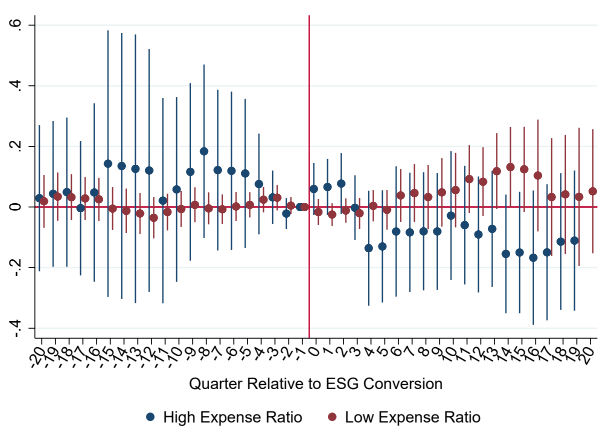
(a) Industry-Adjusted ESG Ratings



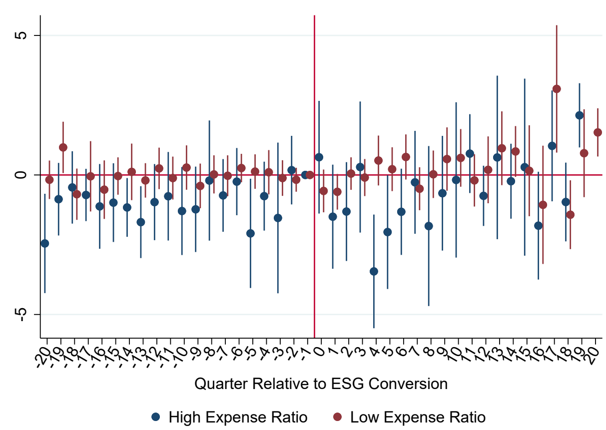
(b) Log of Annual Flows



(c) Expense Ratio



(d) Fama-French Three-Factor Alpha



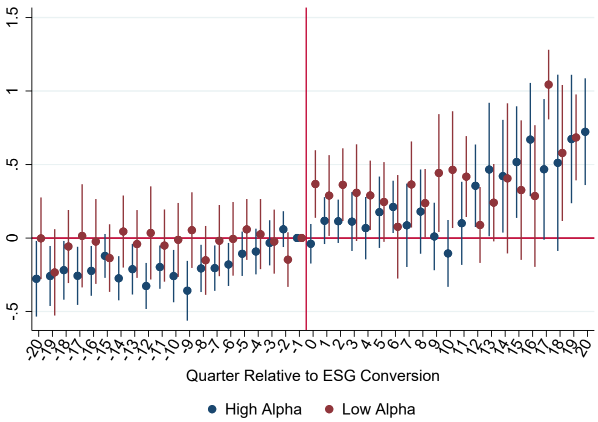
Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on different outcomes estimated from the following specification:

$$Outcome_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + Conversion_i \times \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

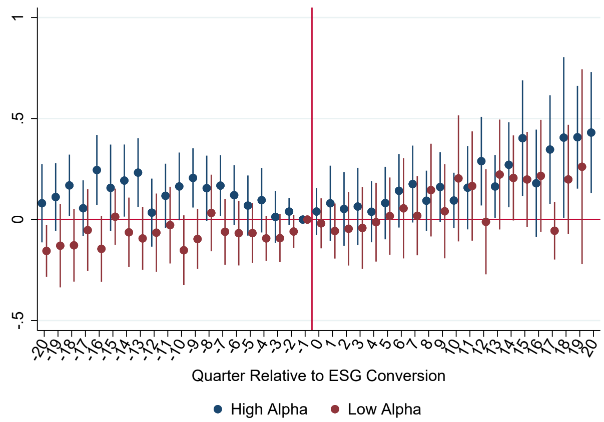
where $Outcome_{it}$ is (a) Industry-Adjusted ESG Ratings; (b) Log of Annual Flows; (c) Expense Ratio; (d) Quarterly Fama-French Three-Factor Alpha. λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. $Conversion_i$ is a dummy if fund i were repurposed by the end of 2021, and X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. The sample is divided into two groups: funds with above and below median expense ratio in 2011-2017. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Figure 2.10: Effect on Select Outcomes, by Fund Alpha

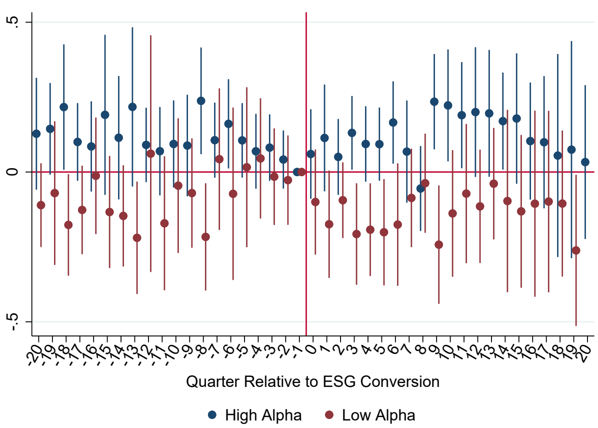
(a) Industry-Adjusted ESG Ratings



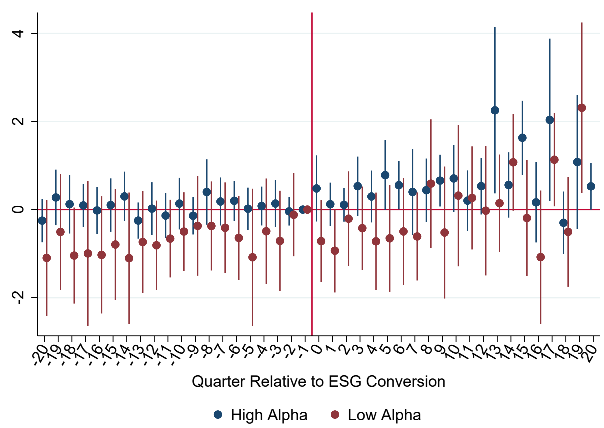
(b) Log of Annual Flows



(c) Expense Ratio



(d) Fama-French Three-Factor Alpha



Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on different outcomes estimated from the following specification:

$$Outcome_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + Conversion_i \times \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

where $Outcome_{it}$ is (a) Industry-Adjusted ESG Ratings; (b) Log of Annual Flows; (c) Expense Ratio; (d) Quarterly Fama-French Three-Factor Alpha. λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. $Conversion_i$ is a dummy if fund i were repurposed by the end of 2021, and X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. The sample is divided into two groups: funds with above and below median Fama-French three-factor quarterly alpha in 2011-2017. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Table 2.1: Summary Statistics

	All	Converted	Never Converted
	mean	mean	mean
Fama-French Three-Factor Alpha	-0.37	-0.34	-0.37
Carhart Four-Factor Alpha	-0.34	-0.34	-0.34
Annual returns	11.74	11.49	11.80
Quarterly returns	2.86	2.80	2.87
Expense Ratio	1.15	1.07	1.17
Turnover Ratio	0.70	0.61	0.73
log AUM	4.74	4.58	4.78
Age (years)	77.27	74.88	77.84
Quarterly flows	0.30	0.07	0.36
Annual flows	1.49	0.62	1.71
Quarters with Negative Fund Flows	5.02	5.00	5.03
ESG rating (ind-adj)	3.36	3.67	3.28
ESG rating (w.-av.)	3.21	3.37	3.16
Social Fund Rating	3.00	3.14	2.97
Environmental Fund Rating	3.72	4.00	3.64
Governance Fund Rating	3.65	3.78	3.61
Number of Funds	257	58	199
Observations	9868	2017	7851

Note: This table reports summary statistics of fund characteristics. The sample consists of 257 funds over the period from 2011 to 2022.

Table 2.2: Balance Test for Matching

Variable	Mean (Treated)	Mean (Control)	t-test	p-value
Average AUM	473.31	439.83	0.94	0.347
Age (quarters)	74.594	75.073	-0.23	0.822
Expense Ratio (%)	1.0651	1.2221	-8.74	0.000
Quarterly returns	.03335	.03272	0.21	0.835
Quarters with outflows	4.9493	5.0128	-0.66	0.508

Note: This table reports balance test for matching procedure. The sample consists of 257 funds, 58 of which are in the treated group.

Table 2.3: Fund Repurposing and Fund Ratings: Diff-in-Diff

	(1) ESG rating (ind-adj)	(2) ESG rating (w.-av.)	(3) Environmental Fund Rating	(4) Social Fund Rating	(5) Governance Fund Rating
D_{it}	0.473***	0.238***	0.206**	0.268***	0.168*
	(0.091)	(0.060)	(0.077)	(0.068)	(0.066)
Percent in Equities	0.026***	0.025***	0.032***	0.022***	0.029***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Log Age (in quarters)	-0.004	0.000	-0.009	0.011	-0.013
	(0.028)	(0.023)	(0.031)	(0.023)	(0.028)
Log AUM	0.020	0.005	0.004	0.006	0.002
	(0.015)	(0.012)	(0.014)	(0.012)	(0.014)
N	9185	9185	9185	9185	9185
R^2	0.774	0.776	0.801	0.726	0.770
Y mean	3.588	3.346	3.935	3.129	3.740
Y sd	1.192	1.019	1.318	0.951	1.248

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the estimates of the treatment effect of fund repurposing on fund ratings according to equation 2.4. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. D_{it} is a dummy variable that equals one if the fund has already been repurposed, and zero otherwise. AUM is average assets under management over the last 12 months. All regressions include time and fund-type fixed effects. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Table 2.4: Fund Repurposing and Fund Performance: Diff-in-Diff

	(1) Fama-French Three-Factor Alpha	(2) Annual returns	(3) Quarterly returns	(4) Log of annual flows	(5) Log of quarterly flows	(6) Quarters with Negative Fund Flows	(7) Expense Ratio	(8) Profit
D_{it}	0.104	0.356	0.033	0.065	0.016	-0.018	-0.093	13.321
	(0.098)	(0.709)	(0.175)	(0.043)	(0.009)	(0.293)	(0.057)	(146.198)
Percent in Equities	0.002	0.016	0.003	-0.001	0.000	-0.000	-0.003	-1.896
	(0.003)	(0.023)	(0.005)	(0.001)	(0.000)	(0.007)	(0.002)	(3.976)
Log Age (in quarters)	-0.095*	0.181	0.081	-0.132***	-0.028***	1.634***	0.219***	101.476
	(0.040)	(0.231)	(0.069)	(0.018)	(0.003)	(0.119)	(0.030)	(69.985)
Log AUM	-0.006	0.112	-0.003	0.011	0.000	0.015	-0.077***	299.474***
	(0.021)	(0.095)	(0.022)	(0.006)	(0.001)	(0.058)	(0.016)	(47.537)
N	9185	9185	9185	8673	8964	9185	8123	8123
R^2	0.130	0.708	0.779	0.186	0.114	0.506	0.374	0.446
Y mean	-0.351	12.081	2.937	-0.095	-0.017	5.303	1.120	529.702
Y sd	2.444	17.592	8.968	0.342	0.085	2.646	0.491	1052.684

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the estimates of the treatment effect of fund repurposing on different outcomes according to equation 2.4. D_{it} is a dummy variable that equals one if the fund has already been repurposed, and zero otherwise. AUM is the average assets under management over the last 12 months. All regressions include time and fund-type fixed effects. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Table 2.5: Fund Repurposing and Fund Flows (six quarters leads)

	(1) F6.Log of annual flows	(2) F6.Log of quarterly flows	(3) F6.Quarters with Negative Fund Flows
D_{it}	0.151**	0.033**	-0.889*
	(0.057)	(0.011)	(0.443)
Percent in Equities	0.001	0.000	-0.012
	(0.001)	(0.000)	(0.007)
Log Age (in quarters)	-0.089***	-0.016***	1.147***
	(0.015)	(0.003)	(0.122)
Log AUM	-0.014	-0.003*	0.195***
	(0.007)	(0.001)	(0.057)
N	7458	7449	7518
R^2	0.161	0.098	0.376
Y mean	-0.093	-0.020	5.760
Y sd	0.324	0.082	2.387

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the estimates of the treatment effect of fund repurposing on six quarters leads of fund flows according to equation 2.4. D_{it} is a dummy variable that equals one if the fund has already been repurposed, and zero otherwise. AUM is average assets under management over the last 12 months. All regressions include time and fund-type fixed effects. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

CHAPTER 3

Whistleblowing and Fraud Deterrence

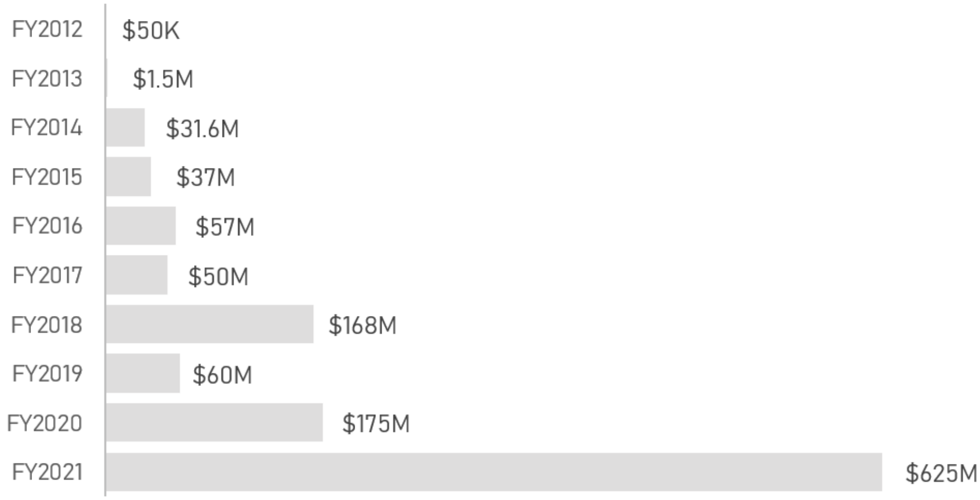
3.1 Introduction

Whistleblower employees play a pivotal role in uncovering corporate fraud. Dyck, Morse and Zingales (2010) show that employee-whistleblowers brought up more fraud cases than any other external actor such as the SEC, auditors or industry regulators in 1996-2004. Since that time, the introduction of new whistleblower programs, notably the Dodd-Frank Act, has complemented existing state and federal regulations. These programs not only protect whistleblowers from retaliation but also offer financial rewards based on the recovered funds. As a result, the frequency of whistleblower cases has surged in the past decade, with the monetary rewards issued by the SEC in 2021 surpassing the total sum awarded from 2012 to 2020 (Figure 3.1). This increase has sparked a policy debate in Europe regarding the adoption of similar regulations, alongside discussions about the potential costs and benefits of whistleblower programs.

Despite the significant costs associated with whistleblower programs in the US, relatively little is known about their effectiveness in exposing and deterring corporate fraud and misbehavior. The challenge lies in the inherent unobservability of committed fraud and the potential biases in measures based on discovered fraud, which depend on both violators' behavior and the disclosure incentives of internal monitors.

This chapter investigates the impact of whistleblower programs on employee disclosures and their side effects on internal audit and management disclosures. We also explore whether

Figure 3.1: SEC whistleblower rewards



Source: SEC Office of the Whistleblower. The value for FY2021 is computed from January 1st to November 29th, 2021.

a decrease in ex-post fraud detection could be accompanied by an increase in unidentified fraud. To address these questions, we develop a model that provides a framework for analyzing fraud disclosure within the context of whistleblower programs.

In our model, there are three risk-neutral agents who work in the same firm: the CEO, the Manager, and the Employee. The Manager oversees a project with an unknown payoff x, and he is the only one who always observes x. The Manager may attempt to divert a fixed portion α of the payoff, but this effort might be thwarted by internal monitoring. Then the Manager issues an internal report r, which equals either x or $(1-\alpha)x$, depending on whether fraud was committed. Next, the CEO receives a signal about the true cash flow x with some probability, which she can affect through costly effort. There are two types of CEOs: the high-type CEO has lower costs for internal audits that detect fraud and possesses better internal monitoring that prevents fraud.

If the CEO's signal coincides with the Manager's report (i.e. if there is no fraud), then the CEO has empty action space. If the CEO's signal differs from the Manager's report r, then the CEO decides whether to disclose fraud to the public. Disclosure launches the fraud investigation that returns stolen funds to the firm but may harm the CEO's reputation, i.e., the public belief that she is high-type. If the CEO remains silent and the Employee is informed, the Employee can blow the whistle and expose the fraud, gaining a fraction of recovered funds but suffering retaliation from the firm.

The model produces the following results. The whistleblower has a single-cutoff strategy of exposing fraud only for high project values. This is because retaliation costs are fixed, but his compensation is a fraction of recovered funds and, hence, of the project payoff. Then, we show that there are no separating equilibria in the game and that there exists a single-threshold pooling equilibrium where both CEO types disclose fraud only for high payoff values. Intuitively, fraud increases with the project value, as does the CEO's gain from exposing fraud, but reputation costs do not depend on project value. Note that the single-cutoff equilibrium is the standard result in the disclosure literature. The Manager implements a non-monotone fraud rule, stealing only from project payoffs that are low enough (when the probability of being caught is zero) or high enough (when the potential payoff is high).

We focus on the case where the CEO discloses more fraud than the whistleblower (i.e., the whistleblower's threshold is higher than that of the CEO). This is consistent with empirical evidence. Dyck et al. (2010) show that 34% of corporate fraud is reported by firm management (for example, by a press release or resignation) or the board of directors. Employees are responsible for a much lower share of around 18% of fraud disclosure.

The chapter has important empirical predictions for the impact of increased whistleblower compensation. We show that although the Employee is willing to expose more fraud, this also has side-effects on CEO disclosure and might lead to more fraud in the new equilibrium. The intuition for this result is as follows. First, higher compensation lowers whistleblower's disclosure threshold. Second, because the whistleblower's reward is paid from the recovered funds, the CEO has more to lose if if the Employee discloses the fraud instead of her. Thus, the CEO exerts more effort to discover fraud, especially if she is a low-type CEO. This, in turn, affects CEO reputation under different scenarios. As effort increases more for low-type CEO, the gap in the probability of being informed for the high-type and the -ow type CEOs shrinks, lowering the belief in high-type if the CEO is informed and discloses fraud. Thus, in the CEO disclosure scenario, her reputation declines, and staying silent becomes more attractive. Because of that, the CEO exposes less fraud, and her disclosure threshold increases. This expands the range of low project payoffs where no one discloses fraud, leading to more fraud in this region.

At the same time, fraud for high project payoffs declines due to higher exposure probability. The distribution of project values and whether high or low values are more probable determines whether total fraud declines or increases in the new equilibrium.

Most empirical papers studying the effect of whistleblower programs on fraud overlook the distinction between committed fraud and disclosed fraud, considering only disclosed fraud, which can be observed. Some papers show that whistleblower cases decline after increased compensation, which is consistent with our findings. The contribution of this chapter is to demonstrate that in some cases, this decline can occur despite an increase in committed fraud. This highlights limitations of current empirical studies and underscores the need to look beyond exposed fraud and find better proxies for committed fraud.

The equilibrium where the CEO discloses more fraud than the Employee occurs when whistleblower compensation is relatively low and retaliation costs are relatively high. This scenario is likely in countries with no or minor whistleblower programs. For instance, Nyreröd and Spagnolo (2021) discuss that whistleblower protection or rewards are mostly non-existent in Europe. Another example is industries with high retaliation costs, such as those with high employer power that greatly restrict future employment opportunities. Note that in the US, whistleblowers can receive a share of recovered funds only if they exceed \$1 or \$2 million, depending on the program. Although not modelled explicitly, this is consistent with our prediction that the whistleblower does not disclose low fraud due to insufficient reward. Therefore, fraud disclosure for low payoffs is solely determined by the CEO and is negatively affected by whistleblower compensation.

When policymakers anticipate this scenario, they should be cautious in administering whistleblower programs. One solution is to significantly increase whistleblower rewards to shift to another equilibrium where the Employee discloses more fraud than the CEO. In this new equilibrium, the disclosure of low fraud is undertaken by the whistleblower and increases with the reward, unambiguously leading to less fraud.

However, it might not be politically feasible to increase the whistleblower reward high enough to induce the new equilibrium. Another policy lever is measures aimed to reduce retaliation costs. This would increase the area of whistleblower disclosure but would not affect the CEO's effort or the amount of fraud if the change is relatively small. However, if retaliation costs are substantially reduced, the equilibrium changes. Then, if whistleblower compensation is increased at this point, it will lead to higher willingness to disclose and lower fraud without adverse side effects.

The rest of the chapter is organized as follows. Section 2 discusses the related literature and the background of whistleblower programs in the US. In Section 3, we describe the model and then provide the solution in Section 4. Section 5 reports the relevant comparative statics and discusses the intuition of the main results. Next, in Section 6, we discuss the limitations of the assumptions made throughout the chapter and consider potential extensions of the model. Section 7 briefly concludes.

3.2 Background and Related Literature

Background

The history of whistleblower programs in the US started in 1863 when Congress passed the False Claims Act which allows citizens who are not directly affiliated with the government to initiate legal actions against potentially fraudulent federal contractors. If successful, these lawsuits reward whistleblowers with 10% to 30% of any funds recovered by the government. Three recent Congressional acts substantially restructured whistleblower incentives.

The Sarbanes-Oxley Act of 2002 was passed following corporate scandals of the early 2000s (e.g., WorldCom or Enron cases). It determines certain protection of whistleblowers, for instance, it explicitly prohibits retaliation against employees of public companies who disclose questionable accounting or auditing matters. Second, the Tax Relief and Health Care Act of 2006 provides significant financial incentives to potential whistleblowers who inform the IRS about delinquent taxpayers: the whistleblowers may receive up to 30% of the proceeds exceeding $2 million. Finally, the Dodd-Frank Act extends protections of employee whistleblowers and specifies significant monetary incentives. The whistleblowers may receive from 10% to 30% of the funds that are recovered by the regulators based on the disclosure when the monetary sanctions exceed $1 million. According to the SEC press release, 233 individuals received more than $1.2 billion from the SEC for whistleblowing from 2012 to the end of November 2021.

The following empirical papers describe various aspects of whistleblowing that motivate our analysis. An influential study by Dyck et al. (2010) documents the effectiveness of different monitors in discovering fraud inside firms. They collect a sample of shareholder lawsuits related to potential accounting fraud from 1996 to 2004, and find that employee whistleblowers reveal more cases of financial wrongdoing than any other external monitor, including auditors and the SEC (almost a third of all cases for external actors). Call et al. (2018) use the data on employee whistleblower allegations and enforcement actions for financial misrepresentation. They document that the presence of an employee who blows the whistle is associated with higher monetary penalties and quicker enforcement proceedings. These findings, along with the SEC reports about the results of the whistleblower programs motivate our focus on the employee-whistleblower who is a major actor in fraud exposure.

Dyck et al. (2010) and Dey et al. (2021) document that employees who blew the whistle under the False Claim Act experienced severe negative consequences, such as lower future earnings, necessity to work in a different industry, demotion, harassment, threats and intimidation. This finding is reflected in our model as the retaliation costs for whistleblowing.

Related Literature

The first related strand of literature tries to empirically assess whether whistleblower programs are effective in deterring fraud. Cordis and Lambert (2017) analyze whether expanding whistleblower programs to private employees deters corporate fraud, using corporate fraud convictions in 2003-2015. The cross-section regression results indicate that higher awareness about whistleblower laws is associated with less corporate fraud cases.

Dey et al. (2021) study the effect of increased financial incentives for whistleblowers under the FCA cash-for-information programs. The authors find no impact on the propensity to blow the whistle internally (before disclosing to the outside monitor) or the number of lawsuits involving corporate fraud. However, their results indicate an increase in the quality of lawsuits: investigation length, the share of settled lawsuits, and total settlement amounts increase. This refutes the widespread concern that greater incentives to blow the whistle may lead to higher number of false or frivolous claims. Corroborated by this finding, we assume that the whistleblower can only make a truthful disclosure.

Wilde (2017) examines the effect of whistleblowing by employees of public US companies on the future firm conduct. The paper shows that firms subject to whistleblowing allegations exhibit significant decreases in financial misreporting and tax aggressiveness in subsequent years (compared to control firms). Note that this chapter focuses on reported fraud and its effect on the future firm engagement in fraud while we are interested in ex-ante effect of whistleblower programs.

All these papers suffer from a limitation of only looking at reported fraud which might not be a good proxy for committed fraud. A rare exception is the study by Berger and Lee (2021) which attempts to develop a measure of committed fraud by predicting fraud probability from firm financial and performance data (they use both F-score and M-score models). The authors investigate whether the Dodd-Frank whistleblower law increased deterrence of accounting fraud. Their quasi-experimental design utilizes the state False Claim Acts (FCA) passed before the Dodd-Frank. The state FCAs protect and reward whistleblowing about fraud at a public firm if a state's pension fund is invested in this firm – which is interpreted as defrauding the state government (Rapp, 2007). Using firms exposed to a state FCA as a control group, the authors show that exposure to the Dodd-Frank reduces probability of accounting fraud by 12-22%.

Our contribution to this empirical literature is to provide a theoretical framework explaining the mechanisms of how whistleblower compensation affects fraud not only through the employee disclosure but also through the management disclosure. Moreover, we show that because of this additional channel, committed fraud can increase even when reported fraud declines, which highlights the importance of developing measures of committed fraud in the vein of Berger and Lee (2022).

The second strand of literature that we relate to is theoretical papers exploring the optimal design of the whistleblower policies and their potential side effects.

Givati (2016) develops a model studying the optimal size of the whistleblower reward given that it might encourage false reports. There are two players: the employer and the employee. The employer decides whether to violate the law, and the employee decides whether to report this violation. Importantly, the employee can make false claims that has a lower probability of success than a true claim. If the report is successful, then the government punishes the employer and rewards the whistleblower. Interestingly, the optimal reward from the social planner's perspective is a non-monotone function of the probability of false claim success: it increases for low probabilities, but once the false claim success becomes high enough, the optimal whistleblower reward is zero. In our analysis, we assume away false claims (which importance is not supported by empirical evidence) and focus on interaction with the CEO disclosure.

Heyes and Kapur (2009) analyze how responsive regulators should be whistleblower reports in terms of (i) probability of investigation, and (ii) punishment for firms if they violated the law. Importantly, the authors assume that there are no monetary rewards for whistleblowers but they may have different intrinsic motivations. Three competing theories are: (a) the decision to blow the whistle is based on "moral defensibility" of the violation (e.g., if cost of compliance is very high then the employee will not report the violation); (b) report the violation if the social benefit from doing so is sufficiently high; (c) disgruntled employees who want to punish the firm as much as possible. The main insight of the model is that the design of whistleblower programs should carefully take into account the motivation of the employees to blow the whistle: for example, for (b) the optimal policy is to always investigate all whistleblower tips and impose maximum possible punishment for the firms, however, the same is not optimal for punishment-motivated employees (case (c)). We complement their analysis by studying another type of whistleblower who is motivated by the monetary reward.

The following papers examine potential side effects of whistleblower programs. For example, Ting (2008) develops a model showing that in some circumstances whistleblower protections may undermine the overall output by diluting the employee's incentives to incur effort. The two-period model has a three-tier principal agent structure: there is the principal, the manager, and the employee. The manager approves or rejects a project which may be of high or low quality which is determined by employee's effort. There is agency conflict: the manager is "aggressive" and wants to approve any project while the principal and the employee only want the high-quality project approved. The manager wishes to provide incentives for the employee to incur high effort. However, when the employee is able to blow the whistle, the manager also wants to avoid whistleblowing. Moreover, the employee who can make a disclosure about the project quality does not face all the consequences of the bad project, because the principal will replace an exposed manager in the second period. That is, both effects lower the employee's incentives to work hard and, hence, make the high-quality project less likely. In this case, the principal may want to ban whistleblowing altogether. This setting is more appropriate for analyzing internal whistleblowing, while our main interest is external whistleblowing.

Another side effect is analyzed in Iwasaki (2018) that explores whether increased rewards for external whistleblowing on corporate crimes might discourage internal reporting and hence undermine internal governance systems.[1] To define a socially optimal level of whistleblower reward, we need to consider the trade-off between external whistleblowing as a means of fraud deterrence and internal whistleblowing as a means of fraud prevention. The main result of the chapter is that the probability of internal reporting is a non-monotone function of rewards for external whistleblowing: as the reward increases, the probability of internal reporting first increases but drops to zero after a certain threshold.

In addition, some authors informally discuss that financial incentives for whistleblowers may potentially exacerbate corporate fraud. For example, Howse and Daniels (1995) describe the following dynamic issue. If the whistleblower has information about fraud in the firm, he may decide to withhold his signal and wait till the wrongdoing gets even worse and only then disclose the information to the regulators. Disclosure of higher fraud might potentially increase the whistleblower's payoff but also lead to more violations. This concern is partially addressed by the extensive report issued by the Government Accountability Project. The report is based on online survey of 1,366 whistleblowers, whistleblower right lawyers and whistleblower organizations throughout the world and indicates that timing of disclosure is mostly driven by time required to obtain evidence about potential wrongdoing, time and effort to find an experienced lawyer, importantly, the personal situation of potential whistleblowers. Moreover, according to 2011 National Business Ethics Survey, 97% of whistleblowers first file their complaints via internal compliance, so the concern that some whistleblowers may strategically withhold their information should be minimal. This high share might also indicate that crowding-out of internal disclosure is a minor issue.

Thus, this chapter abstracts from internal monitoring and solely focuses on external reporting. Our contribution is to analyze external reporting not only by the employee but also by the firm management – voluntary disclosure of firms; and to study behaviour of potentially fraudulent managers in the presence of whistleblowing risk and reputation concerns.

[1] Recall, however, that Dey et al. (2021) investigate this question empirically and do not find any effect of increased whistleblower compensation on internal reporting.

In addition, we complement the ongoing discussion of potential side effects of whistleblower programs by showing that higher whistleblower reward can crowd-out the CEO disclosure and lead to more fraud.

Finally, we contribute to the literature on voluntary disclosure which goes back to Grossman (1981) Milgrom (1981), Dye (1985), and Jung and Kwon (1988). The following papers consider voluntary disclosure with informed third-party in similar settings. Dye (2017) studies a decision of potentially informed seller on whether to disclose or withhold her private signal about the value of the product. There is also a non-strategic third-party (e.g., the regulator) who, if informed, immediately discloses the information. If it turns out that the buyer overpaid for the product, then the seller must make damage payments to the buyer. In equilibrium, the seller discloses high signals and withholds low signals. Frenkel et al. (2020) consider a model where the firm chooses its optimal disclosure strategy in the presence of a non-strategic analyst who has the same information as the manager. In equilibrium, the firm uses a single-cutoff disclosure rule. Importantly, depending on the information production function of the analyst, greater analyst coverage may crowd out or crowd in the firm's voluntary disclosure. Our contribution is to explore voluntary disclosure when the third-party is strategic and apply it to a policy-relevant setting of whistleblowing.

Another related paper is Langberg and Sivaramakrishnan (2008). They explore voluntary disclosure in the presence of the analyst who may have private information about the firm. They consider a special case of strategic analyst whose objective is to maximize the accuracy of his announcements at the lowest possible cost. Our modeling of whistleblower's incentives is different since there are retaliation costs and rewards from regulators. This is a more realistic assumption in the context of whistleblowing. Another difference is that Langberg and Sivaramakrishnan (2008) assume that the analyst's information is orthogonal to the information of the firm, while we in our setting, the CEO and the employee receive the same signal.

There are also continuous-time models which consider a dynamic version of Dye (1985).

For example, Acharya et al. (2011) develop a model explaining the disclosure clustering in bad states of the world, since negative public news provide more incentives for the Manager to disclose the firm private information. Other examples of dynamic disclosure models include Guttman, Kremer and Skrzypacz (2014) and Marinovic and Varas (2016). While this literature focuses on the timing of disclosure, our goal is to study the interaction of two strategic informed parties.

Our contribution to this literature is to study the interaction of voluntary disclosure of two strategic players: the CEO and the employee. We show that higher disclosure incentives for the employee lead to higher CEO effort to acquire information but, at the same time, to lower region of CEO disclosure.

3.3 Model

Consider a firm with three risk-neutral players – the CEO, the Manager, and the Employee. The Manager is a relatively high-level employee, but lower in the hierarchy than the CEO, for example, a head of the firm's division or department. The Employee is a mid-level specialist who works under the supervision of the Manager and therefore might have access to the information about the Manager's actions.

The firm has a project with unknown payoff that is supervised by the Manager. Because of that he has a superior information about the project payoff compared to the CEO and the Employee.

The game starts at $t = 0$ when the project's payoff x is drawn according to the distribution $x \sim F[0, \infty]$. At the same time, the CEO is drawn to be high-type (H) with probability $\mu \in (0, 1)$, and low-type (L) otherwise. The CEO's type is her private information but the prior probability μ of the high-type CEO is public knowledge. The CEO type affects the efficiency of preventing fraud in the firm as well as the CEO's cost of learning about fraud.

The CEO does not observe x. At $t = 1$, she can exert costly effort $e \in [0, 1]$ in order to receive a signal later in the game with probability pe where $p \in (0, 1]$. One can think about p as the highest possible probability to receive a signal, i.e., the limit to obtain information. The signal structure is described below at $t = 3$. The cost of effort is $c(e) = \frac{1}{2}k_\tau e^2$, where $\tau \in \{L, H\}$, and $0 \leq k_H < k_L \leq \infty$. This implies that the high-type CEO has lower cost of learning about the true value of the project and, as discussed below, whether the Manager committed fraud. For simplicity, we assume that $k_H = 0$, that is, the high-type CEO always chooses $e_H = 1$.

At $t = 2$, the Manager observes the payoff x and decides whether to attempt to divert α percent of the cash flow or not. The parameter $\alpha \in (0, 1)$ is assumed to be fixed, that is, the Manager only decides whether to divert funds or not but cannot affect how many percents of the project payoff to steal. If the Manager attempts to steal from the firm, he is successful with probability w_τ, where $\tau \in \{L, H\}$ and $0 \leq w_H \leq w_L$. In practice, the Manager might not be able to steal from the firm because of high quality of internal monitoring that prevents fraud when it is attempted. Moreover, the high-type CEO is better at managing fraud prevention, i.e., under $\tau = H$ attempted fraud is successful with lower probability.

Irrespective of the fraud decision, the Manager makes an internal disclosure to the CEO and the employee about value of the project for the firm (after-fraud). In other words, if the Manager tries to steal and is successful (probability w_τ), then he discloses $(1-\alpha)x$; if not, he announces the true payoff x. Because only the Manager knows the true value of the project, neither the CEO nor the employee know whether fraud happened at this stage. Note that the public does not observe the Manager's report. In practice, this corresponds to internal report or presentation on the state of the project or submission of ongoing accounting documents that are not yet aggregated and released to the public. Alternatively, if the firm is private, it does not have to disclose internal information to the public. We discuss this and other assumptions in more detail in Section 6.

At $t = 3$, the CEO receives the signal $s \in \{\emptyset, (Fraud, x)\}$. $Fraud \in \{0, 1\}$ is a binary part of the signal, and $Fraud = 1$ means that the Manager diverted funds from the firm at $t = 2$ while $Fraud = 0$ means that there is no fraud in the firm (that is, the Manager's report at $t = 2$ is truthful). The second component of the non-empty signal s, i.e. x, is the true payoff of the project. Given that the CEO exerted effort e at $t = 1$, she receives an informative signal with probability pe, i.e., $P(s = (F, x)) = pe$. In other words, if the Manager engaged in fraud at $t = 2$, then with probability pe the CEO will receive an informative signal $s = (Fraud = 1, x)$, and with probability $1 - pe$ she will receive an empty signal, i.e., $s = \emptyset$. However, if the Manager does not engage in fraud at $t = 2$ then with probability pe the CEO will receive an informative signal $s = (Fraud = 0, x)$ and with probability $1 - pe$ she will receive an empty signal, i.e., $s = \emptyset$.

Then, if informed, the CEO decides whether to disclose her signal to the public or not. The binary component $Fraud$ of the signal is always verifiable, while the second component x is verifiable only for $Fraud = 1$. Put it differently, the informed CEO might choose to disclose her signal if $Fraud = 1$. On the contrary, if the CEO is informed and $Fraud = 0$ then she cannot make a disclosure. This can be viewed as a consequence of fraud investigation that is only launched in case of disclosed fraud and that verifies both the fraud amount and the true project value. If there is fraud and the disclosure is made, then the stolen money is recovered from the Manager and returned to the firm. However, it is possible that the informed CEO decides not to disclose fraud because she cares not only about the project payoff but also about her reputation of being the high-type CEO. The CEO's utility increases in the belief of the public on whether she is the high-type who is better at fraud prevention and internal monitoring, conditional on all available public information.

Next, at $t = 4$, the Employee receives the signal $s \in \{\emptyset, (Fraud, x)\}$, where $P(s = (Fraud, x)) = q$. If the Employee is informed and the CEO did not make a disclosure, he decides whether to disclose the fraud to the public. As with the CEO's signal, the Employee decides whether to disclose his signal only if his signal $s = (Fraud = 1, x)$ and CEO did not disclose, i.e., if the Manager engaged in Fraud, the CEO remained silent and the Employee is informed. We assume that $q < p$, i.e, if the CEO exerts highest possible effort $e = 1$ then she will be more likely to be informed than the Employee.

If the Manager's wrongdoing is exposed to the public, then the Manager incurs costs C and the stolen funds are taken away from him. Costs C can be broadly viewed as any costs associated with committing fraud such as litigation costs, regulatory costs, costs of potentially adverse career prospects, etc. If the disclosure was made by the Employee, then (i) he is entitled by law to receive a fraction β of the recovered funds while the rest is returned to the firm, (ii) the CEO incurs a regulatory penalty $L \geq 0$. If this is the CEO who voluntarily disclosed fraud, then all the money goes back to the firm. I assume that there is no regulatory penalty if the disclosure is made by the CEO.

If neither the CEO nor the Employee make an announcement, then the public calculate reputation of the CEO conditional only on a non-disclosure event. That is, we assume that the public has no access to the announcement of the Manager and therefore observe nothing if both the CEO and Employee choose to remain silent.[2] However, if the CEO or Employee disclose then the public observes their disclosure and update their beliefs about the CEO's type taking into account both the fact of disclosure and the exact value of disclosed x.

Payoffs of the players are defined as follows:

The informed **Employee**: $\begin{cases} 0, \ if \ withholds \ information \ (both \ for \ Fraud \ = \ 0 \ or \ 1) \\ \beta \alpha x - r, \ if \ discloses \ fraud \end{cases}$

The **Manager**: $\begin{cases} 0, \ if \ he \ does \ not \ engage \ in \ fraud \\ \alpha x, \ if \ he \ commits \ fraud \ but \ \textbf{the fraud is not disclosed} \\ -C, \ if \ he \ commits \ fraud \ and \ \textbf{it is disclosed} \end{cases}$

The **informed CEO** when she observes that fraud has been committed (in other cases the

[2]See Section 6 for the discussion of this assumption.

CEO has empty action space since we assumed that x is verifiable only if $Fraud = 1$):

$$\begin{cases} (1-\alpha)x + P(\tau = H|no\ disclosure\ neither\ by\ E\ nor\ by\ CEO), if\ both\ CEO\ and\ E\ do\ not\ disclose \\ (1-\beta\alpha)x - P + P(\tau = H|\ Employee\ discloses\ x),\ if\ CEO\ withholds\ and\ E\ discloses \\ x + P(\tau = H|\ CEO\ discloses\ x),\ if\ CEO\ discloses \end{cases}$$

Finally, at $t = 5$, the payoffs are realized.

3.4 Model Solution

The solution concept for the game is Perfect Bayesian Nash Equilibrium (PBNE) in pure strategies. We will start solving the model by backward induction.

Strategy of the Employee/Whistleblower

The optimal strategy of the Employee depends on the value of signal he has received, the whistleblower reward parameter β, the retaliation costs r and the fraction of the payoff which could be diverted by the Manager α. The whistleblower strategy is described in Proposition (1).

Proposition 1. *Assume that the Manager has engaged in fraud. Then, the Employee makes a disclosure if and only if the following conditions are satisfied:*

1. *The CEO did not make a disclosure.*
2. *The Employee has received the signal about true project's payoff x*
3. *The true payoff of the project x is higher than the announcement of the Manager (i.e., fraud has been committed)*
4. *The true project's payoff x is high enough: $x > \frac{r}{\beta\alpha} := \sigma_w$*

Proof. The first three conditions are ensured by the model assumptions: (i) if the CEO has disclosed the true state of the world then there is nothing to disclose; (ii) the signal is verifiable and, hence, an uninformed player cannot make a disclosure; (iii) no disclosure is possible if there is no fraud. The Employee gets $\beta\alpha x - r$ if he discloses fraud, and 0 if he withholds his signal. Then, given the first three conditions, the Employee exposes fraud if $x > \frac{r}{\beta\alpha}$. □

The whistleblower only discloses fraud for high payoff values. Intuitively, retaliation costs are fixed but, as project value increases, fraud and the fraction of stolen funds that is awarded to the Employee both rise as well, which makes blowing the whistle more attractive. Note that this results depends on the implicit assumption that there are no commitment types of employees who would disclose fraud irrespective of their personal gain or loss.

Strategy of the CEO

Proposition 2. *There are no separating equilibria in this game. That is, in every equilibrium both types of the CEO choose to disclose (or withhold) the same signals.*

Proof. Assume that there is a separating equilibrium in this game and there exists a value of the project payoff x such that, conditional on fraud and being informed, $\tau = H$ prefers to make a disclosure while $\tau = L$ remains silent. Note that conditional on fraud and being informed, the payoff functions of both types of CEOs are identical. Hence, if $\tau = H$ wants to make a disclosure conditional on observing x and fraud happening, this is also true for the other type. Hence, there is no such x, and the set of project values where, conditional on fraud and being informed, $\tau = H$ discloses fraud is a subset of the set of project values where $\tau = L$ discloses fraud. Similarly, one can show that the opposite is true. Hence, both sets coincide which contradicts the initial assumption on separating equilibrium. This proves that there is no separating equilibrium in this game. □

We do not analyze all pooling equilibria of the game and only focus on single-threshold equilibria. Recall that a single-threhold equilibrium is the standard outcome in the disclosure literature (e.g., Dye (1985, 2017) or Frenkel et al. (2020)). Specifically, we assume that there is a threshold σ_c such that, conditional on fraud and being informed, both types of the CEO disclose their signal (F, x) if and only if $x \geq \sigma_c$. We focus on the case where $\sigma_c < \sigma_w$, and the CEO is willing to disclose fraud for a higher range of payoffs than the employee. We briefly discuss the case of $\sigma_c \geq \sigma_w$ at the end of the section.

Case 1: $\boldsymbol{\sigma_c < \sigma_w}$

For brevity, we introduce the following notations for CEO's reputation under different scenarios:

- $R_1 = P(\tau = H | \textit{no disclosure neither by E nor by CEO}) = P(\tau = H|ND)$
- $R_2(x) = P(\tau = H| \textit{ CEO discloses } x) = P(\tau = H|D_c)$
- $R_3(x) = P(\tau = H| \textit{ Employee discloses } x) = P(\tau = H|D_w)$

Then we can rewrite the CEO's interim payoff V_c (conditional on fraud and being informed) in the following form:

$$V_c(x) = \begin{cases} (1-\alpha)x + R_1, \; if \; no \; disclosure \\ x + R_2(x), \; if \; CEO \; discloses \; x \\ (1-\beta\alpha)x - P + R_3(x), \; if \; CEO \; withholds \; and \; E \; discloses \; x \end{cases}$$

Lemma 1. *Equilibrium beliefs of the public about the CEO's type do not depend on the project value disclosed by the CEO or the employee. That is, conditional on disclosure of x by either player, $R_2(x)$ and $R_3(x)$ do not depend on x, and are equal to the following:*

a) $R_2(x) = \frac{\mu}{\mu + (1-\mu)\frac{e w_L}{w_H}} =: R_2$ *for* $x \geq \sigma_c$

b) $R_3(x) = \frac{\mu}{\mu+(1-\mu)\frac{1-pe}{1-p}\frac{w_L}{w_H}} =: R_3$ *for* $x \geq \sigma_w$

Proof. Follows immediately from Bayes' rule. $\square$

Lemma 1 allows us to treat $R_2(x)$ as a constant in each region where disclosure decision of the CEO is the same. First, $R_2(x) = R_2$ when CEO discloses fraud and $x \geq \sigma_c$. Second, off-equilibrium reputation of the CEO, i.e. when the realized payoff x of the project is lower than the disclosure threshold σ_c and the CEO discloses such x, are assumed to be equal to equilibrium beliefs R_2.

Similarly, we treat as constant the CEO reputation $R_3(x) = R_3$ when disclosure is made by the whistleblower and $x \geq \sigma_w$. Again, we assume that off-equilibrium beliefs of the public are the same as equilibrium beliefs, i.e. if the whistleblower discloses x lower than σ_w (which never happens in equilibrium), then $R_3(x) = R_3$.

Hence, $R_2(x) = R_2$ and $R_3(x) = R_3$ for all realizations of project value.

Strategy of the Manager

Recall that the Manager decides whether to engage in fund diversion or not after observing the realized project value x. Naturally, in the model his decision depends on the particular value of x since the probability of being caught (and then punished) depends on project value: both the CEO and the employee are going to disclose fraud if they are informed and the project value x is high enough (i.e., if $x > max(\sigma_c, \sigma_w) = \sigma_w$ since we are considering Case 1). The next proposition describes the optimal strategy of the Manager.

Proposition 3. *The Manager engages in fraud in the following cases:*

a) *if* $x < \sigma_c$*;*

b) *if* $\sigma_c \leq x < \sigma_w$ *and* $x > t_1(e)$*, where* $t_1(e) := \frac{C[\mu p+(1-\mu)pe]}{(1-\mu p-(1-\mu)pe)\alpha}$*;*

c) *if* $x > \sigma_w$ *and* $x > t_2(e)$*, where* $t_2(e) := \frac{Cy(e)}{\alpha(1-y(e))}$ *and* $y(e) := \mu(p + (1-p)q) + (1-\mu)(pe + (1-pe)q)$*.*

Proof. See the Proofs. □

Intuitively, the Manager always engages in fraud for low project values because there is no chance that the fraud will be revealed. Otherwise, he considers benefits of fraud (a share of project payoff), fixed punishment and a probability of getting caught. Note that even very strict punishment cannot prevent fraud for high payoffs – for arbitrary high C there exists x s.t. $x > \max\left(\sigma_w, \frac{Cy(e)}{\alpha(1-y(e))}\right)$, and, hence, the Manager will always engage in fraud for some high values of x.

Note that for any $e \in [0, 1]$, $t_1(e) < t_2(e)$ and both $t_1(e)$ and $t_2(e)$ are increasing in effort e. To simplify the analysis, we make assumptions on parameter values so that $t_1(e) > \sigma_w$. Then, the Manager engages in fraud only if $x < \sigma_c$ or $x > t_2(e)$, i.e., we rule out the possibility of fraud in the intermediate region between the thresholds σ_c and σ_w.

Reputation of the CEO in the Case of Non-disclosure

Knowing the strategies of the Manager, the employee and disclosure rule of the CEO, we can calculate the CEO's reputation R_1 when neither the CEO nor the Employee disclose their signals (ND):

$$R_1 = P(\tau = H|ND) = P(ND|\tau = H)\frac{P(\tau = H)}{P(ND)}$$

$$P(ND|\tau = H) = F(\sigma_c) + [F(t_2(e) - F(\sigma_c)] + (1 - F(t_2(e))[(1 - w_H) + w_H(1 - p)(1 - q)]$$

$$P(ND|\tau = L) = F(\sigma_c) + [F(t_2(e) - F(\sigma_c)] + (1 - F(t_2(e))[(1 - w_L) + w_L(1 - pe)(1 - q)]$$

$$=> R_1 = \frac{\mu}{\mu + (1 - \mu)\frac{P(ND|\tau=L)}{P(ND|\tau=H)}}. \tag{3.1}$$

Effort Problem of the CEO

The high-type CEO has zero cost of effort and, thus, she always chooses the highest effort level $e_H = 1$.[3] The low-type CEO, $\tau = L$, chooses effort $\hat{e}$ to maximize her expected payoff from the game.

Recall that we consider the case where $\sigma_c < \sigma_w$, $t_1(e) > \sigma_w$ and $t_2(e) > \sigma_w$. The ex-ante (i.e. before receiving the signal) objective function of the low-type CEO is the following:

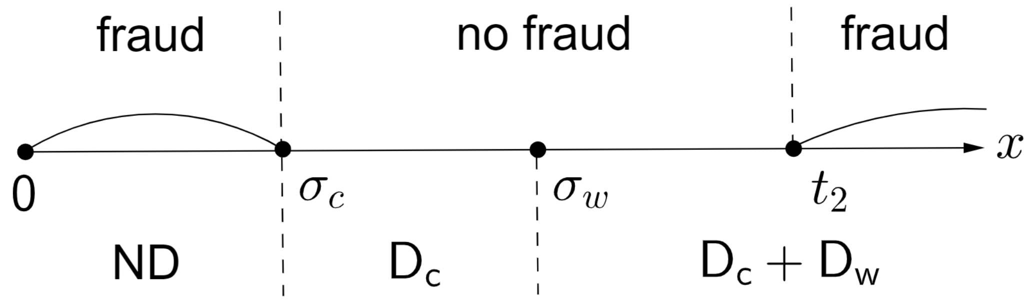
$$
\begin{aligned}
V_c(\hat{e}) = p\hat{e}\Bigg[\int_0^{\sigma_c} \{w_L(1-\alpha)z + R_1) + (1-w_L)(z+R_1)\}\, dF(z) + \int_{\sigma_c}^{t_2(e)} (z+R_1) dF(z) + \\
+ \int_{t_2(e)}^{\infty} \{w_L(z+R_2) + (1-w_L)(z+R_1)\}\, dF(z)\Bigg] + \\
+ (1-p\hat{e})\Bigg[\int_0^{\sigma_c} \{w_L((1-\alpha)z + R_1) + (1-w_L)(z+R_1)\}\, dF(z) + \int_{\sigma_c}^{t_2(e)} (z+R_1) dF(z) + \\
+ \int_{t_2(e)}^{\infty} [w_L\{q(z(1-\alpha\beta)+R_3) + (1-q)((1-\alpha)z+R_1)\} + (1-w_L)(z+R_1)]\, dF(z)\Bigg] - \frac{k_L}{2}\hat{e}^2 = \\
= \int_0^{\sigma_c} \{w_L(1-\alpha)z + R_1) + (1-w_L)(z+R_1)\}\, dF(z) + \int_{\sigma_c}^{t_2(e)} (z+R_1) dF(z) + \\
+ p\hat{e}\int_{t_2(e)}^{\infty} \{w_L(z+R_2) + (1-w_L)(z+R_1)\}\, dF(z) + (1-p\hat{e})\int_{t_2(e)}^{\infty} [w_L\{q(z(1-)z+R_1)\}] - \frac{k_L}{2}\hat{e}^2
\end{aligned}
$$

Note that the CEO's payoff depends not only on her effort $\hat{e}$ but also other's beliefs about effort of the CEO e. However, since the CEO's choice of effort is unobserved by other players, we can treat other's beliefs e as constant and maximize $V_c(\hat{e})$ in $\hat{e}$. The first-order condition is the following:

$$
\begin{aligned}
\frac{\partial V_c(\hat{e})}{\partial \hat{e}} = p\int_{t_2(e)}^{\infty} \{w_L(z+R_2) + (1-w_L)(z+R_1)\}\, dF(z) - \\
-p\int_{t_2(e)}^{\infty} [w_L\{q(z(1-\alpha\beta)+R_3) + (1-q)((1-\alpha)z+R_1)\} + (1-w_L)(z+R_1)]\, dF(z) - k_L\hat{e} = 0
\end{aligned}
$$

[3] All previous derivations implicitly use $e_H = 1$.

Figure 3.2: Equilibrium Fraud and Disclosure, Case 1



Note: This graph illustrates the equilibrium described in Theorem 1: the Manager engages in fraud only when the project's payoff x is either lower than σ_c or higher than t_2. The CEO discloses fraud if and only if x is greater than her threshold σ_c, conditional on fraud and being informed. The whistleblower discloses fraud if and only if x is not less than his threshold σ_w, conditional on fraud, being informed, and no disclosure by the CEO.

We can rewrite the FOC as follows:

$$\hat{e} = \frac{pw_L}{k_L} \int_{t_2(e)}^{\infty} \{z\alpha(1 - q + q\beta) + [R_2 - qR_3 - (1 - q)R_1]\} \, dF(z), \tag{3.2}$$

where reputation values R_1, R_2, and R_3 are functions of beliefs e.

The equilibrium level of effort e^* is s.t. the expression above holds when $e^* = e = \hat{e}$. In the Proofs we introduce conditions Σ and Ω. As Lemma 2 shows, Σ ensures the existence and uniqueness of such $e^* \in [0, 1]$.

Lemma 2. *If the conditions Σ are satisfied, then there exists unique e^* s.t. (i) Equation (3.2) holds for $e = e^*$ and $e = \hat{e}$; (ii) $e^* \in [0, 1]$.*

Proof. See the Proofs. □

Equilibrium

Theorem 1. *If the condition Ω is satisfied, then the game has the pooling equilibrium where:*

1. *The informed Employee discloses his signal if and only if $Fraud = 1$ and $x \geq \sigma_w^* = \frac{r}{\beta\alpha}$.*

2. *The informed CEO discloses his signal if and only if $Fraud = 1$ and $x \geq \sigma_c^* = \frac{R_1(e^*)-R_2(e^*)}{\alpha}$ where R_1 and R_2 are defined in Equation 3.1 and Lemma 1, respectively, and $\sigma_c^* < \sigma_w^*$.*

3. *The Manager engages in fraud if and only if $x \in [0, \sigma_c) \cup (t_2(e^*), \infty)$.*

4. *The equilibrium effort level e^* of the low-type CEO is defined in Equation (3.2). The equilibrium effort of the high-type CEO is $e_H = 1$.*

5. *The off-equilibrium beliefs coincide with equilibrium beliefs.*

Proof. See the Proofs. □

Theorem 1 describes the equilibrium which is interesting because: (i) the Manager's strategy is non-monotone: he engages in fraud only low or high enough project payoffs while for does not misbehave in the intermediate range; (ii) as it will be shown in the next section, the CEO's threshold σ_c increases with the whistleblower compensation – hence, if one increases β it increases the amount of fraud for the lower payoff values.

Case 2: $\sigma_c \geq \sigma_w$

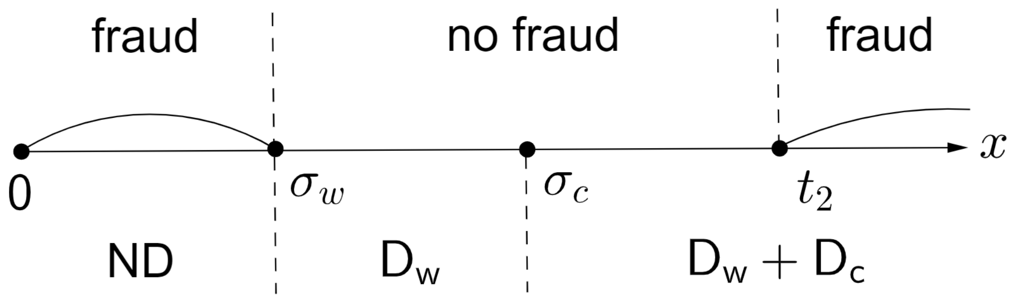
Another case is when the whistleblower discloses more fraud than the CEO: the whistleblower's threshold σ_w is lower than the CEO threshold σ_c. In the interest of space, we only describe the equilibrium in this case. The proofs of the results are similar to Case 1 and are available by request.

Theorem 2. *If the condition $\hat{\Omega}$ is satisfied, then the game has the following pooling equilibrium:*

1. *The informed Employee discloses his signal if and only if $Fraud = 1$ and $x \geq \sigma_w^* = \frac{r}{\beta\alpha}$*

2. *The informed CEO discloses his signal if and only if $Fraud = 1$ and $x \geq \sigma_c^* = \frac{(1-q)R_1(e^*)+qR_3(e^*)-R_2(e^*)}{\alpha(1-q+q\beta)}$*

Figure 3.3: Equilibrium Fraud and Disclosure, Case 2



Note: This graph illustrates the equilibrium described in Theorem 2: the Manager engages in fraud only when the project's payoff x is either lower than σ_w or higher than t_2. The CEO discloses fraud if and only if x is greater than her threshold σ_c, conditional on fraud and being informed. The whistleblower discloses fraud if and only if x is not less than his threshold σ_w, conditional on fraud, being informed, and no disclosure by the CEO.

3. *The Manager engages in fraud if and only if* $x \in [0, \sigma_w) \cup (t_2(e), \infty)$.

4. *The equilibrium effort level* e^* *of the low-type CEO satisfies the following equation:*

$$\int_{t_2(e^*)}^{\infty} z+R_2(e^*)dF(z)+\int_{t_2(e^*)}^{\infty} q(z(1-\alpha\beta)+R_3(e^*)))+(1-q)((1-\alpha)z+R_1(e^*))dF(z) = \frac{k_L e^*}{p w_L}$$

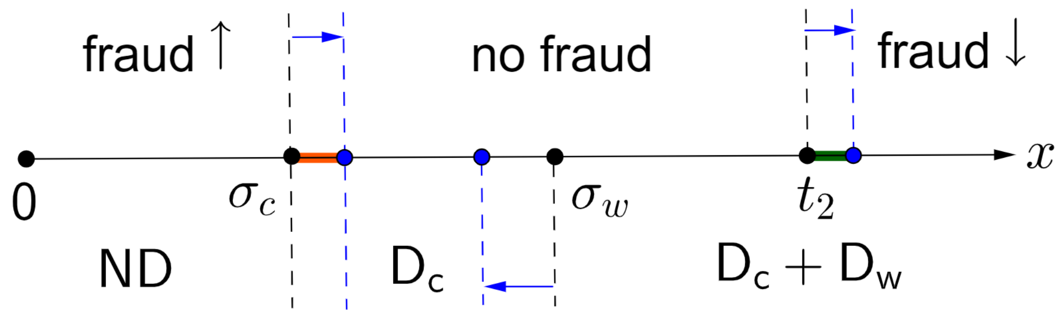
5. *The off-equilibrium beliefs coincide with the equilibrium beliefs.*

Note that reputation values $R_i(e)$ depend on the type of equilibrium we are considering. As we discuss below, although Theorem 1 and 2 describe very similar equilibria, there is an important difference between these two cases: the effect of increasing whistleblower's compensation on the Manager's decision to engage in fraud.

3.5 Comparative statics

Consider the policy intervention of increasing the whistleblower compensation by raising the share of recovered funds β that are awarded to the Employee.

Figure 3.4: Equilibrium change due to increase in β, Case 1



Note: This graph illustrates the comparative statics described in Proposition 4. If the whistleblower's compensation paramater β increases then: (i) the corresponding threshold σ_w (mechanically) decreases; (ii) the equilibrium effort e^* increases => the CEO is informed with higher probability, hence, the Manager's upper threshold also increases (the green segment) as he needs to have higher potential rewards to justify increased risk of being disclosed; (iii) the CEO's threshold increases (the red segment), that is, the Manager engages in fraud more for the lower values of the project value. Does the ex-ante probability of fraud decrease when β goes up? This depends on the distribution of the project value: if the probability of the red segment is greater than the probability of the green segment then the ex-ante likelihood of fraud increases; if not, then the increased compensation of the whistleblower indeed deters fraud.

Proposition 4. *If the share of recovered funds β that are awarded to the Employee increases, the equilibrium described in Theorem 1 changes in the following way (as long as the type of equilibrium does not change and $\sigma_c < \sigma_w$):*

1. *The Employee disclosure threshold $\sigma_w^* = \frac{r}{\beta\alpha}$ declines.*

2. *The CEO disclosure threshold $\sigma_c^* = \frac{R_1(e^*)-R_2(e^*)}{\alpha}$ increases.*

3. *The equilibrium effort level e^* of the low-type CEO increases.*

4. *The Manager's fraud threshold $t_2(e^*)$ increases. Thus, the first fraud region $x \in [0, \sigma_c)$ expands and the second fraud region $x \in (t_2, \infty)$ shrinks.*

Proof. See the Proofs. □

The rest of this section discusses the intuition for these changes and relevant policy implications.

First, recall that the Employee has a single-cutoff strategy: he discloses fraud only for high payoff values, $x > \sigma_w = \frac{r}{\beta\alpha}$. If whistleblower compensation β increases, the Employee disclosure threshold declines and the range of disclosed project payoffs increases.

Second, the optimal CEO's effort to acquire information increases. Note that the whistleblower compensation is paid by the firm. Thus, in case of whistleblower disclosure, the share of the project value that is awarded to the Employee and taken from the CEO increases with the rise in whistleblower reward. In addition, the Employee exposes higher range of fraud values. As a result of both higher frequency and higher cost of whistleblowing, it becomes costlier for the CEO to be uninformed, inducing more effort to learn about fraud.

Third, the CEO disclosure threshold $\sigma_c^* = \frac{1}{\alpha}(R_1(e^*) - R_2(e^*))$ increases. Note that σ_c does not directly depend on the whistleblower's compensation because in the equilibrium there is no Employee disclosure neither right above nor right below the threshold. However, a rise in the whistleblower's compensation increases CEO's effort and changes the CEO's reputation under different disclosure scenarios.

The CEO's reputation if she discloses fraud declines in effort (i.e., $\frac{\partial R_2(e)}{\partial e} < 0$). Recall that the high-type's effort is always 1 and only the low-type's effort changes. When the CEO discloses fraud, the public learns not only that there is fraud but also that the CEO is informed, and updates the probability of the CEO being the high-type. The CEO's effort does not affect the probability of fraud itself but the low-type CEO becomes more informed, reducing the expected share of the high-types among informed and disclosing CEOs, implying reduced CEO reputation.

Next, the CEO's reputation if there is no disclosure increases in effort (i.e., $\frac{\partial R_1(e)}{\partial e} > 0$). Under no disclosure, there are three possibilities: (i) there is no fraud; (ii) there is fraud but the CEO is uninformed; (iii) there is fraud, the CEO is informed but does not want to disclose it. When the low-type CEO increases effort and becomes more informed, the probability of the high-type in the second scenario (when the CEO is uninformed) increases. Moreover, as the CEO becomes more informed overall, the probability of the second scenario declines, making the other two more probable. The first case of no fraud increases the CEO reputation, but the third one decreases it. We show that the overall effect if positive.

Finally, consider fraud decision by the Manager. He always steals when there is no chance of getting caught, i.e. when $x < \sigma_c$. Because, as we discussed earlier, the CEO's disclosure threshold σ_c increases in the whistleblower's compensation β, there is more fraud for low payoff values, but in this region it is never disclosed. Recall that the second fraud region is when the project payoff is high enough, i.e. when $x \in (t_2, \infty)$. In this region, both the Employee and the CEO disclose fraud if they are informed. Because the CEO's effort increases and the CEO is more informed, fraud is exposed more often. As the probability of getting caught increases, the Manager engages in fraud only when the potential payoff is higher, that is, the fraud threshold $t_2(e)$ increases.

Overall, when the whistleblower compensation β increases, fraud increases for low payoff values and declines for high payoff values. The overall frequency of fraud might increase or decline depending of the payoff distribution: whether high or low payoffs are more probable (the green and red segments in Figure 4, respectively). Fraud for low payoff values is never disclosed, but fraud for high payoff values can be disclosed by both the CEO and the Employee.

Importantly, there is a distinction between committed fraud and instances of fraud disclosure. When the whistleblower compensation β increases, there is less fraud for high payoff values, but the CEO is informed more often due to higher effort. Therefore, the probability of CEO disclosure in this region (which combines the probability of fraud and of CEO being informed) can change in any direction. However, the probability of whistleblower disclosure declines unambiguously because the Employee can only blow the whistle if the CEO is silent as well as because fraud declines. This is consistent with empirical findings of higher whistleblower compensation leading to lower whistleblower cases (e.g., Cordis & Lambert (2017)), despite the fact that overall fraud might have increased or decreased in new equilibrium.

Because protection of whistleblowers is at the center of recent whistleblower programs (e.g., the Sarbanes-Oxley or Dodd-Frank Acts), we also consider the effect of retaliation costs on fraud. Decline in retaliation costs does not affect the equilibrium, unless it changes the type of equilibrium, from the one described in Theorem 1 (Case 1). Although the Employee disclosure threshold declines, this happens in no fraud region and does not affect fraud. Moreover, retaliation has no impact on CEO's payoff and does not change her effort or disclosure. However, if the decline is significant enough to shift the equilibrium to the one described in Theorem 2 (i.e. Case 2), retaliation costs become important. Decline in retaliation costs r decreases the whistleblower's threshold $\sigma_w = \frac{r}{\beta\alpha}$ and leads to more disclosure for low project values. Consequently, fraud declines both for low and high payoffs.

Similarly, if whistleblower compensation β increases in Case 2, it lowers the Employee disclosure threshold $\sigma_w = \frac{r}{\beta\alpha}$ and leads to less fraud.

3.6 Discussion

This section discusses the model assumptions and potential extensions. First, we assume that the Manager's announcement is not observed by the public. This simplifies the analysis because the CEO's reputation in the case of non-disclosure does not depend on the Manager's report. In practice, we view this announcement as an internal report that updates the actors within the firm about the project development. If the firm is private, there is no obligation to share its information with the public. In the case of a public firm, the internal report could be issued before the firm must file a periodic report with the SEC. This assumption is partially motivated by the recent Nissan whistleblower case in which the CEO Carlos Ghosn is accused of funneling funds from the firm.

Second, we assume that the CEO can only disclose verifiable project payoff if there is fraud. The interpretation is that the announcement of fraud makes the authorities start a fraud investigation that reveals the stolen amount αx to the public along with the project payoff x. In practice, public firms publish their statements and, if the internal audit shows no signs of fraud, they do not issue an additional announcement confirming the statement's accuracy. A limitation of our setup is that we also assume that the Manager's report is not observed by the public – at least one of these assumptions should be relaxed in the extension of our model.

Third, the CEO's payoff is assumed to depend on the project payoff x and not the public belief about it. The standard assumption in the disclosure literature is that the firm/management wants to maximize the stock price, which is affected by the beliefs of its investors. In contrast, in our model, the public affects the behaviour of the CEO only through the reputation channel. This allows to simplify the analysis of the model but potentially limits the generalizability of its implications. The assumption might be reasonable in the following contexts. First, x can be viewed as realized dividends, which are not affected by the public beliefs. Second, if the firm is private and there is one main VC/PE fund, then x represents realized profits from the enterprise, which could not be affected by the beliefs of the investor. Finally, the CEO's payoff is realized before the CEO or Employee disclosure and before the report is disclosed to the public. For example, the internal Manager's report is observed by the board of directors that rewards or punishes the CEO based on the report.

There are several potential extensions of the model. Many authors (e.g., Heyes and Kapur (2008) or Howse and Daniels (1995)) argue that whistleblowers often have intrinsic motivation to disclose information about corporate wrongdoings regardless of potential financial payoff. The report by the GAP indicates that sometimes whistleblowers are not even aware of potential rewards. Therefore, a natural extension of the model is to consider two types of whistleblowers: commitment (C) and strategic (S). If $\theta = C$, then the Employee discloses the wrongdoing as soon as he learns about it. If $\theta = S$, then the Employee takes into account the reward he receives (i.e. fraction of the recovered funds) and retaliation costs r. If the fraction of the commitment type employees is not too high, then the equilibrium described in Theorem 1 could remain qualitatively the same. Of course, if the fraction of the commitment type employees is very high, we might get different predictions. For example, in the extreme case where $P(\theta = C) = 1$, the equilibrium does not depend on the strategic whistleblower at all.

In addition, we assume that there is no regulatory penalty for the CEO in case of fraud disclosure by the whistleblower, i.e. $L = 0$. In reality, the CEO might want to avoid potential litigation costs or any involvement in fraud lawsuits (Dyck et al., 2010). In our model, disclosure by the whistleblower lowers the CEO's payoff through both the reputation and project payoff channel. The results would remain qualitatively the same if the regulatory punishment L is not too high, however, the natural extension is to consider $L > 0$.

Finally, in our model the project payoff is random and cannot be affected by the players. In practice, either the Manager or the Employee (or both) may incur costly effort to improve the outcome of the project (see Ting, 2008). Specifically, if the Manager both affects the success of the project and decides whether to steal a portion of the cash flow, we can explore the impact of the whistleblower programs on generated output and fraud involvement by the Manager. Is it optimal for the social planner to allow the Manager to engage in fraud up to some level? In other words, is it better to forgive low levels of fraud to maximize output (similar to Manso (2011) in the context of innovations)?

3.7 Conclusion

The US whistleblower cases have become increasingly frequent in recent years, especially after the Dodd-Frank Act whistleblower provision. Although regulators believe that whistleblower programs decrease corporate fraud, relatively little is known about the effectiveness of the whistleblower programs in terms of deterring fraud. The existing empirical literature on whistleblowing mainly focuses on the ex-post measures of fraud and shows that whistleblowing is associated with lower reported fraud. However, because committed fraud is unobserved and the interested parties (firms, regulators, media, etc) rely on the ex-post fraud measures, the true effect remains ambiguous. It might be the case that increased external monitoring curbs the incentives of the firm to disclose fraud due to reputation concerns of the management or the CEO. What is the effect of whistleblower programs on internal monitoring? Could they have the side effect of increasing the amount of fraud, at least under some circumstances, inducing the opposite of the intended result?

We address these questions with a three-tier model where the Manager chooses whether to engage in fraud, the CEO first exerts costly effort to acquire information about potential fraud and then, if informed, decides whether to disclose the Manager's misbehaviour, and finally, the Employee/whistleblower decides whether to expose fraud. The CEO also cares about her reputation, which may keep her from disclosing fraud. We show that there are no separating equilibria in the game and then focus on one pooling equilibria in which the CEO discloses more fraud then the whistleblower. Interestingly, the equilibrium effort of the CEO increases with the whistleblower compensation but the effect on the ex-ante probability of fraud is obscure: depending on the distribution of cash flows, the ex-ante amount of fraud might both increase or decrease.

3.8 Proofs

PROOF OF PROPOSITION 3:

Consider **a)**: probability of being fraud detection and disclosure is zero => it is optimal for the Manager to engage in fraud.

Consider **b)**: if $x \in [\sigma_c, \sigma_w)$ then there is a chance that the CEO will be informed about the fraud and disclose it to the law enforcement agency. Then the Manager chooses between 0 (if no fraud) and $[\mu p + (1-\mu)pe](-C) + [1 - \mu p - (1-\mu)pe]\alpha x$ (if fraud) – this gives the condition in part b) of Proposition 3.

Consider **c)**: similarly to the previous case, the Manager chooses between 0 (if no fraud) and $y(e)(-C) + (1 - y(e))\alpha x$ where $y(e) = \mu(p + (1-p)q) + (1-\mu)(pe + (1-pe)q)$ is the probability of fraud disclosure under the beliefs e about the low-type CEO's effort. $\square$

For convenience, let's introduce some notations:

- We will say that the condition Σ is satisfied if all inequalities below hold:

  1. $\frac{(w_L - w_H)(1-(1-p)(1-q))}{w_L p(1-q)} < \frac{(1-F(t_2))(F(t_2)+(1-F(t_2))g_H)}{f(t_2)}$
  2. $\int_{t_2(0)}^{\infty} z\alpha(1-q+q\beta) - [(1-q)R_1(e=0) + qR_3(e=0) - R_2(e=0)]dF(z) > \frac{k_L 0}{p w_L} = 0$
  3. $\int_{t_2(1)}^{\infty} z\alpha(1-q+q\beta) - [(1-q)R_1(e=1) + qR_3(e=1) - R_2(e=1)]dF(z) < \frac{k_L}{p w_L}$

- We will say that the condition Ω is satisfied if:

  1. $\sigma_c < \sigma_w$
  2. $t_1(e) > \sigma_w$
  3. $t_2(e) > \sigma_w$
  4. $\sigma_w + R_2 \geq q(\sigma_w(1-\alpha\beta) + R_3) + (1-q)((1-\alpha)\sigma_w) + R_1)$
  5. Condition Σ is satisfied

PROOF OF LEMMA 2: The first-order condition with respect to the choice of effort $\hat{e}$:

$$\frac{\partial V_c(\hat{e})}{\partial \hat{e}} = 0 \iff p\Bigg(\int_0^{\sigma_c} \{w_L(1-\alpha)z + R_1 + (1-w_L)(z+R_1)\}\, dF(z) +$$
$$+\int_{\sigma_c}^{t_2(e)} (z+R_1)dF(z) + \int_{t_2(e)}^{\infty} \{w_L(z+R_2) + (1-w_L)(z+R_1)\}\, dF(z)\Bigg) -$$
$$-\, p\Bigg(\int_0^{\sigma_c} \{w_L((1-\alpha)z + R_1) + (1-w_L)(z+R_1)\}\, dF(z) + \int_{\sigma_c}^{t_2(e)} (z+R_1)dF(z) +$$
$$+\int_{t_2(e)}^{\infty} \{w_L\left(q(z(1-\alpha\beta)+R_3) + (1-q)(z+R_1)\right) + (1-w_L)(z+R_1)\}\, dF(z)\Bigg) = k_L\hat{e}$$

This equality could be rewritten in the following way:

$$\int_{t_2(e)}^{\infty} z\alpha(1-q+q\beta) - [(1-q)R_1 + qR_3 - R_2]dF(z) = \frac{k_L\hat{e}}{pw_L}$$

Next we will show that the LHS(e) of the latter equality is decreasing in e. With proper restrictions on LHS($e=0$) and LHS($e=1$), this will ensure that there is the unique e^* s.t. LHS($e=e^*$) = RHS($\hat{e}=e^*$).

$$\frac{\partial LHS(e)}{\partial e} = -\frac{\partial t_2(e)}{\partial e}\left[t_2(e)\alpha(1-q+q\beta) + R_2 - qR_3 - (1-q)R_1\right] +$$
$$+\int_{t_2(e)}^{\infty}\left[\frac{\partial R_2}{\partial e} - q\frac{\partial R_3}{\partial e} - (1-q)\frac{\partial R_1}{\partial e}\right] dF(z) =$$
$$= -\frac{\partial t_2(e)}{\partial e} t_2(e)\alpha(1-q+q\beta) + \frac{\partial t_2(e)}{\partial e}(qR_3 + (1-q)R_1 - R_2) +$$
$$+\,(1-F(t_2(e)))\left[\frac{\partial R_2}{\partial e} - q\frac{\partial R_3}{\partial e} - (1-q)\frac{\partial R_1}{\partial e}\right]$$

Next, recall that $t_2(e) = \frac{Cy(e)}{\alpha(1-y(e))}$, where $y(e) := \mu(p+(1-p)q) + (1-\mu)(pe+(1-pe)q)$

$=> \frac{\partial t_2(e)}{\partial e} = C\frac{p(1-\mu)(1-q)}{\alpha(1-y(e))^2} > 0 \ \forall e \in [0,1]$.

$$\frac{\partial R_2}{\partial e} = -\frac{\mu(1-\mu)\frac{w_L}{w_H}}{\left(\mu + (1-\mu)\frac{ew_L}{w_H}\right)^2} < 0$$

$$\frac{\partial R_3}{\partial e} = \frac{\mu(1-\mu)p\frac{w_L}{(1-p)w_H}}{\mu + (1-\mu)\frac{1-pe}{1-p}\frac{w_L}{w_H}\Big)^2} > 0$$

$$\frac{\partial R_1}{\partial e} = -\frac{\mu(1-\mu)\left(\frac{P(ND|\tau=L)}{P(ND|\tau=H)}\right)'_e}{\left(\mu + (1-\mu)\frac{P(ND|\tau=L)}{P(ND|\tau=H)}\right)^2}$$

$$\left(\frac{P(ND|\tau=L)}{P(ND|\tau=H)}\right)'_e = \frac{[g_h - g_L]\frac{\partial t_2(e)}{\partial e}f(t_2(e)) - w_Lp(1-q)(1-F(t_2(e)))[F(t_2(e) + (1-F(t_2(e))g_h]}{(P(ND|\tau=H))^2},$$

where $g_H = (1-w_H) + w_H(1-p)(1-q)$ and $g_L = (1-w_L) + w_L(1-pe)(1-q)$.

It is sufficient to ensure that the latter derivative is negative – in that case we have that LHS decreases monotonically with e.

Assumption A : $\frac{(w_L-w_H)(1-(1-p)(1-q))}{w_Lp(1-q)} < \frac{(1-F(t_2))(F(t_2)+(1-F(t_2))g_H)}{f(t_2)}$

If this assumption is satisfied then LHS monotonically decreases with e[4]. Now let's consider the boundary conditions:

- LHS(0) $= \int_{t_2(0)}^{\infty} z\alpha(1-q+q\beta) - [(1-q)R_1(e=0) + qR_3(e=0) - R_2(e=0)]dF(z) > \frac{k_L 0}{pw_L} = 0$
- LHS(1) $= \int_{t_2(1)}^{\infty} z\alpha(1-q+q\beta) - [(1-q)R_1(e=1) + qR_3(e=1) - R_2(e=1)]dF(z) < \frac{k_L}{pw_L}$

Hence, if the above three conditions are satisfied (we will denote them as Σ), the CEO's effort problem has the unique solution. □

PROOF OF THEOREM 1:

First, recall that we assumed the following: $\sigma_c < \sigma_w$, $t_1(e) > \sigma_w$ and $t_2(e) > \sigma_w$. Combined with Lemma 2, we need only show that the CEO's equilibrium strategy is indeed a

[4]It can be shown that there exist distribution functions F with pdf f s.t. *Assumption A* is satisfied for all possible values of effort e.

single-threshold disclosure rule with the cutoff $\sigma_c = \frac{R_1 - R_2}{\alpha}$.

Note that at $x = \sigma_c$ the CEO is indifferent between disclosure and non-disclosure. When x is increasing, then it is clearly better for the CEO to disclose her signal since in the case of non-disclosure she loses a portion of the payoff as long as this signal x is lower than σ_w.

At $x = \sigma_w$, we need to ensure that

$$\sigma_w + R_2 \geq q(\sigma_w(1 - \alpha\beta) + R_3) + (1 - q)((1 - \alpha)\sigma_w) + R_1). \tag{3.3}$$

This inequality holds for σ_w large enough (recall that σ_w depends only on exogenous parameters). Overall, we have to ensure that the following conditions Ω are satisfied[5]: (i) $\sigma_c < \sigma_w$; (ii) $t_1(e) > \sigma_w$; (iii) $t_2(e) > \sigma_w$; (iv) equation (3); (v) conditions Σ from Lemma 2.

[5]It can be shown that these conditions are compatible with each other. Note also that some conditions include the equilibrium effort but this is not a concern since effort is from 0 to 1. For example, for (ii) and (iii) we can consider $e = 0$ since both $t_1(\cdot)$ and $t_2(\cdot)$ are increasing in e.

PROOF OF PROPOSITION 4:

1. The derivative of the optimal CEO effort with respect to β is equal to

$$\frac{de^*(\beta)}{d\beta} = -\frac{\partial^2 V(e^*(\beta),\beta)}{\partial e \partial \beta} \bigg/ \frac{\partial^2 V(e^*(\beta),\beta)}{\partial^2 e} = \frac{pq\alpha w_L}{k_L} \int_{t_2(e)}^{\infty} z dF(z) > 0 \tag{3.4}$$

   All the parameters and the integral are positive, hence, effort increases in β.

2. $\frac{\partial t_2(e)}{\partial e} = -\frac{p(1-\mu)(1-q)}{\alpha(1-y(e))^2} > 0$ Fraud threshold t_2 increases in effort.

3. CEO disclosure threshold:

$$\frac{d\sigma_c(e^*(\beta))}{d\beta} = \frac{\partial \sigma_c(e^*(\beta))}{\partial e} \frac{\partial e^*(\beta)}{\partial \beta} \tag{3.5}$$

   From Equation (3.4), the first derivative is positive and effort increases in β. The derivative of σ_c with respect to effort is:

$$\frac{\partial \sigma_c(e)}{\partial e} = \frac{1}{\alpha} \left(\frac{\partial R_1(e)}{\partial e} - \frac{\partial R_2(e)}{\partial e} \right) \tag{3.6}$$

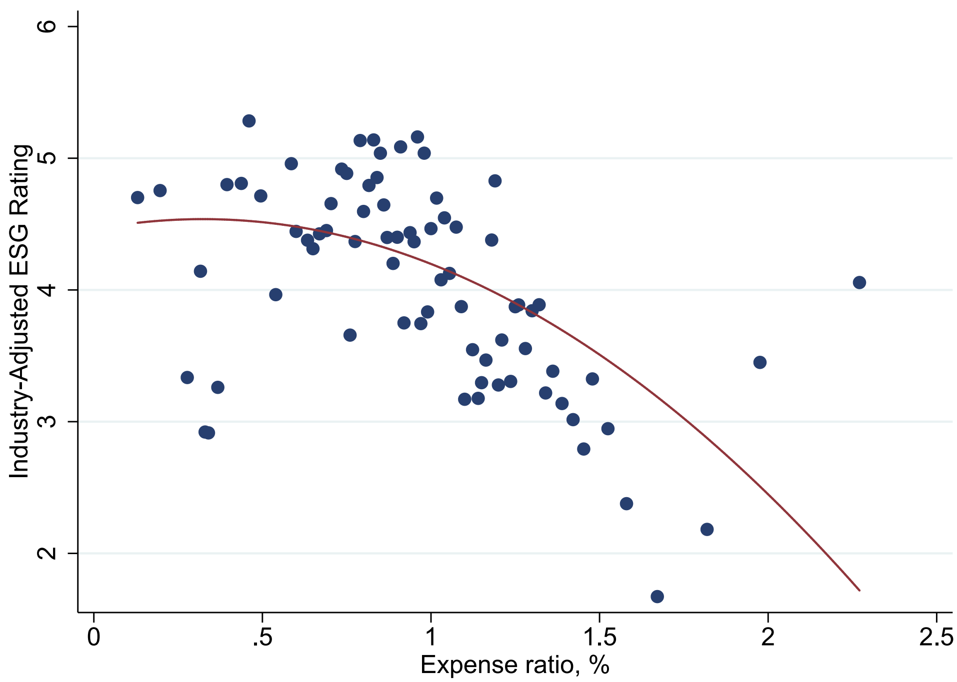
   $\frac{\partial R_2(e)}{\partial e} = -\frac{\mu(1-\mu)\frac{w_L}{w_H}}{\left(\mu+(1-\mu)\frac{ew_L}{w_H}\right)} < 0$ and $\frac{\partial R_1(e)}{\partial e} > 0$ under *Assumption A* from Lemma 2. Hence, $\frac{\partial \sigma_c(e)}{\partial e} > 0$ and $\frac{d\sigma_c(e^*(\beta))}{d\beta} > 0$.

□

APPENDIX A

Additional Figures and Tables for Chapter 1

Figure A.1: Industy-Adjusted ESG Fund Rating and Fund Expense Ratio: Quadratic Fit



Note: The scatterplot illustrates the relationship between the industry-adjusted ESG fund ratings and the fund expense ratios. A total of 6,351 observations are uniformly distributed across 70 bins. The scatterplot includes month fixed effects but does not control for any fund characteristics.

Table A.1: Fund ESG ratings and Expense Ratio

	(1) ESG rating (ind-adj)	(2) ESG rating (w.-av.)	(3) Environmental Fund Score	(4) Social Fund Score	(5) Governance Fund Score
Expense Ratio, percent	-0.382***	-0.168	-0.415**	-0.185*	-0.115
	(0.131)	(0.113)	(0.162)	(0.101)	(0.156)
Percent in Equities	-0.036	-0.049**	-0.052**	-0.051**	-0.077***
	(0.024)	(0.020)	(0.024)	(0.020)	(0.026)
Percent in Equities Squared	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Monthly Returns (after fee)	0.014**	0.014***	0.021***	0.011***	0.014**
	(0.006)	(0.005)	(0.005)	(0.004)	(0.006)
Monthly Returns Squared (after fee)	-0.001	0.000	-0.000	0.000	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(Total Net Assets)	0.062	0.068	-0.000	0.063	0.125
	(0.073)	(0.061)	(0.085)	(0.056)	(0.076)
Log(Total Net Assets) Squared	-0.002	-0.003	0.006	-0.001	-0.009
	(0.009)	(0.007)	(0.009)	(0.007)	(0.008)
Log(Age in quarters)	-0.452***	-0.296***	-0.464***	-0.210**	-0.251**
	(0.128)	(0.090)	(0.122)	(0.091)	(0.110)
Log(Age in quarters) Squared	0.095***	0.065***	0.100***	0.041**	0.057**
	(0.028)	(0.020)	(0.025)	(0.020)	(0.023)
Constant	3.593***	3.500***	4.577***	3.347***	4.291***
	(0.911)	(0.740)	(0.828)	(0.736)	(0.982)
Month FE	Yes	Yes	Yes	Yes	Yes
R-Squared	0.550	0.556	0.508	0.579	0.567
Observations	6281	6281	6281	6281	6281

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Note: The table shows estimated slope coefficients for the OLS regression of funds' ESG performance on expense ratios in the period from January 2011 to March 2022. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. In columns (3)-(5) the dependent variables are Environmental, Social, and Governance Scores, respectively. All regressions include time fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table A.2: Fund ESG ratings and Expense Ratio

	(1) ESG rating (ind-adj)	(2) ESG rating (w.-av.)	(3) Environmental Fund Score	(4) Social Fund Score	(5) Governance Fund Score
Expense Ratio, percent	0.035	0.275	-0.176	0.343	0.508
	(0.438)	(0.358)	(0.450)	(0.335)	(0.447)
Expense Ratio Squared	-0.194	-0.207	-0.112	-0.246*	-0.290
	(0.171)	(0.154)	(0.197)	(0.141)	(0.215)
Percent in Equities	-0.036	-0.050**	-0.053**	-0.052***	-0.079***
	(0.024)	(0.020)	(0.024)	(0.020)	(0.026)
Percent in Equities Squared	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Monthly Returns (after fee)	0.014**	0.014***	0.021***	0.011***	0.014**
	(0.006)	(0.005)	(0.005)	(0.004)	(0.006)
Monthly Returns Squared (after fee)	-0.000	0.000	-0.000	0.000	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(Total Net Assets)	0.058	0.064	-0.002	0.058	0.119
	(0.074)	(0.062)	(0.086)	(0.058)	(0.076)
Log(Total Net Assets) Squared	-0.001	-0.003	0.006	-0.000	-0.009
	(0.009)	(0.007)	(0.009)	(0.007)	(0.009)
Log(Age in quarters)	-0.447***	-0.290***	-0.461***	-0.204**	-0.243**
	(0.127)	(0.089)	(0.121)	(0.091)	(0.111)
Log(Age in quarters) Squared	0.094***	0.063***	0.099***	0.039*	0.054**
	(0.028)	(0.020)	(0.024)	(0.020)	(0.024)
Constant	3.439***	3.337***	4.489***	3.153***	4.062***
	(0.926)	(0.752)	(0.834)	(0.760)	(0.991)
Month FE	Yes	Yes	Yes	Yes	Yes
R-Squared	0.552	0.559	0.509	0.584	0.572
Observations	6281	6281	6281	6281	6281

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Note: The table shows estimated slope coefficients for the OLS regression of funds' ESG performance on expense ratios in the period from January 2011 to March 2022. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. In columns (3)-(5) the dependent variables are Environmental, Social, and Governance Scores, respectively. All regressions include time fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

Table A.3: Fund ESG ratings and Expense Ratio: No Controls

	(1) ESG rating (ind-adj)	(2) ESG rating (w.-av.)	(3) Environmental Fund Score	(4) Social Fund Score	(5) Governance Fund Score
Expense Ratio, percent	0.145	0.314	-0.109	0.320	0.512
	(0.706)	(0.590)	(0.710)	(0.568)	(0.676)
Expense Ratio Squared	-0.313	-0.295	-0.230	-0.315	-0.379
	(0.280)	(0.240)	(0.300)	(0.227)	(0.287)
Constant	4.551***	3.845***	4.857***	3.597***	4.033***
	(0.412)	(0.336)	(0.404)	(0.325)	(0.368)
Month FE	Yes	Yes	Yes	Yes	Yes
R-Squared	0.179	0.075	0.138	0.105	0.184
Observations	6351	6351	6351	6351	6351

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

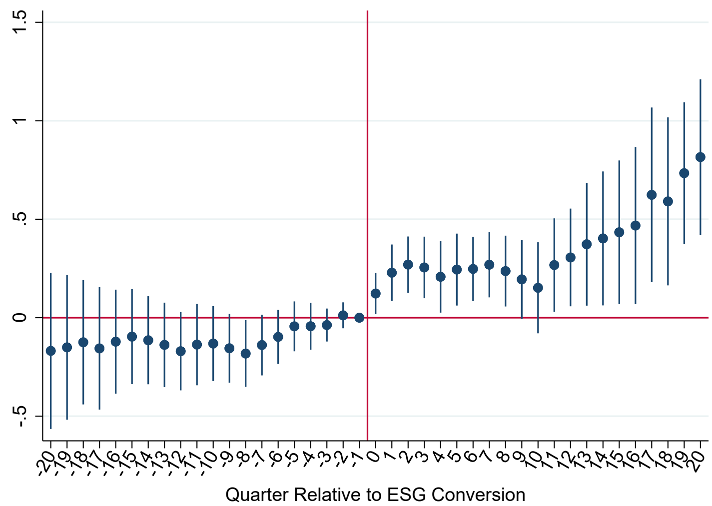
Note: The table shows estimated slope coefficients for the OLS regression of funds' ESG performance on expense ratios in the period from January 2011 to March 2022. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. In columns (3)-(5) the dependent variables are Environmental, Social, and Governance Scores, respectively. All regressions include time fixed effects. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. t-statistics are reported in parentheses, standard errors are clustered at the fund level.

APPENDIX B

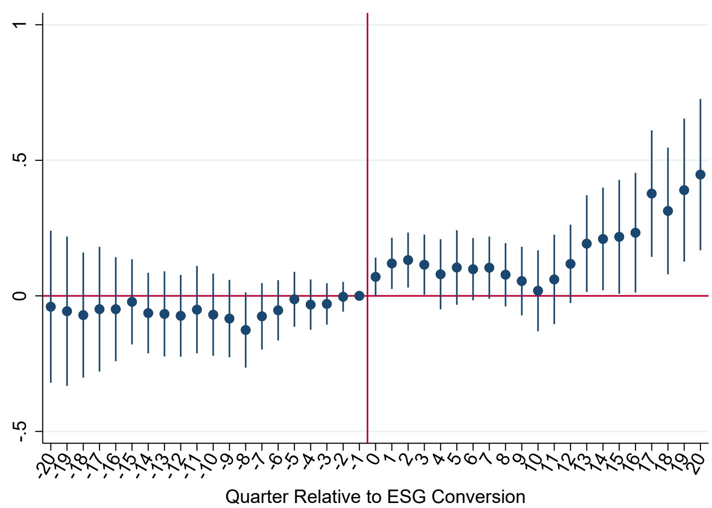
Additional Figures and Tables for Chapter 2

Figure B.1: Treatment Effect of Fund Repurposing on ESG Ratings, Only Funds that Eventually Converge

(a) Industry-Adjusted ESG Ratings



(b) Weighted-Average ESG Ratings



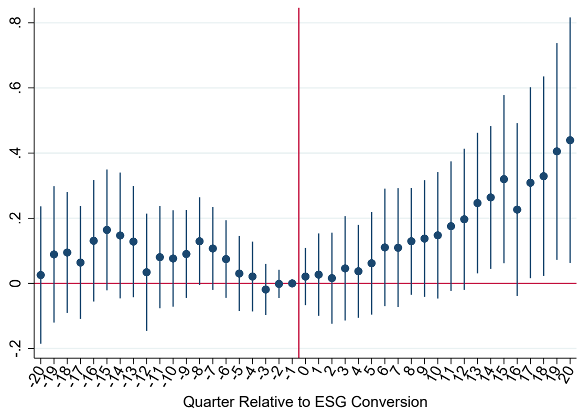
Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on funds ESG ratings estimated from the following specification:

$$Rating_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

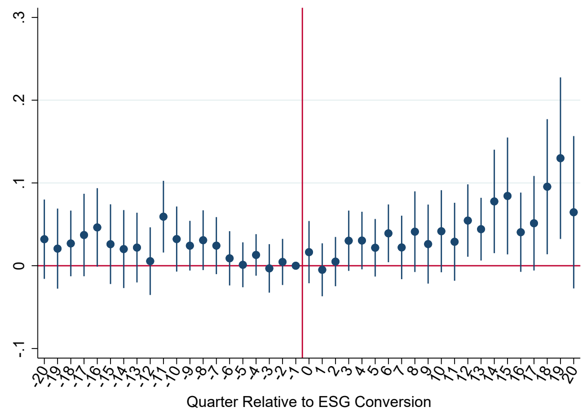
where $Rating_{it}$ is fund's i ESG rating, λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 58 repurposed funds.

Figure B.2: Treatment Effect of Fund Repurposing on Fund Flows, Only Funds that Eventually Converge

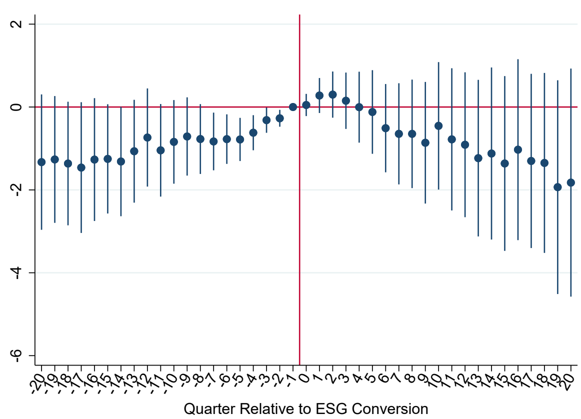
(a) Log of Annual Flows



(b) Log of Quarterly Flows



(c) Number of Quarters with Negative Fund Flows over the Last 2 Years



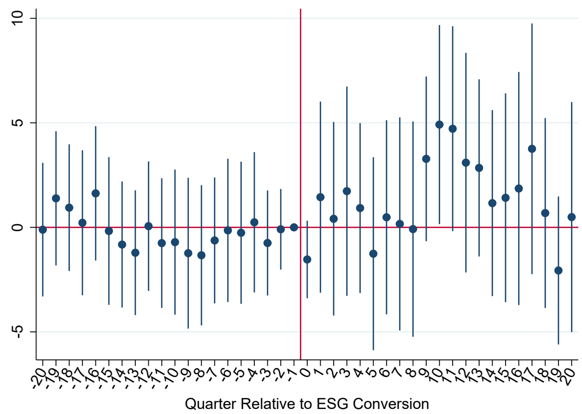
Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on fund flows estimated from the following specification:

$$Log(1 + Flow_{it}) = \lambda_t + \alpha_{j(i)} + X_{it} + \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

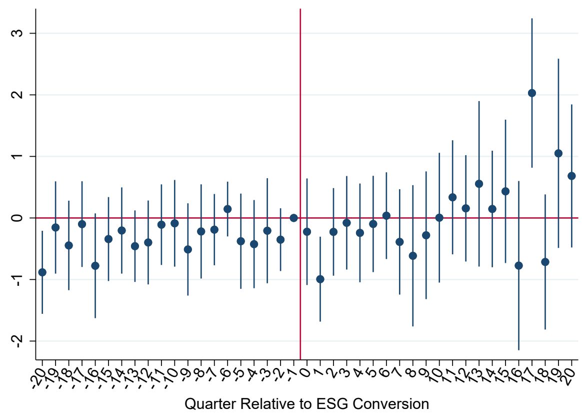
where $Flow_{it}$ is i fund's flows in period t, λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 58 repurposed funds.

Figure B.3: Treatment Effect of non-ESG Fund Repurposing on Select Outcomes

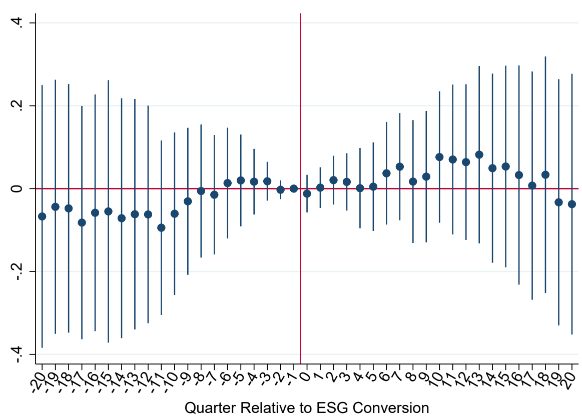
(a) Annual Returns



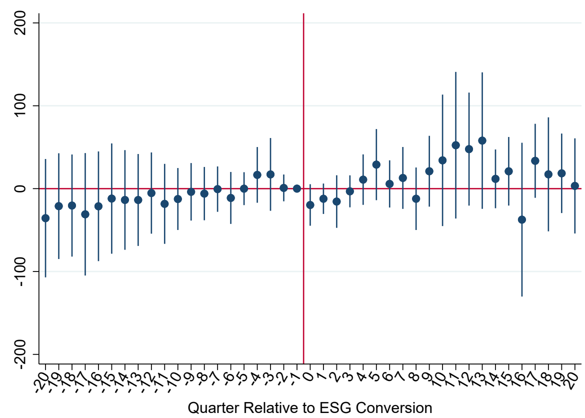
(b) Fama-French Three-Factor Alpha



(c) Expense Ratio



(d) Profits



Note: These figures show the dynamic treatment effect of non-ESG funds conversion to ESG on select outcomes estimated from the following specification:

$$Outcome_{it} = \lambda_t + \alpha_{j(i)} + X_{it} + \left[\sum_{j=-20, j\neq -1}^{20} \beta_j I\{t - t_i^* = j\} \right] + \varepsilon_{it},$$

where λ_t and $\alpha_{j(i)}$ are time and industry fixed effects, respectively. X_{it} are fund characteristics: percent of holdings in common stocks, fund age, quarterly returns, quarterly flows, average AUM over last year, expense ratio. Vertical bars show 95% confidence intervals. Standard errors are clustered at the fund level. The sample covers 58 repurposed funds.

Table B.1: Fund Repurposing and Fund Ratings: Event Study

	(1) Industry-Adjusted	(2) Weighted-Average	(3) Environmental	(4) Social	(5) Governance
β_{-6}	-0.118	-0.069	0.023	-0.101	-0.050
	(0.077)	(0.061)	(0.080)	(0.064)	(0.074)
β_{-5}	-0.062	-0.027	0.090	-0.055	-0.042
	(0.068)	(0.054)	(0.078)	(0.055)	(0.068)
β_{-4}	-0.055	-0.041	0.017	-0.046	-0.077
	(0.063)	(0.049)	(0.074)	(0.048)	(0.064)
β_{-3}	-0.043	-0.034	0.008	-0.037	-0.073
	(0.042)	(0.039)	(0.061)	(0.037)	(0.056)
β_{-2}	0.007	-0.008	0.012	-0.013	-0.024
	(0.033)	(0.027)	(0.044)	(0.026)	(0.050)
β_{-1}	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)
β_0	0.127*	0.074*	0.063	0.102**	0.038
	(0.052)	(0.036)	(0.051)	(0.037)	(0.049)
β_1	0.237**	0.127*	0.081	0.168**	0.056
	(0.073)	(0.049)	(0.067)	(0.054)	(0.069)
β_2	0.282***	0.142*	0.046	0.198**	0.084
	(0.074)	(0.053)	(0.075)	(0.058)	(0.076)
β_3	0.272**	0.131*	0.009	0.188**	0.073
	(0.081)	(0.059)	(0.077)	(0.063)	(0.091)
β_4	0.227*	0.098	-0.040	0.165*	0.065
	(0.096)	(0.069)	(0.090)	(0.073)	(0.098)
β_5	0.271*	0.128	-0.007	0.187*	0.112
	(0.102)	(0.075)	(0.101)	(0.079)	(0.105)
β_6	0.276**	0.121	0.014	0.183*	0.101
	(0.099)	(0.070)	(0.094)	(0.075)	(0.100)
Percent in Equities	0.023**	0.025***	0.031***	0.023***	0.029***
	(0.007)	(0.005)	(0.006)	(0.006)	(0.005)
Log Age (in quarters)	-0.057	-0.042	0.014	-0.051	-0.107*
	(0.044)	(0.036)	(0.051)	(0.041)	(0.042)
Log AUM	0.023	0.014	-0.031	0.022	0.034
	(0.038)	(0.029)	(0.029)	(0.031)	(0.035)
N	9185	9185	9185	9185	9185
R^2	0.781	0.782	0.807	0.731	0.776
Y mean	3.588	3.346	3.935	3.129	3.740
Y sd	1.192	1.019	1.318	0.951	1.248

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports estimates of the effect of fund repurposing on fund ratings from equation 2.3 that includes twenty pre and twenty post-periods. Only six are reported due to space constraints. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. In columns (3)-(5), the dependent variables are the fund's environmental, social, and governance ratings, respectively. Coefficients β_i caputure the difference between treated and control funds i periods after the treatment. $\beta_{-1} = 0$ since $t = -1$ is set as the reference period. AUM is average assets under management over the last 12 months. All regressions include time and fund-type fixed effects. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Table B.2: Fund Repurposing and Fund Performance: Diff-in-Diff with no controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fama-French Three-Factor Alpha	Annual returns	Quarterly returns	Log of annual flows	Log of quarterly flows	Quarters with Negative Fund Flows	Expense Ratio	Profit
D_it	0.048	-0.633	-0.051	0.041	0.011	0.246	-0.035	-105.701
	(0.109)	(0.903)	(0.206)	(0.040)	(0.009)	(0.334)	(0.055)	(161.241)
N	9367	9367	9367	8800	9094	9367	8257	8253
R^2	0.097	0.614	0.772	0.058	0.014	0.148	0.051	0.011
Y mean	-0.344	12.021	2.959	-0.100	-0.019	5.251	1.115	506.262
Y sd	2.412	17.626	8.914	0.354	0.091	2.616	0.492	1034.038

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the estimates of the treatment effect of fund repurposing on different outcomes according to equation 2.4. D_{it} is a dummy variable that equals one if the fund has already been repurposed, and zero otherwise. AUM is the average assets under management over the last 12 months. All regressions include time fixed effects. Standard errors are clustered at the fund level. The sample covers 257 funds, including 58 repurposed funds.

Table B.3: Fund Repurposing and Fund Ratings: Diff-in-Diff, Only Funds that Eventually Converge

	(1) ESG rating (ind-adj)	(2) ESG rating (w.-av.)	(3) Environmental Fund Rating	(4) Social Fund Rating	(5) Governance Fund Rating
D_it	0.333**	0.169*	-0.027	0.239**	0.151
	(0.111)	(0.079)	(0.097)	(0.088)	(0.094)
Percent in equities	0.022**	0.024***	0.032***	0.022**	0.029***
	(0.008)	(0.005)	(0.006)	(0.006)	(0.005)
log Age (in quarters)	-0.027	-0.029	-0.007	-0.035	-0.082*
	(0.036)	(0.031)	(0.041)	(0.035)	(0.038)
log AUM	0.014	0.010	-0.021	0.018	0.026
	(0.032)	(0.025)	(0.026)	(0.026)	(0.030)
N	1386	1386	1386	1386	1386
R^2	0.840	0.875	0.868	0.830	0.868
Y mean	3.938	3.551	4.265	3.313	3.888
Y sd	1.061	0.874	1.126	0.821	1.076

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the estimates of the treatment effect of fund repurposing on fund ratings according to equation 2.4. The dependent variable in column (1) is the fund's industry-adjusted ESG rating, while in column (2), the dependent variable is the fund's weighted-average ESG rating. D_{it} is a dummy variable that equals one if the fund has already been repurposed, and zero otherwise. AUM is average assets under management over the last 12 months. All regressions include time and fund-type fixed effects. Standard errors are clustered at the fund level. The sample covers 58 repurposed funds.

Table B.4: Fund Repurposing and Fund Performance: Diff-in-Diff, Only Funds that Eventually Converge

	(1) Fama-French Three-Factor Alpha	(2) Annual returns	(3) Quarterly returns	(4) Log of annual flows	(5) Log of quarterly flows	(6) Quarters with Negative Fund Flows	(7) Expense Ratio	(8) Profit
D_it	0.040	1.147	0.200	0.088	0.021	-0.245	-0.080	-67.809
	(0.193)	(0.905)	(0.284)	(0.068)	(0.014)	(0.451)	(0.083)	(319.696)
Percent in equities	0.008	0.089	0.002	0.000	0.000	-0.013	-0.016**	2.888
	(0.014)	(0.062)	(0.013)	(0.003)	(0.001)	(0.019)	(0.005)	(8.889)
log Age (in quarters)	-0.014	0.000	0.084	-0.114**	-0.025***	1.630***	0.201***	184.905
	(0.085)	(0.406)	(0.137)	(0.040)	(0.007)	(0.209)	(0.047)	(135.402)
log AUM	-0.059	0.091	-0.012	0.005	-0.002	0.218*	-0.039	296.684**
	(0.048)	(0.195)	(0.044)	(0.016)	(0.003)	(0.103)	(0.026)	(107.995)
N	1386	1386	1386	1366	1376	1386	1377	1377
R^2	0.195	0.788	0.845	0.204	0.166	0.605	0.464	0.534
Y mean	-0.325	12.441	3.020	-0.083	-0.015	5.599	1.066	456.284
Y sd	2.361	16.182	8.441	0.341	0.082	2.599	0.460	981.170

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the estimates of the treatment effect of fund repurposing on different outcomes according to equation 2.4. D_{it} is a dummy variable that equals one if the fund has already been repurposed, and zero otherwise. AUM is the average assets under management over the last 12 months. All regressions include time and fund-type fixed effects. Standard errors are clustered at the fund level. The sample covers 58 repurposed funds.

Table B.5: Fund Repurposing and Fund Flows (six quarters leads), Only Funds that Eventually Converge

	(1) F6.Log of annual flows	(2) F6.Log of quarterly flows	(3) F6.Quarters with Negative Fund Flows
D_it	0.158*	0.034*	-1.171*
	(0.069)	(0.013)	(0.526)
Percent in equities	0.003	0.000	-0.034
	(0.003)	(0.001)	(0.020)
log Age (in quarters)	-0.066*	-0.012	1.092***
	(0.032)	(0.006)	(0.208)
log AUM	-0.025	-0.006	0.433***
	(0.017)	(0.004)	(0.108)
N	1092	1092	1092
R^2	0.262	0.162	0.607
Y mean	-0.081	-0.016	5.897
Y sd	0.324	0.079	2.430

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the estimates of the treatment effect of fund repurposing on six quarters leads of fund flows according to equation 2.4. D_{it} is a dummy variable that equals one if the fund has already been repurposed, and zero otherwise. AUM is average assets under management over the last 12 months. All regressions include time and fund-type fixed effects. Standard errors are clustered at the fund level. The sample covers 58 repurposed funds.

BIBLIOGRAPHY

Abadie, Alberto (2021) "Using synthetic controls: Feasibility, data requirements, and methodological aspects," *Journal of Economic Literature*, 59 (2), 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010) "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American statistical Association*, 105 (490), 493–505.

——— (2015) "Comparative politics and the synthetic control method," *American Journal of Political Science*, 59 (2), 495–510.

Abadie, Alberto and Javier Gardeazabal (2003) "The economic costs of conflict: A case study of the Basque Country," *American Economic Review*, 93 (1), 113–132.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous (2012) "The environment and directed technical change," *American Economic Review*, 102 (1), 131–166.

Acharya, Viral V, Peter DeMarzo, and Ilan Kremer (2011) "Endogenous information flows and the clustering of announcements," *American Economic Review*, 101 (7), 2955–2979.

Albuquerque, Rui A, Yrjo Koskinen, and Raffaele Santioni (2024) "Mutual Fund Trading, Greenwashing, and ESG Clientele."

Albuquerque, Rui, Yrjö Koskinen, and Chendi Zhang (2019) "Corporate social responsibility and firm risk: Theory and empirical evidence," *Management Science*, 65 (10), 4451–4469.

Bailey, Warren, Alok Kumar, and David Ng (2011) "Behavioral biases of mutual fund investors," *Journal of Financial Economics*, 102 (1), 1–27.

Baker, Malcolm, Daniel Bergstresser, George Serafeim, and Jeffrey Wurgler (2022) "The pricing and ownership of US green bonds," *Annual Review of Financial Economics*, 14, 415–437.

Barber, Brad M, Adair Morse, and Ayako Yasuda (2021) "Impact investing," *Journal of Financial Economics*, 139 (1), 162–185.

Barber, Brad M, Terrance Odean, and Lu Zheng (2005) "Out of sight, out of mind: The effects of expenses on mutual fund flows," *The Journal of Business*, 78 (6), 2095–2120.

Bauer, Rob, Kees Koedijk, and Roger Otten (2005) "International evidence on ethical mutual fund performance and investment style," *Journal of banking & finance*, 29 (7), 1751–1767.

Bauer, Rob, Tobias Ruof, and Paul Smeets (2021) "Get real! Individuals prefer more sustainable investments," *The Review of Financial Studies*, 34 (8), 3976–4043.

Van der Beck, Philippe (2021) "Flow-driven ESG returns," *Swiss Finance Institute Research Paper* (21-71).

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song (2022) "What do mutual fund investors really care about?" *The Review of Financial Studies*, 35 (4), 1723–1774.

Berg, Florian, Julian F Koelbel, Anna Pavlova, and Roberto Rigobon (2022a) "ESG confusion and stock returns: Tackling the problem of noise,"Technical report, National Bureau of Economic Research.

Berg, Florian, Julian F Koelbel, and Roberto Rigobon (2022b) "Aggregate confusion: The divergence of ESG ratings," *Review of Finance*, 26 (6), 1315–1344.

Berger, Philip G and Heemin Lee (2022) "Did the Dodd–Frank whistleblower provision deter accounting fraud?" *Journal of Accounting Research*, 60 (4), 1337–1378.

Berk, Jonathan B and Richard C Green (2004) "Mutual fund flows and performance in rational markets," *Journal of Political Economy*, 112 (6), 1269–1295.

Berk, Jonathan and Jules H Van Binsbergen (2021) "The impact of impact investing."

Bloomberg (2022) "ESG by the Numbers: Sustainable Investing Set Records in 2021," https://www.bloomberg.com/news/articles/2022-02-03/esg-by-the-numbers-sustainable-investing-set-records-in-2021/, accessed May 2024.

Bolton, Patrick and Marcin Kacperczyk (2021) "Do investors care about carbon risk?" *Journal of Financial Economics*, 142 (2), 517–549.

Broccardo, Eleonora, Oliver Hart, and Luigi Zingales (2022) "Exit versus voice," *Journal of Political Economy*, 130 (12), 3101–3145.

Carhart, Mark M (1997) "On persistence in mutual fund performance," *The Journal of Finance*, 52 (1), 57–82.

Cen, Xiao, Winston Wei Dou, Leonid Kogan, and Wei Wu (2023) "Fund flows and income risk of fund managers,"Technical report, National Bureau of Economic Research.

Cheng, Ing-Haw, Harrison Hong, and Kelly Shue (2023) "Do managers do good with other people's money?" *The Review of Corporate Finance Studies*, 12 (3), 443–487.

Christoffersen, Susan EK and David K Musto (2002) "Demand curves and the pricing of money management," *The Review of Financial Studies*, 15 (5), 1499–1524.

Cochrane, John (2009) *Asset pricing: Revised edition*: Princeton University Press.

Cordis, Adriana S and Elizabeth M Lambert (2017) "Whistleblower laws and corporate fraud: Evidence from the United States," in *Accounting Forum*, 41, 289–299, Elsevier.

Dey, Aiyesha, Jonas Heese, and Gerardo Pérez-Cavazos (2021) "Cash-for-information whistleblower programs: Effects on whistleblowing and consequences for whistleblowers," *Journal of Accounting Research*, 59 (5), 1689–1740.

Dikolli, Shane S, Mary Margaret Frank, Zhe Michael Guo, and Luann J Lynch (2022) "Walk the talk: ESG mutual fund voting on shareholder proposals," *Review of Accounting Studies*, 27 (3), 864–896.

Doudchenko, Nikolay and Guido W Imbens (2016) "Balancing, regression, difference-in-differences and synthetic control methods: A synthesis," Technical report, National Bureau of Economic Research.

Dyck, Alexander, Adair Morse, and Luigi Zingales (2010) "Who blows the whistle on corporate fraud?" *The Journal of Finance*, 65 (6), 2213–2253.

Dye, Ronald A (1985) "Disclosure of nonproprietary information," *Journal of Accounting Research*, 123–145.

——— (2017) "Optimal disclosure decisions when there are penalties for nondisclosure," *The RAND Journal of Economics*, 48 (3), 704–732.

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

Elton, Edwin J, Martin J Gruber, and Christopher R Blake (2003) "Incentive fees and mutual funds," *The Journal of Finance*, 58 (2), 779–804.

Fama, Eugene F and Kenneth R French (1993) "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33 (1), 3–56.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu (2020) "Taming the factor zoo: A test of new factors," *The Journal of Finance*, 75 (3), 1327–1370.

Flammer, Caroline (2021) "Corporate green bonds," *Journal of Financial Economics*, 142 (2), 499–516.

Frenkel, Sivan, Ilan Guttman, and Ilan Kremer (2020) "The effect of exogenous information on voluntary disclosure and market quality," *Journal of Financial Economics*, 138 (1), 176–192.

Friede, Gunnar, Timo Busch, and Alexander Bassen (2015) "ESG and financial performance: aggregated evidence from more than 2000 empirical studies," *Journal of Sustainable Finance & Investment*, 5 (4), 210–233.

Gibson Brandon, Rajna, Simon Glossner, Philipp Krueger, Pedro Matos, and Tom Steffen (2022) "Do responsible investors invest responsibly?" *Review of Finance*, 26 (6), 1389–1432.

Gil-Bazo, Javier and Pablo Ruiz-Verdu (2008) "When cheaper is better: Fee determination in the market for equity mutual funds," *Journal of Economic Behavior & Organization*, 67 (3-4), 871–885.

Gil-Bazo, Javier and Pablo Ruiz-Verdú (2009) "The relation between price and performance in the mutual fund industry," *The Journal of Finance*, 64 (5), 2153–2183.

Gil-Bazo, Javier, Pablo Ruiz-Verdú, and André AP Santos (2010) "The performance of socially responsible mutual funds: The role of fees and management companies," *Journal of Business Ethics*, 94, 243–263.

Gillan, Stuart L, Andrew Koch, and Laura T Starks (2021) "Firms and social responsibility: A review of ESG and CSR research in corporate finance," *Journal of Corporate Finance*, 66, 101889.

Gillan, Stuart L, Naciye Sekerci, and Laura T Starks (2022) "Managerial Response to Investor Environmental Demand: The Role of Firm Ownership Structure."

Givati, Yehonatan (2016) "A theory of whistleblower rewards," *The Journal of Legal Studies*, 45 (1), 43–72.

Goldstein, Itay, Alexandr Kopytov, Lin Shen, and Haotian Xiang (2022) "On ESG investing: Heterogeneous preferences, information, and asset prices,"Technical report, National Bureau of Economic Research.

Grossman, Sanford J (1981) "The informational role of warranties and private disclosure about product quality," *The Journal of law and Economics*, 24 (3), 461–483.

Guttman, Ilan, Ilan Kremer, and Andrzej Skrzypacz (2014) "Not only what but also when: A theory of dynamic voluntary disclosure," *American Economic Review*, 104 (8), 2400–2420.

Hartzmark, Samuel M and Abigail B Sussman (2019) "Do investors value sustainability? A natural experiment examining ranking and fund flows," *The Journal of Finance*, 74 (6), 2789–2837.

Heyes, Anthony and Sandeep Kapur (2009) "An economic model of whistle-blower policy," *The Journal of Law, Economics, & Organization*, 25 (1), 157–182.

Hirshleifer, David, Sonya S Lim, and Siew Hong Teoh (2011) "Limited investor attention and stock market misreactions to accounting information," *The Review of Asset Pricing Studies*, 1 (1), 35–73.

Hong, Harrison and Marcin Kacperczyk (2009) "The price of sin: The effects of social norms on markets," *Journal of Financial Economics*, 93 (1), 15–36.

Howse, Robert and Ronald J Daniels (1995) "Rewarding whistleblowers: the costs and benefits of an incentive-based compliance strategy," *Corporate decision-making in Canada*, 4, 525–549.

Iwasaki, Masaki (2018) "Effects of external whistleblower rewards on internal reporting," *Harvard John M. Olin Fellow's Discussion Paper Series* (76).

Jung, Woon-Oh and Young K Kwon (1988) "Disclosure when the market is unsure of information endowment of managers," *Journal of Accounting research*, 146–153.

Kempf, Elisabeth, Alberto Manconi, and Oliver Spalt (2017) "Distracted shareholders and corporate actions," *The Review of Financial Studies*, 30 (5), 1660–1695.

Khorana, Ajay, Henri Servaes, and Peter Tufano (2009) "Mutual fund fees around the world," *The Review of Financial Studies*, 22 (3), 1279–1310.

Kim, Soohun and Aaron Yoon (2023) "Analyzing active fund managers' commitment to ESG: Evidence from the United Nations Principles for Responsible Investment," *Management Science*, 69 (2), 741–758.

Krueger, Philipp, Zacharias Sautner, and Laura T Starks (2020) "The importance of climate risks for institutional investors," *The Review of Financial Studies*, 33 (3), 1067–1111.

Langberg, Nisan and K Sivaramakrishnan (2008) "Voluntary disclosures and information production by analysts," *Journal of Accounting and Economics*, 46 (1), 78–100.

Liang, Hao, Lin Sun, and Melvyn Teo (2022) "Responsible hedge funds," *Review of Finance*, 26 (6), 1585–1633.

Ma, Linlin, Yuehua Tang, and Juan-Pedro Gomez (2019) "Portfolio manager compensation in the US mutual fund industry," *The Journal of Finance*, 74 (2), 587–638.

Manso, Gustavo (2011) "Motivating innovation," *The Journal of Finance*, 66 (5), 1823–1860.

Marinovic, Iván and Felipe Varas (2016) "No news is good news: Voluntary disclosure in the face of litigation," *The RAND Journal of Economics*, 47 (4), 822–856.

Milgrom, Paul R (1981) "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 380–391.

Nofsinger, John R and Richard W Sias (1999) "Herding and feedback trading by institutional and individual investors," *The Journal of Finance*, 54 (6), 2263–2295.

Nofsinger, John and Abhishek Varma (2014) "Socially responsible funds and market crises," *Journal of Banking & Finance*, 48, 180–193.

Nyreröd, Theo and Giancarlo Spagnolo (2021) "Myths and numbers on whistleblower rewards," *Regulation & Governance*, 15 (1), 82–97.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor (2015) "Scale and skill in active management," *Journal of Financial Economics*, 116 (1), 23–45.

——— (2021) "Sustainable investing in equilibrium," *Journal of Financial Economics*, 142 (2), 550–571.

Raghunandan, Aneesh and Shiva Rajgopal (2022) "Do ESG funds make stakeholder-friendly investments?" *Review of Accounting Studies*, 27 (3), 822–863.

Renneboog, Luc, Jenke Ter Horst, and Chendi Zhang (2008) "The price of ethics and stakeholder governance: The performance of socially responsible mutual funds," *Journal of Corporate Finance*, 14 (3), 302–322.

Revelli, Christophe and Jean-Laurent Viviani (2015) "Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis," *Business Ethics: A European Review*, 24 (2), 158–185.

Riedl, Arno and Paul Smeets (2017) "Why do investors hold socially responsible mutual funds?" *The Journal of Finance*, 72 (6), 2505–2550.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang (2023) "Firm-level climate change exposure," *The Journal of Finance*, 78 (3), 1449–1498.

Security and Exchange Commission (2021) "The Division of Examinations' Review of ESG Investing," https://www.sec.gov/files/esg-risk-alert.pdf, accessed May 2024.

Shanken, Jay (1992) "On the estimation of beta-pricing models," *The Review of Financial Studies*, 5 (1), 1–33.

Sirri, Erik R and Peter Tufano (1998) "Costly search and mutual fund flows," *The Journal of Finance*, 53 (5), 1589–1622.

Starks, Laura T (2023) "Presidential address: Sustainable finance and ESG issues—Value versus values," *The Journal of Finance*, 78 (4), 1837–1872.

Ting, Michael M (2008) "Whistleblowing," *American Political Science Review*, 102 (2), 249–267.

Wilde, Jaron H (2017) "The deterrent effect of employee whistleblowing on firms' financial misreporting and tax aggressiveness," *The Accounting Review*, 92 (5), 247–280.