# The Effects of ESG Ratings on Firms' Financial Decisions

**Sahand Davani** *

**Abstract**

This study examines the effects of ESG (Environmental, Social, and Governance) ratings on firms' financial decisions. Theoretical models predict that Socially Responsible Investing (SRI) positively impacts firms with high ESG performance by reducing the cost of capital for these firms, thereby providing them with better investment and growth opportunities. In this paper, I provide evidence that SRI can have positive impact on firms with high ESG performance through another channel as well. In a regression discontinuity design, I show that firms with higher ESG ratings (high-ESG firms) have higher ownership by ESG institutional investors, have lower perceived cost of equity capital, and consequently, issue more net equity than net debt compared to similar firms with lower ESG ratings (low-ESG firms). These results imply that access to cheaper equity financing acts as another impact mechanism of SRI. Consistently, I find that high-ESG firms try to maintain their high ESG ratings at the current levels, while the ESG ratings of similar low-ESG firms decline.

*ESADE Business School. sahand.davani@esade.edu

# 1 Introduction

Socially Responsible Investing (SRI) is an approach in which the investors consider Environmental, Social, and Governance (ESG) factors in their portfolio selection and management. In recent years, this approach has experienced considerable growth in popularity. One of the main strategies of SRI is divestment, in which responsible investors refuse to invest in companies with low ESG performance (low-ESG firms). Consequently, the demand for the shares of companies with high ESG performance (high-ESG firms) increases. The rise in demand pushes up the stock prices of high-ESG firms and, thus, reduces their cost of capital. This would provide high-ESG firms with more investment opportunities, and help them grow more.[1] Consequently, low-ESG firms would be forced to either improve their ESG performance to lower their cost of capital or risk losing the competition. This "investment" channel is the primary impact mechanism of divestment strategy in the literature (Heinkel et al., 2001; Pástor et al., 2021). To provide investors with information about the firms' ESG performance, a growing market of ESG rating agencies has formed. By reporting the firms' ESG ratings, these agencies can affect the financial decisions of responsible investors.

In a novel empirical setting, I investigate how the ESG ratings, through their impact on the investors' portfolio selection process, affect the financial decisions of the firms. Moreover, I provide evidence for the existence of a "financing" channel that can work as another impact mechanism of divestment strategy alongside the "investment" channel. This is an important finding considering the current debate regarding the effectiveness of the divestment strategy. For example, Berk and Van Binsbergen (2021) show that the reduction in the cost of capital resulting from the divestment strategy is too small to have a meaningful effect on firms' investment decisions. Consistent with this argument, while the recent empirical studies support the idea that divestment strategy reduces the cost of capital for high-ESG firms (Bolton and Kacperczyk, 2021; Pástor et al., 2022; Hsu et al., 2023; Pástor et al., 2022), there is no evidence that this lower cost

[1] Because of lower cost of capital, high-ESG firms would have access to more positive NPV projects than low-ESG firms.

of capital improves the investment opportunities of these firms. My results show that even if the divestment strategy does not provide high-ESG firms with better investment opportunities, it can still positively impact these firms by providing them with cheaper equity capital to finance their activities without necessarily providing them with more investment opportunities. Consequently, high-ESG firms are motivated to maintain their high ESG performance.

I employ a Regression Discontinuity (RD) design motivated by the methodology that MSCI uses to report its ESG ratings. Based on the assessment of publicly available data on several ESG key issues of the firms, MSCI assigns an industry-adjusted *numerical* ESG score in the range of 0 (lowest) to 10 (highest) to each firm. At two arbitrary cutoffs, MSCI divides this range into three parts, each corresponding to an ESG *label* that can be "Laggard," "Average," or "Leader." I focus on a small neighbourhood around the cutoff where the ESG label changes from "Average" to "Leader." The treatment is the change in the ESG label. I call the firms that are above the cutoff in the year *t*, the *treated firms*, and the firms that are below the cutoff, the *control firms*. I compare the outcome variables of these two groups of firms in the year *t+1*. My sample consists of the firm-year observations with MSCI ESG score in the period 2011 to 2019 in the US stock market.

The results of this study shed light on the effect of ESG ratings on firms' financial decisions. First, I investigate the differences in firms' ownership by ESG institutional investors between the treated and control firms. In order to identify ESG institutional investors, I calculate each institutional investor's ESG score as the value-weighted ESG score of its portfolio holdings (Cao et al., 2022; Hwang et al., 2022). Then, I divide institutional investors into five quantiles based on their ESG scores and define ESG institutional investors as those in the top quantile. My results show that in the year following the treatment, ESG institutional investors hold significantly higher ownership in treated firms than control firms. This result also implies that ESG institutional investors do consider firms' MSCI ESG labels in their financial decision-making.

Second, I study the perceived cost of capital difference between the treated and control firms. Perceived cost of capital is the firms' estimated financial cost of capital based on observed asset prices and interest rates that companies use to set the discount rate (or the required rate of return) for making investment decisions (Gormsen and Huber, 2023). If high-ESG firms believe they would have a lower cost of capital, this belief will be reflected in their perceived cost of capital. I compare treated and control firms' perceived cost of capital and show that in the year following the treatment, the perceived cost of capital of the treated firms is significantly lower than that of the control firms. To check the validity of this belief, I compare the actual return on equity of treated and control firms. The treated firms have significantly lower actual returns, which can be interpreted as a lower cost of equity than the control firms in the year following the treatment. Moreover, I do not find any difference in the cost of debt between the treated and control firms. These results indicate that treated firms correctly believe that their higher ESG ratings in the current year would reduce their cost of capital in the following year.

Third, I study the effect of ESG ratings on firms' financial decisions. I argue that since treated firms have a lower cost of equity than the control firms, they should prefer to issue more net equity. Analyzing the differences in financing preferences between the treated and control firms show that treated firms issue more net equity than net debt compared to control firms in the year following the treatment. However, I do *not* find any evidence that lower cost of capital provides the treated firms with better investment opportunities. Also, I do *not* find any evidence that the sum of net equity issuance and net debt issuance is different between treated and control firms. These results imply that treated firms can use the lower cost of equity to replace their debt with equity, which is now cheaper for them, without necessarily being able to make more investments. This result is in line with the findings in Berk and Van Binsbergen (2021), which argue that the impact of investing in high-ESG firms on their cost of capital is too small to affect their real investment decisions. Consistent with his claim, further analysis does not reveal any significant difference in the actual discount rate[2] between the treated and

[2]Actual discount rate (required rate of return) that is used in making investment decisions is equal to the perceived cost of capital plus a "wedge factor" that includes the effects of beliefs about value creation,

control firms.

Finally, I explore how ESG ratings affect firms' ESG performance. I show that following the treatment, the change in the MSCI ESG scores is less negative for treated firms compared to control firms. While for both groups of firms, the ESG score, on average, decreases in the year following the treatment, this reduction is smaller for the treated firms. It implies that treated firms try to maintain their ESG scores at the current levels to take advantage of the benefits, i.e., moderately lower cost of equity capital.

Employing RD design to investigate the effect of ESG ratings on financial markets and firms' behaviour can address the empirical challenge of studying this relationship. If the required conditions of the RD setting hold, the distribution of firms in a small neighbourhood around the cutoff, and consequently, the assignment of firms to the treated and control groups, can be considered as good as random (Lee and Lemieux, 2010). Therefore, on average, the treated and control firms are similar in every aspect other than their ESG labels. Because the cutoff is determined arbitrarily by MSCI, the difference between the labels of treated and control firms has no economic underpinning and is thoroughly exogenous. Therefore, any potential differences in the outcome variables between the treated and control firms are isolated from the possible effects of other (observable or unobservable) variables and have a causal interpretation.

However, the results of the RD analysis, while providing strong internal validity, are essentially local and should be interpreted in this sense. Thus, the findings of this study have low external validity by design. Considering this, these findings show that ESG ratings have causal effects on the financial markets and firms' behaviour, at least in a local neighbourhood around the cutoff. ESG institutional investors are more attracted to firms with better ESG labels and increase their ownership in these firms, compared to similar firms with lower ESG labels. The higher demand for firms with better ESG labels pushes up their share prices and reduces their actual returns, which can be interpreted as a lower cost of equity capital. These firms correctly lower their perceived cost of capital. Moreover, they issue more net equity than net debt, compared to similar

risk, and financial constraints on the firm's discount rate (Gormsen and Huber, 2023).

firms with lower ESG labels. However, the decrease in the cost of capital is not large enough to impact the real investment decisions of the high-ESG firms. Finally, these firms try to maintain their ESG scores at their current levels to resume benefiting from the lower cost of capital.

My results contribute to the literature on the impact of SRI on the financial markets and firms' behaviour. First, my results contribute to the empirical literature on the impact of ESG ratings on the financial markets and firms' behaviour. Hartzmark and Sussman (2019) provide strong evidence that the introduction of Morningstar mutual fund's ESG ranking resulted in a significant net inflow for high sustainability mutual funds, and a net outflow for low sustainability ones. Rzeźnik et al. (2022) find that a change in the Sustainalytics ESG rating methodology affects the firms' monthly returns. Glück et al. (2021) show that downgrades in MSCI Environmental and Social scores are followed by negative abnormal returns. In contrast, upgrades in the MSCI Environmental and Governance scores lead to lower downside risks and systematic risks, respectively. My results are closely related to those of Berg et al. (2022), who show that downgrades in the MSCI ESG letter ratings reduce the ownership by ESG mutual funds, while upgrades increase it. Moreover, they document a negative abnormal return following ESG rating downgrades. While they do not find any significant effect of changes in MSCI ESG letter rating on firms' capital expenditure, they find a positive change in the Governance Pillar score following MSCI ESG letter downgrades and a negative change following upgrades. While they focus on the effect of changes in MSCI ESG letter rating in a panel event study setting, I focus on the effect of the ESG labels in an RD setting. I show that everything else being equal, firms with better ESG labels have higher ownership by ESG institutional investors and lower cost of equity, which helps them finance their current activities by issuing more net equity than net debt. However, the reduction in the cost of equity is not large enough to affect their real investment decisions. Moreover, high-ESG firms maintain their high ESG ratings in the following year to further take advantage of the lower cost of equity. Importantly, these effects are independent of differences in the firms' actual ESG performance or other fundamental characteristics. Therefore, ESG rating agencies can affect the financial markets and firms' behaviour.

Second, I add to the literature on the optimal SRI strategy. There are two main investment strategies through which SRI can reduce the firms' negative externalities: 1) Engagement (also known as voice), that is, investing in firms with negative externalities and trying to reduce the externalities through active ownership (Dimson et al., 2015; Broccardo et al., 2022; Krueger et al., 2020) and 2) Divestment (also known as exit), that is, refusing to invest in firms with negative externalities and investing in firms with positive externalities. This study is related to the ongoing debate regarding the impact mechanism and the effectiveness of the divestment strategy. Some theoretical models of divestment strategy show that investors' preferences for green holdings would push up the prices of high-ESG firms through higher demand, thereby reducing their cost of capital compared to low-ESG firms. This would create a positive social impact by inducing firms to improve their ESG performance and by shifting real investments towards high-ESG firms (Heinkel et al., 2001; Pástor et al., 2021). Other theoretical works, however, show that the effect of divestment strategy on firms' cost of capital depends on several other factors and can be positive, negative, or neutral. For example, Pedersen et al. (2021) show that depending on the relative weight of different investor types in the market, the relationship between the firm's ESG performance and investors' expected returns can be positive, negative, or neutral. Berk and Van Binsbergen (2021) develop a formula to approximate the change in the cost of capital resulting from divestment strategy. They calibrate the formula with the current data and show that the change in the cost of capital is too small to have a meaningful effect on firms' real investment decisions. Moreover, Goldstein et al. (2022) find that when investors have heterogeneous preferences over the financial and ESG performance of the firms, a higher fraction of ESG investors can raise the information risk regarding the firms' financial payoff, driving up their cost of capital. The empirical evidence on the relationship between ESG performance and the financial performance of the firms is also mixed. While El Ghoul et al. (2011), Chava (2014), Bolton and Kacperczyk (2021), Pástor et al. (2022), Hsu et al. (2023), and Gormsen et al. (2023) find negative relationship between firms' ESG performance and their cost of equity, Edmans (2011) and Krüger (2015) find a positive relationship. More recently, Pástor et al. (2022) show that green stocks outperform brown stocks; however, they attribute this outperformance to an ex-post realization due to climate adverse events and not the higher expected returns. Even if divestment lowers the cost of capital, there is no solid empirical evidence that high-ESG firms can benefit from it by having better investment opportunities. I add to this literature by providing evidence that a divestment strategy benefits high-ESG firms by giving them access to cheaper equity financing without necessarily providing them with better investment opportunities. I also show that this "financing" channel can motivate high-ESG firms to keep their ESG ratings at the current levels while the control firms, on average, let their ESG scores decline.

Finally, This paper contributes to the recent empirical literature exploring the effect of investors' ESG preferences on the firms' behaviour in the real side of the economy. Theoretically, Fama and French (2007) provide a rationale on how investors' tastes for certain assets deviate asset prices from the classical asset pricing models, like CAPM. Moreover, Bond et al. (2012) discuss how the feedback from the market prices affects the firms' decisions in the real economy. Therefore, the investors' ESG preferences could affect firms' behaviour by impacting asset prices. The empirical literature, however, has not yet found compelling evidence supporting this argument. Briere and Ramelli (2022) show that an upward shift in the investors' ESG preferences increases environmentally friendly firms' capital investments and cash holdings. Gantchev et al. (2022) argue that following adverse environmental and social incidents, firms improve their environmental and social policies, and the magnitude of this response depends on the firms' shareholders' ESG preferences. However, they do not find any effects on the firms' governance practices. Heath et al. (2023) find that while SRI mutual funds do select firms with better Environmental and Social (E&S) conduct, they do not improve the E&S performance of their portfolio firms. My findings show that investors' ESG preferences motivate firms to pay attention and try to maintain their ESG performance at high levels to enjoy a lower cost of capital. Moreover, I find that these preferences affect firms' financing decisions by inducing high-ESG firms to issue more net equity than net debt.

# 2 Data and Variables

## 2.1 Data and Sample

In this study, I use MSCI ESG rating to investigate the effect of ESG ratings on financial markets and firms' behaviour. MSCI ESG rating is one of the most widely used ESG ratings in the market. This rating evaluates firms' management of financially important ESG risks and opportunities. In coming up with this rating, MSCI takes into account the materiality of ESG risks and the capability of each firm to manage those risks (MSCI, 2023). Each firm is evaluated on a collection of two to seven Environmental and Social "Key Issues," which are selected from a total of 33 Key Issues based on the firm's industry and market factors (Key Issue Exposure Score). Then, the firm's capability to manage their aggregate ESG risks and opportunities is evaluated based on its governance structure, policies and targets, quantitative performance measures, and relevant controversies (Key Issue Management Score). Based on these scores, each company receives a 0-10 score for each selected Environmental and Social Key Issue (Key Issue Scores). Moreover, all the companies are evaluated on all the Key Issues in the Governance pillar, and each firm receives a 0-10 score (Governance Pillar Score). Then, a Weighted Average Key Issue Score (WAKIS) is calculated for each firm based on the Key Issue Scores and the Governance Pillar Score. The WAKIS is then normalized relative to the ESG ratings of other firms in the same industry to calculate the Industry-Adjusted Company Score, which is in the range 0-10. Finally, this range is divided into 7 equal parts, and each firm receives an ESG letter rating from CCC (lowest) to AAA (highest) (see Table 1). Moreover, MSCI labels firms with ESG labels and letter ratings of AA and AAA as "Leader," BB, BBB, and A as "Average," and CCC and B

**Table 1: The MSCI method of mapping the Final Industry-Adjusted Company ESG Score to ESG Letter Rating and ESG Label**

| Final Industry-Adjusted Company ESG Score | ESG Letter Rating | ESG Label |
|---|---|---|
| 8.571 - 10.000 | AAA | Leader |
| 7.143 - 8.571 | AA | Leader |
| 5.714 - 7.143 | A | Average |
| 4.286 - 5.714 | BBB | Average |
| 2.857 - 4.286 | BB | Average |
| 1.429 - 2.857 | B | Laggard |
| 0.000 - 1.429 | CCC | Laggard |

This table shows how MSCI maps the final industry-adjusted company ESG score to ESG letter rating and ESG label. The 0-10 numerical scale is divided into seven equal parts, and each part is assigned an ESG letter rating, changing from CCC to AAA. Moreover, at the ESG score of 2.857, where the letter rating changes from B to BB, the ESG label changes from Laggard to Average; similarly, at the ESG score of 7.143, where the ESG letter rating changes from A to AA, the ESG label changes from Average to Leader. Also note that the overlap in the score ranges is because of the rounding error (MSCI, 2023).

as "Laggard."[3] These letter ratings, along with other tools measuring climate risks and opportunities that the firm may face, are publicly available in the MSCI's website[4]. For the purpose of this study, it is important to note that in this process, MSCI uses only publicly available data from sources like company financial and sustainability disclosures, specialized government and academic data sets, media searches, etc. Moreover, MSCI publicly declares on its website that it does not send surveys or conduct interviews with firms, and they do not accept any data provided by firms that is not publicly available. Therefore, firms do not have the ability to manipulate their ratings, or appeal their assigned ratings.

For this study, I use the MSCI ESG ratings for several reasons. First, MSCI is one of the largest providers of ESG ratings in the market. These ratings are widely used by academics and practitioners. Second, MSCI uses only publicly available data to evaluate firms, and firms can *not* manipulate their ratings. Finally, having both an ESG numerical score which changes continuously, and an ESG letter rating that changes discontinuously at the arbitrarily chosen cutoffs make MSCI ESG rating system suitable to be employed in the specific Regression Discontinuity design of this study. I combine this dataset with the institutional ownership data from the Thompson Reuters 13F

[3]For more details, see https://www.msci.com/esg-and-climate-methodologies.

[4]https://www.msci.com/our-solutions/esg-investing/esg-ratings-climate-search-tool

database and the firms' financial data from Compustat. I also use the data for firms' perceived cost of capital and discount rates from Gormsen and Huber (2023). I have removed the companies in the finance sector (SIC Industry Codes 6000 to 6999), and winsorized all the variables at the 1st and 99th percentiles to remove the effect of outliers. The final sample in this study has 13,268 unique firm-year observations for the period of 2011 to 2019, for 3186 unique firms in the US stock market.

## 2.2 Outcome Variables

### 2.2.1 Firms' Ownership by ESG Institutional Investors

In order to calculate the ownership by ESG institutional investors, I first define ESG institutional investors. Several methods have been employed in the literature to identify these investors. For example, Berg et al. (2022) identify ESG mutual funds by screening their names and strategies for keywords related to ESG, like "SRI," "social," "ESG," "green," etc. However, as Dumitrescu et al. (2022) show, 24% of self-labeled ESG funds are "greenwashers," that is, they do not necessarily follow their commitment to SRI in practice. Another approach, used by Cao et al. (2022) and Hwang et al. (2022), is to calculate each investor's ESG score as the value-weighted ESG score of its portfolio holdings. This approach has the advantage of identifying investors that are committed to SRI in practice, without necessarily stating it in their names and strategies. Moreover, consistent with the focus of this study, the latter approach focuses on investors that use divestment (and not engagement) strategy by investing in high-ESG firms and divesting in low-ESG ones. I use the second approach to assign an ESG score to each fund at the end of each quarter. For each fund, I take the average of these quarterly ESG scores to come up with its yearly ESG score. Then, in each year, I divide funds into five quantiles, and define ESG funds in that year as those that are in the top quantile. I calculate, for each firm in each year, the sum of shares owned by ESG funds to create *ESG_Ownership* variable.

### 2.2.2 Firms' Perceived Cost of Capital

According to the stylized view in economics, firms should invest in any project that offers returns above a threshold, that is the firm's discount rate. In theory, this discount rate is equal to firm's financial cost of capital. However, in reality, estimating this variable is complicated and not straightforward. Therefore, many firms estimate their financial cost of capital based on observed asset prices and interest rates. Gormsen and Huber (2023) call this estimate the firm's perceived cost of capital, and explain that in order to come up with the final discount rate to be used in their investment decisions, firms also include the effect of other factors, like beliefs about value creation, risk, and financial constraints in their perceived cost of capital. Therefore, the firm's actual discount rate that is used as the threshold in making their investment decision is usually different from the firm's perceived cost of capital. Gormsen and Huber (2023) measure firms' perceived cost of capital and their actual discount rate by using the information shared in the corporate conference calls, during which managers inform the public about their firms' operations. In these calls, managers sometimes share their estimate of their perceived cost of capital and discount rates. By collecting and analyzing the transcripts of these conference calls, Gormsen and Huber (2023) have created a database containing the perceived cost of capital and discount rates used by around 2500 firms across 20 countries in the period 2002 to 2021[5]. I use this database to create *Perceived_CoC* and *Discount_Rate* variables.

### 2.2.3 Difference between Firms' Net Debt Issuance and Net Equity Issuance

In order to measure the firms' financing preferences, I use the difference between a firm's *Net Debt Issuance* and its *Net Equity Issuance*, divided by its one-year lagged *Total Assets*. The reason for choosing this variable is that it directly measures the firms' short-term preferences regarding their capital structure choices, in contrast to *Leverage* which is affected by other variables, like *Earnings* (Kisgen, 2019), and is shown to be a persistent characteristic, that is, it changes very slowly over time (Lemmon et al., 2008).

[5] Publicly available at: https://costofcapital.org/

Kisgen (2006) defines *Net Debt Issuance* as debt issuances minus debt reductions (long-term debt issuance minus long-term debt reduction plus changes in current debt), and *Net Equity Issuance* as equity issuances minus share repurchases (sale of common and preferred stock minus purchases of common and preferred stock). I define the variable *NetD-NetE* as:

$$NetD - NetE_{i,t} = \frac{NetDebtIssuance_{i,t} - NetEquityIssuance_{i,t}}{Assets_{i,t-1}} \tag{1}$$

#### 2.2.4 Change in the Firms' ESG Score

This variable measures how the firms' ESG performance varies in each year. I define the change in the firms' MSCI ESG score (*dMSCI_Score*) as:

$$dMSCI\_Score_{i,t} = MSCI\_Score_{i,t} - MSCI\_Score_{i,t-1}$$

### 2.3 Data Description and Summary Statistics

Figure 1 shows the distribution of firm-year observations based on their MSCI ESG letter ratings through the sample period. This figure demonstrates that a majority of firm-year observations are in the middle of the distribution, that is, more than 50 percent of the observations are in the letter ratings of "BB" and "BBB." This implies that firms are not able to manipulate their ratings. If firms could manipulate their ratings, we would have observed majority of firm-year observations in the higher letter ratings. Figure 2 further confirms this argument. The distribution of observations in each year has remained roughly constant during the sample period. In all years, more than 50 percent of firms are in the "BB" and "BBB" categories, implying that firms are not able to precisely manipulate their ratings. As explained in Section 3.2, this is an important point that justifies the validation of the empirical strategy employed in this study.

Table 2 shows the summary statistics of the variables used in this study, separated by their MSCI ESG label. This cursory look at this table shows that there are very small differences among the mean of different variables in the three label categories. This observation provides preliminary evidence that the distribution of firms in the ESG labels is not dependent on fundamental characteristics of the firms. I will provide more

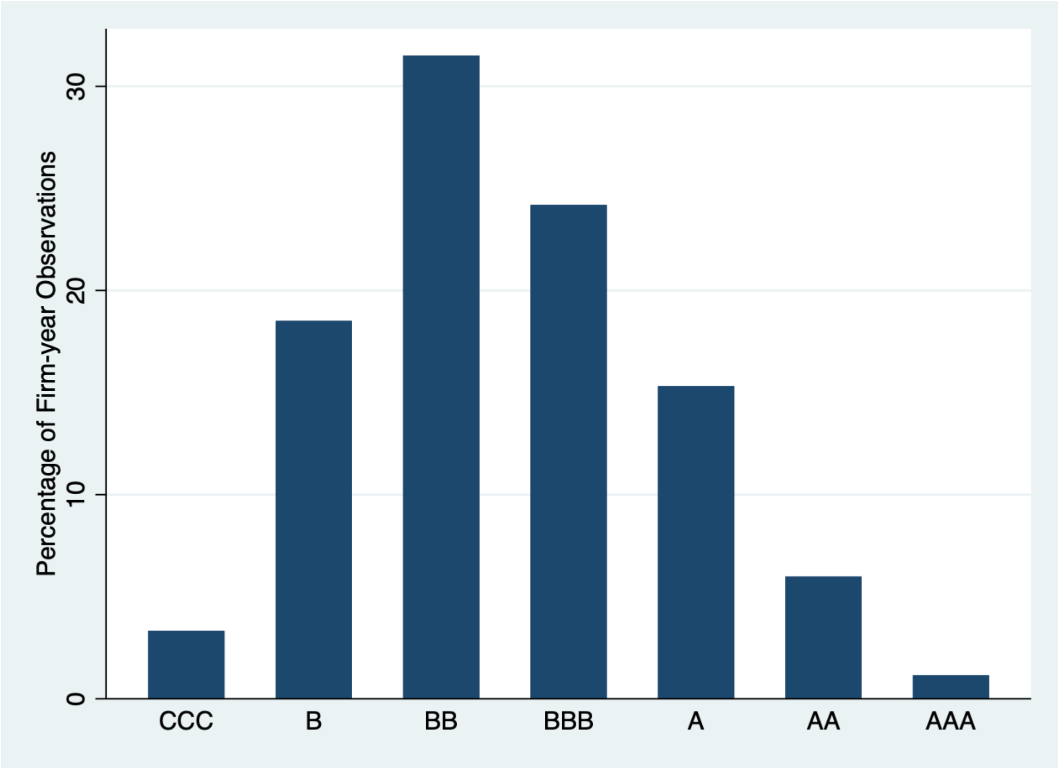



**Figure 1: Distribution of Firm-year Observations**

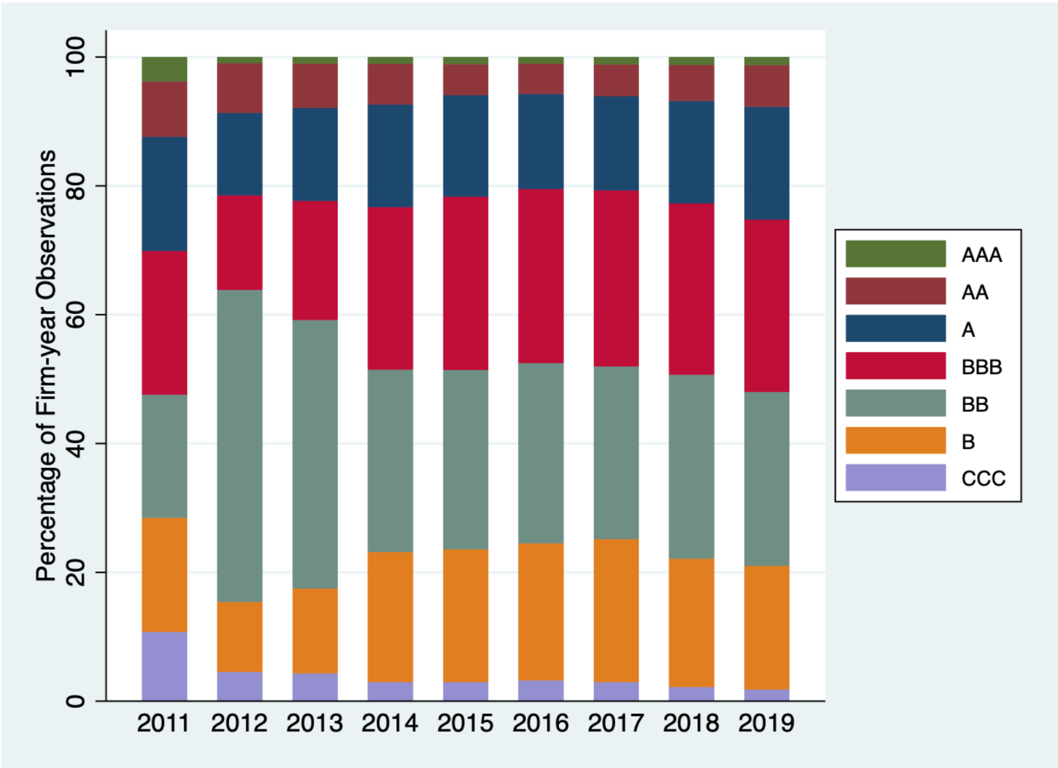



**Figure 2: Distribution of Firm-year Observations in Each Year**

evidence for this argument in Section 3.2.

# 3 Methodology

## 3.1 Empirical strategy

I use a Regression Discontinuity (RD) design to estimate the causal effect of MSCI ESG ratings on the outcome variables. In an RD design, a treatment is assigned to units whose running variables exceed a known cutoff point. If units can not precisely manipulate the running variable, then the variation in the treatment for the units in a small enough neighborhood around the cutoff is as good as random, as though generated from a randomized experiment. Therefore, the effect of treatment on the outcome variables can be analyzed and tested like a randomized experiment, because the units that are above the cutoff (treated units) and those that are below the cutoff (control units) are, on average, similar in every aspect, except for their treatment status (Imbens and Lemieux, 2008; Lee and Lemieux, 2010; Cattaneo and Titiunik, 2022).

In this study, the running variable is the firms' numerical ESG score in the year $t$, and the outcome variables are those introduced in the Section 2.2, in the year $t+1$. The cutoff is the point in the ESG score where the ESG label changes from "Average" to "Leader" (or equivalently, where the ESG letter rating changes from "A" to "AA"), i. e., at the ESG numerical score of 7.143. The treatment is the change in the ESG label (or equivalently, the change in the ESG letter rating). The firms that are above the cutoff are the *treated firms*, and those below the cutoff are the *control firms*. Since the cutoff is determined arbitrarily by dividing the ESG numerical score scale into seven equal parts, and, as will be argued later, firms can not precisely manipulate their running variable (ESG numerical score), the treatment can be considered a randomly assigned variable with regard to firms' characteristics. Therefore, the distribution of firms around the cutoff is as good as random, and any observed differences in the outcome variables between the treated and control firms are results of the treatment status. Moreover, as will be shown later, the treated and control firms are similar on their observable

## Table 2: The Summary Statistics

| | Mean | SD | Min | p25 | p50 | p75 | Max | N |
|---|---|---|---|---|---|---|---|---|
| **Panel A - Firms with "Laggard" MSCI ESG Label** | | | | | | | | |
| CashFlow / 1-y Lagged Assets | .073 | .14 | -.68 | .05 | .091 | .14 | .36 | 2,390 |
| Dividend / 1-y Lagged Assets | .017 | .028 | 0 | 0 | .0054 | .022 | .17 | 2,387 |
| Cash / 1-y Lagged Assets | .13 | .15 | .00061 | .026 | .082 | .18 | .87 | 2,367 |
| CAPEX / 1-y Lagged Assets | .059 | .059 | .00078 | .02 | .04 | .073 | .29 | 2,390 |
| Leverage | .85 | 3 | -15 | .11 | .56 | 1.2 | 16 | 2,890 |
| Return on Assets | .055 | .15 | -.67 | .03 | .07 | .12 | .34 | 2,895 |
| Market to Book Ratio | 1.6 | 1.6 | .12 | .58 | 1.1 | 2 | 9.2 | 2,893 |
| Return on Equity | .01 | .41 | -1.3 | -.18 | .06 | .26 | .96 | 2,408 |
| Cost Of Debt | .065 | .09 | .00047 | .038 | .051 | .07 | 1 | 2,368 |
| dMSCI_Score | -.23 | .92 | -5.3 | -.7 | 0 | .2 | 2.8 | 2,394 |
| ESG_Ownership | .54 | .2 | .038 | .43 | .58 | .68 | .95 | 2,551 |
| Perceived_CoC | 9.4 | .96 | 7.3 | 8.8 | 9.4 | 10 | 12 | 2,404 |
| NetD-NetE | .034 | .18 | -.85 | -.016 | .012 | .074 | .77 | 2,393 |
| **Panel B - Firms with "Average" MSCI ESG Label** | | | | | | | | |
| CashFlow / 1-y Lagged Assets | .06 | .16 | -.68 | .045 | .086 | .13 | .36 | 7,536 |
| Dividend / 1-y Lagged Assets | .016 | .027 | 0 | 0 | .0025 | .022 | .17 | 7,529 |
| Cash / 1-y Lagged Assets | .14 | .16 | .00061 | .033 | .089 | .19 | .87 | 7,435 |
| CAPEX / 1-y Lagged Assets | .049 | .049 | .00078 | .017 | .033 | .063 | .29 | 7,539 |
| Leverage | .79 | 2.9 | -15 | .06 | .49 | 1.1 | 16 | 9,376 |
| Return on Assets | .049 | .16 | -.67 | .034 | .074 | .12 | .34 | 9,423 |
| Market to Book Ratio | 1.8 | 1.7 | .12 | .7 | 1.2 | 2.2 | 9.2 | 9,401 |
| Return on Equity | .044 | .38 | -1.3 | -.14 | .081 | .27 | .96 | 7,574 |
| Cost Of Debt | .071 | .12 | .00047 | .035 | .049 | .067 | 1 | 7,351 |
| dMSCI_Score | .13 | 1 | -5.3 | -.3 | 0 | .6 | 6.2 | 7,543 |
| ESG_Ownership | .59 | .19 | .038 | .49 | .63 | .73 | .95 | 8,029 |
| Perceived_CoC | 9.5 | 1 | 7.3 | 8.8 | 9.5 | 10 | 12 | 8,046 |
| NetD-NetE | .034 | .19 | -.85 | -.017 | .014 | .082 | .77 | 7,543 |
| **Panel C - Firms with "Leader" MSCI ESG Label** | | | | | | | | |
| CashFlow / 1-y Lagged Assets | .1 | .091 | -.49 | .063 | .1 | .14 | .36 | 782 |
| Dividend / 1-y Lagged Assets | .025 | .03 | 0 | 0 | .018 | .038 | .17 | 782 |
| Cash / 1-y Lagged Assets | .12 | .13 | .00061 | .032 | .077 | .16 | .87 | 763 |
| CAPEX / 1-y Lagged Assets | .043 | .037 | .0013 | .02 | .034 | .054 | .29 | 782 |
| Leverage | .91 | 3 | -15 | .2 | .58 | 1.3 | 16 | 944 |
| Return on Assets | .098 | .088 | -.58 | .053 | .097 | .15 | .34 | 947 |
| Market to Book Ratio | 1.8 | 1.6 | .12 | .75 | 1.3 | 2.2 | 9.2 | 940 |
| Return on Equity | .067 | .31 | -1.3 | -.077 | .11 | .26 | .96 | 782 |
| Cost Of Debt | .064 | .13 | .00047 | .031 | .043 | .058 | 1 | 781 |
| dMSCI_Score | .31 | .99 | -2.8 | -.2 | 0 | .8 | 5.5 | 782 |
| ESG_Ownership | .65 | .17 | .038 | .57 | .67 | .76 | .95 | 763 |
| Perceived_CoC | 9.1 | .96 | 7.3 | 8.4 | 9.1 | 9.7 | 12 | 855 |
| NetD-NetE | .063 | .12 | -.39 | -.0025 | .034 | .096 | .77 | 782 |
| **Panel D - All Firms** | | | | | | | | |
| CashFlow / 1-y Lagged Assets | .066 | .15 | -.68 | .048 | .088 | .13 | .36 | 10,708 |
| Dividend / 1-y Lagged Assets | .017 | .028 | 0 | 0 | .005 | .023 | .17 | 10,698 |
| Cash / 1-y Lagged Assets | .14 | .16 | .00061 | .031 | .087 | .18 | .87 | 10,565 |
| CAPEX / 1-y Lagged Assets | .051 | .051 | .00078 | .018 | .035 | .065 | .29 | 10,711 |
| Leverage | .81 | 2.9 | -15 | .078 | .51 | 1.1 | 16 | 13,210 |
| Return on Assets | .054 | .15 | -.67 | .035 | .075 | .12 | .34 | 13,265 |
| Market to Book Ratio | 1.8 | 1.7 | .12 | .67 | 1.2 | 2.2 | 9.2 | 13,234 |
| Return on Equity | .038 | .38 | -1.3 | -.15 | .079 | .27 | .96 | 10,764 |
| Cost Of Debt | .069 | .11 | .00047 | .035 | .049 | .067 | 1 | 10,500 |
| dMSCI_Score | .064 | 1 | -5.3 | -.4 | 0 | .5 | 6.2 | 10,719 |
| ESG_Ownership | .59 | .2 | .038 | .48 | .62 | .72 | .95 | 11,343 |
| Perceived_CoC | 9.5 | 1 | 7.3 | 8.8 | 9.5 | 10 | 12 | 11,305 |
| NetD-NetE | .036 | .19 | -.85 | -.016 | .015 | .081 | .77 | 10,718 |

This table shows the summary statistics for the firm-level characteristics (yearly observations) used in this study. Panel A reports the summary statistics for firms with the MSCI ESG Label of "Laggard," while Panel B and Panel C report the statistics for firms with MSCI ESG Labels of "Average," and "Leader," respectively. Panel D shows the statistics for all the firms.

fundamental characteristics, therefore, the differences in their outcome variables are the results of different ESG labels, independent of the firms' financial performance. Therefore, this framework enables me to capture the causal effect of ESG ratings on financial markets and firms' behaviour.

For example, consider two firms with very similar ESG scores. One has an ESG score of 7.2 and the other has an ESG score of 7.1. These two firms have very similar ESG performance. However, the first firm is assigned an ESG label of "Leader" (treated), because its ESG score exceeds the cutoff point of 7.143; the second firm is assigned an ESG label of "Average" (control), because its ESG score falls below the cutoff. Since the firms cannot "decide" what their ESG scores would be, it can be argued that their assignment to the treated or control group is as good as random. If these two firms are similar on every other characteristics, then any observed differences between their outcome variables in the subsequent year is a causal effect of their treatment status.

I employ a continuity-based RD framework to estimate the effect of ESG ratings on the outcome variables. In this framework, (local) polynomial functions are used to approximate the regression functions on each side of the cutoff separately. This method is by now the standard framework for RD empirical analysis (Cattaneo et al., 2019). In this approach, only observations that have a running variable between $c-h$ and $c+h$, where $c$ is the cutoff and $h(>0)$ is the bandwidth, enter the RD analysis. The local regression model can be specified as

$$Y_{it+1} = \alpha + \tau D_{it} + f_b^p(X_{it} - c) + D_{it} f_a^p(X_{it} - c) + \varepsilon_{it} \tag{2}$$

where $Y_{it+1}$ is the outcome variable for firm $i$ in the year $t+1$, $X_{it}$ is the numerical ESG score (the running variable) for firm $i$ in the yeat $t$, and $D_{it}$ is a dummy variable that equals 1 if the firm $i$ was above the cutoff in year $t$, and zero otherwise. $f_b^p$ and $f_a^p$ are the fitted polynomials of order $p$ below and above the cutoff, respectively, and $c$ is the cutoff point (7.143) in the ESG score, where the ESG label changes from Average to Leader. The coefficient $\tau$ captures the difference in the intercepts of $f_b^p$ and $f_a^p$ at the cutoff, and therefore, shows the difference in the outcome variable between treated and control firms, or equivalently, an estimate of the causal effect of ESG ratings on the outcome variables. Note that since firms perfectly comply with the treatment (all firms that lie above the cutoff are labeled Leader, and all firms below the cutoff are labeled Average), I use a Sharp RD analysis. Moreover, since the distribution of firms around the cutoff is as good as random, no other control variables need to enter the model.

Several issues arise in estimating the regression Equation (2). The first issue is the functional form of the polynomials $f_b^p$ and $f_a^p$ that can affect the treatment effect estimator. While using higher orders generally improves the accuracy of estimation, it also increases the variability of the treatment effect estimator and may lead to overfitting of the data near the cutoff. Therefore, I report the results of the analysis using only linear and quadratic polynomials, and ignore higher order polynomials.

The second issue is the bandwidth around the cutoff that is used to estimate the treatment effect. Choosing a small bandwidth tends to reduce the misspecification error in approximating the polynomial to fit the data below and above the cutoff. However, by using a small bandwidth, fewer observations will be used in the estimation and thus, the variance of the estimated coefficients increases. Therefore, there is a trade off between bias and variance in choosing the appropriate bandwidth. In this regard, since the results of the analysis are highly sensitive to the choice of the bandwidth, using a data-driven approach to choose the optimal bandwidth prevents specification searching and arbitrary decisions. The most popular data-driven approach optimizes the bias-variance trade-off by minimizing the Mean Square Error (MSE) of the local polynomial estimator (Imbens and Kalyanaraman, 2012; Cattaneo et al., 2020). While MSE-optimal bandwidth is appropriate for the estimation purposes, Calonico et al. (2016) and Calonico et al. (2018) show that it is not necessarily optimal for constructing confidence intervals for inference purposes. Instead, they recommend using a different, smaller bandwidth that minimizes the coverage error (CE) probability. It is important to note that both MSE- and CE-optimal bandwidths are sensitive to the total sample size, the order of the polynomial used for estimation, kernel function, variables, and several other factors. Therefore, in order to be consistent and avoid using arbitrary bandwidths, I report the results of the empirical analysis using both MSE- and CE-optimal bandwidths. In Section 5 I check the robustness of my results to choosing different bandwidths around the MSE-optimal one.

The third issue is the choice of the kernel function that is used to assign different weights to each observation based on the distance between the observation's running variable and the cutoff. The intuition behind using different weights is that observations that are closer to the cutoff can be more important in estimating the coefficients compared to farther observations, and thus, the triangular kernel that assigns higher weights to closer observations may be more appropriate. Once again, in order to be consistent in reporting the results and avoid subjective decisions in this regard, I report the results of the empirical analysis using both a triangular kernel and a uniform one.

The last issue is the method for estimating the standard errors of the analysis. The point here is that the MSE- and CE-optimal bandwidths are defined to be optimal for estimation, and are not necessarily appropriate for building confidence intervals and making inferences (Calonico et al., 2020). Therefore, Cattaneo et al. (2019) recommend using different bandwidths for estimation and inference purposes, which result in both a valid point estimation of the coefficient, and a valid robust bias-corrected confidence interval[6].

## 3.2 Validation Tests

One of the main requirements of the RD identification strategy is that the distribution of firm-year observations below and above the cutoff should be as good as random with respect to the firms' other characteristics. If this requirement is satisfied, then any observed difference in the outcome variable between the treated and control firms can be causally attributed to the isolated effect of differences in their ESG label. This requirement would be violated if the firms had the capability to precisely manipulate their ESG ratings near the cutoff (Lee and Lemieux, 2010). A major advantage of RD

[6]To perform the empirical analysis, I have used the RD Packages introduced by Calonico et al. (2014) and Calonico et al. (2017), available at https://rdpackages.github.io/

framework is that there are straightforward validation tests to check for the random distribution assumption. In this section, I show that there are no evidence for rejecting the hypothesis that treated and control firms are similar on every predetermined observable characteristics, and therefore, it is very improbable that some firms could systematically manipulate their ratings around the cutoff.

One of the most important RD validation tests involves checking if near the cutoff, treated and control firms are similar in terms of their observable characteristics prior to the treatment. If this requirement holds, it can be argued that the assignment of treatment to firms around the cutoff is as good as random, and the observed differences in the outcome variable following the treatment is independent of the pre-existing differences between the two groups of firms. To test this, I replace $Y_{it+1}$ with $Y_{it}$ in regression Equation (2) and estimate it by using different observable characteristics as the outcome variables. The results of this analysis are reported in Table 3. The results show that there are no systematic differences in the observable characteristics, near the cutoff, between the treated and control firms prior to the treatment. Overall, I find no evidence to reject the assumption that the distribution of firms around the cutoff is as good as random and firms can not precisely manipulate their ratings around the cutoff. Any possible effect of treatment, therefore, is independent of pre-existing differences between the treated and control firms, and can be interpreted as the causal effect of ESG ratings on firms' behaviour.

Another important method to check the validity of the RD results is to examine the uniform distribution of observations around the cutoff. The idea is that if firms do not have the ability to precisely manipulate their ESG ratings, the number of observations just below the cutoff should not be significantly different from the number of observations just above the cutoff (McCrary, 2008). However, performing this test is not reliable in the specific sample of this study for several reasons. First, in this study, the MSE-optimal bandwidth to infer reliable results from this test (Cattaneo et al., 2020) is so large ($h_{MSE} = 2.3$ in the 0-10 MSCI ESG score scale) that it includes several other cutoffs in which the ESG letter rating changes. As shown in Figure 1, the distribution of

## Table 3: RD Validity Test for Lack of Systematic Differences between Treated and Control Firms Prior to the Treatment

| Variable | (1) Coefficient | (2) t-stat | (3) MSE-opt bw | (4) Eff. Obs. | |
|---|---|---|---|---|---|
| | | | | below | above |
| CashFlow / 1-y Lagged Assets | -0.006 | -0.186 | 1.072 | 1195 | 452 |
| Dividends / 1-y Lagged Assets | 0.004 | 1.279 | 0.898 | 973 | 365 |
| Cash / 1-y Lagged Assets | 0.041 | 1.454 | 1.198 | 1270 | 493 |
| CAPEX / 1-y Lagged Assets | 0.001 | 0.1620 | 1.264 | 1415 | 552 |
| Leverage | -0.373 | -0.964 | 0.874 | 1194 | 434 |
| Return On Assets | 0.002 | -0.115 | 1.471 | 2029 | 797 |
| Return on Equity | -0.058 | -1.194 | 1.060 | 1202 | 452 |
| Market to Book Value | -0.264 | -0.498 | 0.808 | 1195 | 434 |
| Cost of Debt | 0.001 | -0.074 | 1.260 | 1495 | 574 |

This table shows the results of the RD validity test for the lack of difference in observable characteristics between the treated and control firms prior to the treatment. I estimate Equation (2) after replacing $Y_{it+1}$ with $Y_{it}$, and use different firm characteristics as the outcome variable. The coefficient $\tau$ is reported in Column (1). I follow the recommendations of Cattaneo et al. (2019) to use the MSE-optimal bandwidth (Column (3)), and robust bias corrected standard errors to calculate t-statistics (Column (2)). Note that the optimal bandwidth is different when different outcome variables are used. The effective number of observations, below and above the cutoff, is reported in Column (4). I have used triangular kernel and polynomials of order 2 to estimate the results. Unreported analysis using different bandwidths, a uniform kernel, and order 1 polynomials show no significant difference from these results. As this table shows, there are no significant differences in the observable characteristics, near the cutoff, between the treated and control firms *prior* to the treatment. Therefore, there is no evidence to reject the assumption that the distribution of firms around the cutoff is as good as random. (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively) firms in different ESG letter ratings is not uniform, and therefore, the effective number of firm-year observations to run this test would be around 4000 below the cutoff and 800 above the cutoff. Obviously, it cannot be expected from the uniform distribution test with the MSE-optimal bandwidth to yield reliable results in this setting. Another reason is that MSCI ESG rating methodology has changed several times (sometimes more than once in a single year) during the sample period of this study (MSCI, 2023). Therefore, aggregating all the firm-year observations and running the uniform distribution test on them would compare the outcomes of different rating methodologies, and thus, is not reliable. Moreover, MSCI uses industry-adjusted company ESG scores in its rating. These scores are normalized based on the ESG scores of the focal firm's peers in the same industry. Therefore, aggregating firm-year observations which are normalized based on different benchmarks, and based on different procedures in each year, would not yield a reliable sample for performing the uniform distribution test, because, each group of firm-year observations (i. e., those in different industries and in different years) is distributed differently. One way to bypass this issue is to run the uniform distribution test for the observations in each industry and in each year separately. Doing so, however, reduces the number of effective observations that enter the test to an extent that makes the results of the analysis unreliable.

In summary, the purpose of validation tests is to examine whether units have the ability to precisely manipulate their running variable or not. Table 3 shows that there are no systematic differences between the treated and control firms. If some firms have the ability to manipulate their ESG ratings, there should be some differences between the characteristics of firms below and above the cutoff. The lack of such differences implies that firms do not have such ability. Moreover, the ESG rating methodology employed by MSCI is arguably complicated enough that it would not allow firms to precisely manipulate their ESG ratings. Also, MSCI uses only publicly available information in its rating process, and does not accept any information, submitted in other ways, from the firms. Therefore, firms can not appeal to change their ESG ratings. Of course, firms could try to affect their ESG ratings by improving their ESG performance, but they can not *decide* what the outcome of the ESG rating process would be.

# 4 Results

## 4.1 Firms' Ownership by ESG Institutional Investors

In this section, I show that ESG institutional investors have significantly higher ownership in the treated firms compared to control firms, in the year following the treatment. In other words, firms that were above the cutoff in the year $t$, compared to firms that were below the cutoff in the same year, will have higher ownership by ESG institutional investors in the year $t+1$. These two groups of firms are very similar on every fundamental characteristics before the treatment (as shown in Section 3.2), including their actual ESG performance (since their ESG scores are very close to each other). The only difference between them is their different ESG label. These arbitrary ESG labels do not have any economic interpretations (they are predetermined by MSCI to map ESG scores into ESG letters). Therefore, the results of this section show that the ESG labels play an important role in the portfolio decisions of ESG institutional investors. While this is not a surprising finding, it implies that, everything else being equal, ESG institutional investors have preferences over firms with ESG labels of "Leader," compared to firms with ESG labels of "Average." Therefore, the RD setting in this study can capture the isolated effect of ESG ratings, disentangled from other firms' or investors' characteristics, on the firms' ownership by ESG institutional investors.

Table 4 shows the regression coefficients of Equation (2), where the outcome variable is the *ESG_Ownership*. In Panel A, a triangular kernel function has been used to assign different weights to the observations, based on their distance from the cutoff. Observations that are closer to (farther from) the cutoff have higher (lower) weights in the analysis. In Panel B, a uniform kernel has been employed. In columns (1) and (2), I have used an MSE-optimal bandwidth around the cutoff to estimate the coefficient. In columns (3) and (4), CE-optimal bandwidths have been used. In columns (1) and (3), polynomials of order 1 fit the data above and below the cutoff, while in columns (2) and

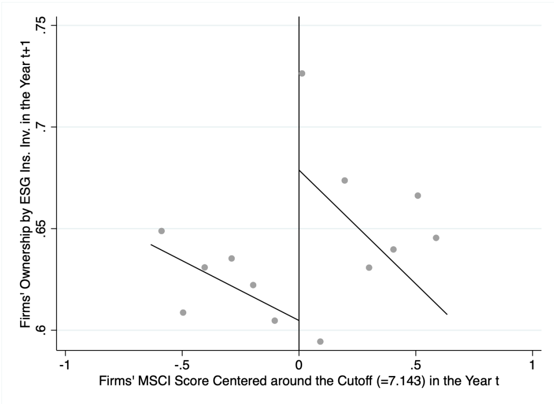
**Table 4: The Difference in Ownership by ESG Institutional Investors between Treated and Control Firms**

| **Panel A: Using *Triangular* Kernel** | | | | |
|---|---|---|---|---|
| | (1)<br>bw = 0.635<br>(MSE) | (2)<br>bw = 0.997<br>(MSE) | (3)<br>bw = 0.430<br>(CE) | (4)<br>bw = 0.640<br>(CE) |
| *Treatment Dummy* | 0.074** | 0.089** | 0.099** | 0.122** |
| | (2.328) | (2.355) | (2.564) | (2.561) |
| Kernel | Triangular | Triangular | Triangular | Triangular |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [493 182] | [560 209] | [493 182] | [560 209] |
| **Panel B: Using *Uniform* Kernel** | | | | |
| | (1)<br>bw = 0.502<br>(MSE) | (2)<br>bw = 0.699<br>(MSE) | (3)<br>bw = 0.340<br>(CE) | (4)<br>bw = 0.448<br>(CE) |
| *Treatment Dummy* | 0.056** | 0.093** | 0.096** | 0.127** |
| | (2.095) | (2.077) | (2.528) | (2.084) |
| Kernel | Uniform | Uniform | Uniform | Uniform |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [493 182] | [518 192] | [493 182] | [518 192] |

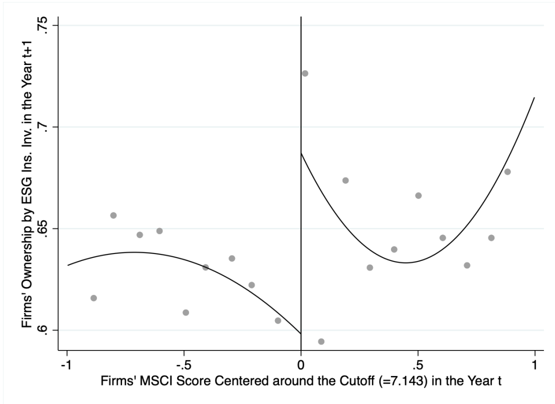
This table shows the regression coefficient $\tau$ (the intercept of the fitted polynomial above the cutoff minus the intercept of the fitted polynomial below the cutoff) in Equation (2) under different settings, where the outcome variable ($Y_{it+1}$) is *ESG_Ownership* for the firm $i$ in the year $t+1$. The running variable ($X_{it}$) is the MSCI ESG score for the firm $i$ in the year $t$. t-statistics based on standard errors clustered at the firm level are shown in parentheses. Panel A reports the results using Triangular kernel, and Panel B reports the results using Uniform kernel. Column (1) and (2) reports the results using MSE optimal bandwidth (bw). Columns (3) and (4) report the results using CE optimal bandwidth. Column (1) and (3) use polynomials of order 1 to fit the data, while columns (2) and (4) use polynomials of order 2. The last row in each panel report the effective number of observations below and above the cutoff that are used in the estimation. In all settings, ESG institutional investors have significantly higher ownership in treated firms ($D_{it} = 1$) than in control firms ($D_{it} = 0$), in the year following the treatment ($t+1$). (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively)

(4), polynomials of order 2 do this. Note that the size of the bandwidth is a function of the outcome variable, sample size, kernel function, the order of the polynomial, and several other factors, and therefore, is different in each of the settings.

In the setting recommended by Cattaneo et al. (2019), i. e., using MSE optimal bandwidth with a triangular kernel (columns (1) and (2) in Panel A), the intercept of the fitted polynomial for treated firms is 7.4% (using polynomial of order 1), or 8.9% (using polynomial of order 2) higher than the intercept of the fitted polynomial for control firms. Figure 3 shows the results from these two settings graphically. Note that because of the local nature of the RD framework, one should be careful in interpreting the magnitude of these results. The magnitude of the regression coefficient represents




**(a)** Polynomial Fit of Order 1




**(b)** Polynomial Fit of Order 2

**Figure 3: Firms' Ownership by ESG Institutional Investors around the Cutoff.** This figure shows the results of Equation (2) on the MSE-optimal bandwidth with triangular kernel, where the outcome variable ($Y_{it+1}$) is *ESG_Ownership* (Section 2.2.1) for the firm $i$ in the year $t+1$. In **(a)** polynomials of order 1 are used to fit the data (the setting reported in column (1) in Panel A of Table 4), while in **(b)**, polynomials of order 2 are employed (the setting reported in column (2) in Panel A of Table 4). Dots in graphs show the average ownership by ESG institutional investors for the firms in each bin.

the difference between the intercepts of the fitted polynomials at the different sides of the cutoff. Therefore, it is highly dependent on the form (the order of the polynomial) of these regression functions, the bandwidth, and the type of the kernel used. Taking this important consideration into account, Figure 3 shows the discontinuity ($\tau$) in the ownership by ESG institutional investors between the treated and control firms, around the cutoff. Similarly, Table 4 shows that the ESG institutional investors have significantly higher ownership in treated firms than in control firms, in the year following the treatment.

## 4.2 Firms' Perceived Cost of Capital and Actual Return

In this section, I first show that treated firms have lower perceived cost of capital compared to control firms, in the year following the treatment. This implies that treated firms expect to have lower cost of capital compared to control firms, therefore, they reduce their perceived cost of capital. Then, I show that, consistent with this expectation, treated firms will have lower actual returns compared to control firms, in the year following the treatment. The underlying economic intuition is that, consistent with the theoretical models of divestment strategy, ESG institutional investors hold higher ownership in treated firms than in control firms, as shown in Section 4.1. This higher demand pushes up the prices of treated firms, and therefore, reduces their cost of equity compared to control firms. Consequently, treated firms correctly reduce their perceived cost of capital in the year following the treatment.

Table 5 reports the regression coefficient of Equation (2), where the outcome variable is the *Perceived_CoC*. Like before, in Panel A, a triangular kernel function has been used to assign different weights to the observations based on their distance from the cutoff. In Panel B, a uniform kernel has been employed. In columns (1) and (2), I have used an MSE-optimal bandwidth around the cutoff to estimate the coefficient. In columns (3) and (4), CE-optimal bandwidths have been used. In columns (1) and (3), polynomials of order 1 fit the data above and below the cutoff, while in columns (2) and (4), polynomials of order 2 do this.

In the setting recommended by Cattaneo et al. (2019) in Column (1) and (2) in Panel A of Table 5, the intercept of the fitted polynomial for the treated firms is 0.34% (using polynomials of order 1), or 0.47% (using polynomials of order 2) lower than the intercept of the fitted polynomial for control firms. Figure 4 represents these results graphically. Again, note that these numbers show the discontinuity of the fitted polynomials around the cutoff and should be interpreted in the this sense. The results, collectively, show that treated firms have lower perceived cost of capital compared to control firms in the year following the treatment.

Next, I analyze the model in Equation (2), under different settings, where the outcome variable is the firms' actual return in the year following the treatment. The results in the Columns (1) and (2) in Panel A of Table 6 show that the fitted polynomial for the treated firms is 10.3% (using polynomials of order 1), or 13.9% (using polynomials of order 2) lower than the intercept of the fitted polynomial for control firms. The lower actual return for treated firms is consistent with the lower perceived cost of capital for these firms compared to control firms. Figure 5 represents these results graphically. Again, the magnitude of the $\tau$ coefficient shows the discontinuity in the fitted polyno-

## Table 5: The Difference in Perceived Cost of Capital between Treated and Control Firms

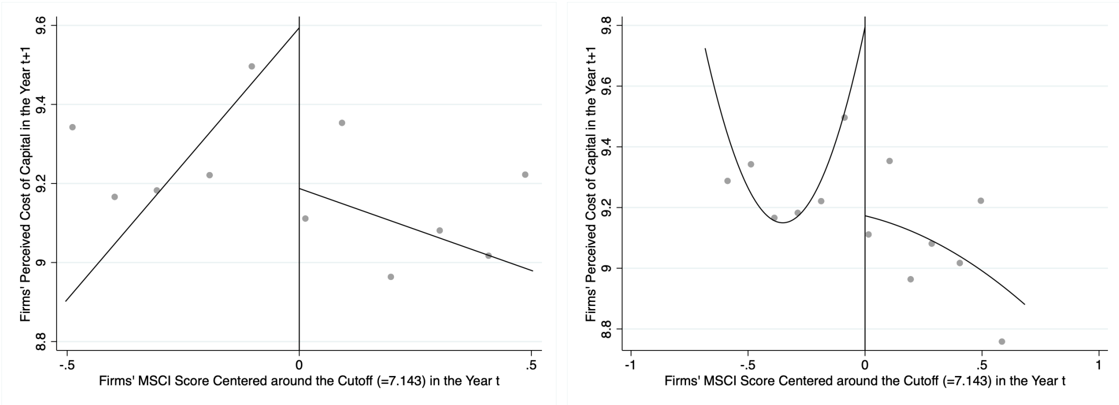
**Panel A: Using *Triangular* Kernel**

| | (1)<br>bw = 0.578<br>(MSE) | (2)<br>bw = 0.782<br>(MSE) | (3)<br>bw = 0.394<br>(CE) | (4)<br>bw = 0.504<br>(CE) |
|---|---|---|---|---|
| *Treatment Dummy* | -0.341** | -0.477*** | -0.481*** | -0.721*** |
| | (-2.514) | (-2.677) | (-2.796) | (-2.666) |
| Kernel | Triangular | Triangular | Triangular | Triangular |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [426 170] | [481 186] | [426 170] | [481 186] |

**Panel B: Using *Uniform* Kernel**

| | (1)<br>bw = 0.335<br>(MSE) | (2)<br>bw = 0.642<br>(MSE) | (3)<br>bw = 0.228<br>(CE) | (4)<br>bw = 0.414<br>(CE) |
|---|---|---|---|---|
| *Treatment Dummy* | -0.432*** | -0.560*** | -0.555** | -0.679** |
| | (-2.840) | (-2.859) | (-2.536) | (-2.534) |
| Kernel | Uniform | Uniform | Uniform | Uniform |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [360 153] | [456 178] | [360 153] | [456 178] |

This table shows the regression coefficient $\tau$ (the intercept of the fitted polynomial above the cutoff minus the intercept of the fitted polynomial below the cutoff) in Equation (2) under different settings, where the outcome variable ($Y_{it+1}$) is *Perceived_CoC* (Section 2.2.2) for the firm $i$ in the year $t+1$. The running variable ($X_{it}$) is the MSCI ESG score for the firm $i$ in the year $t$. t-statistics based on standard errors clustered at the firm level are shown in parentheses. Panel A reports the results using Triangular kernel, and Panel B reports the results using Uniform kernel. Column (1) and (2) reports the results using MSE optimal bandwidth (bw). Columns (3) and (4) report the results using CE optimal bandwidth. Column (1) and (3) use polynomials of order 1 to fit the data, while columns (2) and (4) use polynomials of order 2. The last row in each panel report the effective number of observations below and above the cutoff that are used in the estimation. In all settings, firms' perceived cost of capital is significantly lower for treated firms ($D_{it} = 1$) than for control firms ($D_{it} = 0$), in the year following the treatment ($t+1$). (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively)




**(a)** Polynomial Fit of Order 1

**(b)** Polynomial Fit of Order 2

**Figure 4: Firms' Perceived Cost of Capital around the Cutoff.** This figure shows the result of Equation (2) on the MSE-optimal bandwidth with triangular kernel, where the outcome variable ($Y_{it+1}$) is *Perceived_CoC* (Section 2.2.2) for the firm $i$ in the year $t+1$. In **(a)** polynomials of order 1 are used to fit the data (the setting reported in column (1) in Panel A of Table 5), while in **(b)**, polynomials of order 2 are employed (the setting reported in column (2) in Panel A of Table 5). Dots in graphs show the average perceived cost of capital for the firms in each bin.

mials at the cutoff, and not the difference in the average actual returns for the treated and control firms.

While lower actual returns imply that treated firms have lower cost of equity compared to control firms, the lower perceived cost of capital for treated firms may also be a result of lower cost of debt for these firms. To check if this is actually the case, I analyze Equation (2), with the outcome variable being the *firms' cost of debt capital*, defined as the firms' interest payment divided by their total debt. The (unreported) results show no significant difference in the cost of debt between the treated and control firms in the year following the treatment. Thus, the lower perceived cost of capital for treated firms is a result of lower cost of equity for these firms, and not the lower cost of debt.

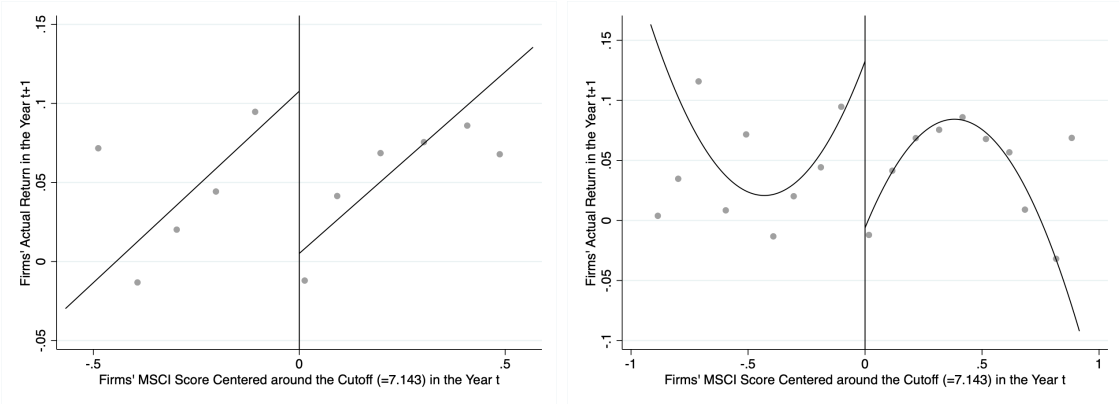
## 4.3 Firms' Financing Decisions

In this section, I focus on the effect of ESG ratings on the firms' financing decisions. In Section 4.2, I show that treated firms have lower perceived cost of capital compared to control firms, as a result of lower cost of equity. Also, I find that there are no differences in the cost of debt between the two groups of firms. Therefore, treated firms have access

## Table 6: The Difference in Actual Return between Treated and Control Firms

| **Panel A: Using *Triangular* Kernel** | | | | |
|---|---|---|---|---|
| | (1)<br>bw = 0.568<br>(MSE) | (2)<br>bw = 0.916<br>(MSE) | (3)<br>bw = 0.385<br>(CER) | (4)<br>bw = 0.587<br>(CER) |
| *Treatment Dummy* | -0.103* | -0.139** | -0.138** | -0.192** |
| | (-2.111) | (-2.389) | (-2.217) | (-2.459) |
| Kernel | Triangular | Triangular | Triangular | Triangular |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [537 196] | [608 222] | [537 196] | [608 222] |
| **Panel B: Using *Uniform* Kernel** | | | | |
| | (1)<br>bw = 0.507<br>(MSE) | (2)<br>bw = 0.791<br>(MSE) | (3)<br>bw = 0.344<br>(CER) | (4)<br>bw = 0.507<br>(CER) |
| *Treatment Dummy* | -0.064 | -0.163** | -0.129** | -0.205*** |
| | (-1.375) | (-2.424) | (-2.054) | (-2.632) |
| Kernel | Uniform | Uniform | Uniform | Uniform |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [511 188] | [583 213] | [511 188] | [583 213] |

This table shows the coefficient $\tau$ (the intercept of the fitted polynomial above the cutoff minus the intercept of the fitted polynomial below the cutoff) in Equation (2) under different settings, where the outcome variable ($Y_{it+1}$) is *Firms' actual Return* for the firm $i$ in the year $t+1$. The running variable ($X_{it}$) is the MSCI ESG score for the firm $i$ in the year $t$. t-statistics based on standard errors clustered at the firm level are shown in parentheses. Panel A reports the results using Triangular kernel, and Panel B reports the results using Uniform kernel. Column (1) and (2) reports the results using MSE optimal bandwidth (bw). Columns (3) and (4) report the results using CE optimal bandwidth. Column (1) and (3) use polynomials of order 1 to fit the data, while columns (2) and (4) use polynomials of order 2. The last row in each panel report the effective number of observations below and above the cutoff that are used in the estimation. In almost all settings, firms' actual returns are significantly lower for treated firms ($D_{it} = 1$) than for control firms ($D_{it} = 0$), in the year following the treatment ($t+1$). (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively)




**(a)** Polynomial Fit of Order 1 **(b)** Polynomial Fit of Order 2

**Figure 5: Firms' Actual Return around the Cutoff.** This figure shows the result of analyzing Equation (2) on the MSE-optimal bandwidth with triangular kernel, where the outcome variable ($Y_{it+1}$) is *Firms' Actual Return* for the firm $i$ in the year $t+1$. In **(a)** polynomials of order 1 are used to fit the data (the setting reported in column (1) in Panel A of Table 6), while in **(b)**, polynomials of order 2 are employed (the setting reported in column (2) in Panel A of Table 6). Dots in graphs show the average actual return for the firms in each bin.

to cheaper equity compared to control firms, in the year following the treatment; therefore, they should prefer to issue more equity than debt, compared to control firms. To test this argument, I analyze Equation (2) with the outcome variable *NetD-NetE*, introduced in the Section 2.2.3. This variable captures the preferences of firms to issue debt versus equity for financing their activities. I expect that the treated firms would issue more equity than debt, compared to control firms, in the year following the treatment; therefore, I expect the coefficient $\tau$ in Equation (2) to be negative.

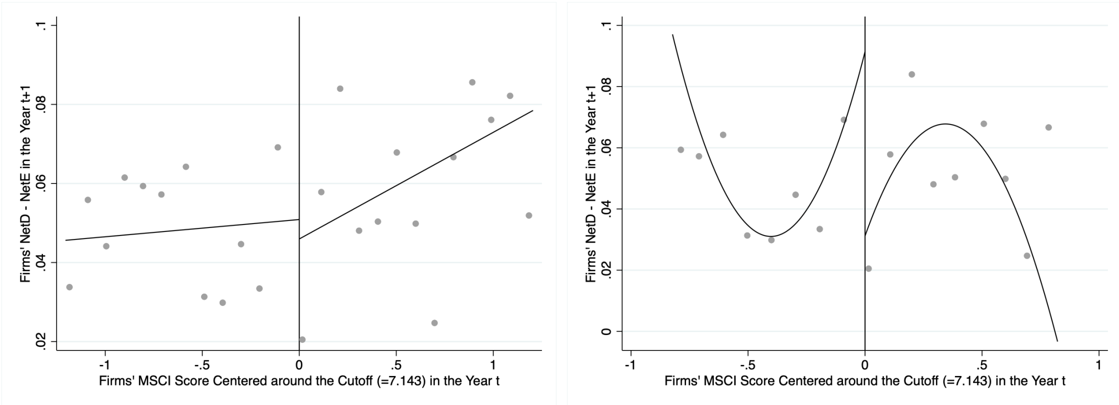
Consistent with this argument, the results in Table 7 show evidence that treated firms prefer to issue more equity than debt, compared to control firms, in the year following the treatment. This is in line with the findings in Section 4.2, which showed that treated firms have lower cost of equity compared to control firms. Figure 6 represents these results graphically.

It can be argued that if treated firms have access to cheaper equity, and thus their cost of capital is lower, their required rate of return (or discount rate) for investment decisions should also be lower. That means, everything else being equal, they should have access to more NPV positive projects, and thus more growth opportunities, compared to control firms. In fact, this "investment" channel is the main impact mechanism

## Table 7: The Difference in Financing Preferences between Treated and Control Firms

| **Panel A: Using *Triangular* Kernel** | | | | |
|---|---|---|---|---|
| | (1)<br>bw = 1.205<br>(MSE) | (2)<br>bw = 0.823<br>(MSE) | (3)<br>bw = 0.816<br>(CER) | (4)<br>bw = 0.527<br>(CER) |
| *Treatment Dummy* | -0.005<br>(-0.331) | -0.060**<br>(-2.473) | -0.015<br>(-0.730) | -0.079**<br>(-2.316) |
| Kernel | Triangular | Triangular | Triangular | Triangular |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [1373 517] | [958 347] | [958 347] | [679 771] |
| **Panel B: Using *Uniform* Kernel** | | | | |
| | (1)<br>bw = 0.724<br>(MSE) | (2)<br>bw = 1.161<br>(MSE) | (3)<br>bw = 0.491<br>(CER) | (4)<br>bw = 0.744<br>(CER) |
| *Treatment Dummy* | 0.001<br>(-0.039) | -0.012<br>(-0.486) | -0.029<br>(-1.200) | -0.056**<br>(-2.009) |
| Kernel | Uniform | Uniform | Uniform | Uniform |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [870 309] | [1246 470] | [551 215] | [870 309] |

This table shows the regression coefficient $\tau$ (the intercept of the fitted polynomial above the cutoff minus the intercept of the fitted polynomial below the cutoff) in Equation (2) under different settings, where the outcome variable ($Y_{it+1}$) is *Firms' Net Debt Issuance Minus Net Equity Issuance* for the firm $i$ in the year $t+1$. The running variable ($X_{it}$) is the MSCI ESG score for the firm $i$ in the year $t$. t-statistics based on standard errors clustered at the firm level are shown in parentheses. Panel A reports the results using Triangular kernel, and Panel B reports the results using Uniform kernel. Column (1) and (2) reports the results using MSE optimal bandwidth (bw). Columns (3) and (4) report the results using CE optimal bandwidth. Column (1) and (3) use polynomials of order 1 to fit the data, while columns (2) and (4) use polynomials of order 2. The last row in each panel report the effective number of observations below and above the cutoff that are used in the estimation. In some settings (specifically, while using polynomials of order two), firms' net equity issuance, compared to net debt issuance, are significantly higher for treated firms ($D_{it} = 1$) than for control firms ($D_{it} = 0$), in the year following the treatment ($t+1$). (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively)




**(a)** Polynomial Fit of Order 1

**(b)** Polynomial Fit of Order 2

**Figure 6: Firms' NetD-NetE around the Cutoff.** This figure shows the result of analyzing Equation (2) on the MSE-optimal bandwidth with triangular kernel, where the outcome variable ($Y_{it+1}$) is *NetQ-NetE* for the firm $i$ in the year $t+1$. In **(a)** polynomials of order 1 are used to fit the data (the setting reported in column (1) in Panel A of Table 7), while in **(b)**, polynomials of order 2 are employed (the setting reported in column (2) in Panel A of Table 7). Dots in graphs show the net debt issuance minus net equity issuance, divided by total assets, for the firms in each bin.

in the theoretical models that support the effectiveness of divestment strategy (Heinkel et al., 2001; Pástor et al., 2021). To check if this is actually the case here, I analyze Equation 2, with the outcome variable being $CAPEX/TotalAssets$. The (unreported) results show no significant differences in the investment ratio between treated and control firms. Moreover, I do not find any evidence that the sum of net equity issuance and net debt issuance is larger for treated firms compared to control firms.

A possible explanation for this result can be the absence of significant difference in the actual discount rates that the two groups of firm use to evaluate their investment decisions. The actual discount rate that firms use as the threshold for accepting or rejecting an investment opportunity is equal to their perceived cost of capital plus a "discount rate wedge" that takes into account several other considerations, like beliefs about value creation, risk, and financial constraints (Gormsen and Huber, 2023). To check if there are any differences in the discount rates used by treated firms and control firms, I analyze Equation 2 with the outcome variable being $Discount\_Rate$. The (unreported) results show no significant differences in the discount rate between treated and control firms, which provides a possible explanation for lack of difference in the investment ratio between these two groups of firms. Therefore, treated firms may issue more net equity, which is cheaper for them, in order to finance their current activities, without necessarily having access to more investment opportunities. This is also consistent with Berk and Van Binsbergen (2021) who argue that the decrease in the cost of capital resulting from a divestment strategy is too small to affect the real investment decisions of firms. However, having access to cheaper equity to finance the current activities may, by itself, motivate firms to improve their ESG performance. In this case, this "financing" channel may be another impact mechanism of divestment strategy that has been overlooked in the literature. Further studies are required to check if this mechanism actually exists, and if yes, what are its consequences. To provide some further evidence for the existence of this channel, in the next section I show that control firms improve their ESG ratings significantly more than the treated firms (for whom the ESG ratings almost remains the same), in the year following the treatment.

## 4.4 Change in firms' ESG Performance

In this section, I investigate the effect of ESG ratings on firms' ESG performance. As I argued in Section 4.2 and Section 4.3, treated firms have lower cost of capital compared to control firms in the year following the treatment, and thus, can finance their current activities by issuing cheaper equity. This can have two consequences. First, if control firms are aware of these benefits, they may have motivation to improve their ESG scores. Second, treated firms may have motivation to improve, or at least keep at the current level, their ESG scores. Focusing on the RD plot at the cutoff where the MSCI ESG label changes from "Average" to "Leader," Figure 7 shows that firms that are above the cutoff have a *less negative* change in their ESG score compared to firms that are below this cutoff. Note that Figure 7 shows the plot of a "global" RD analysis; therefore, I have used a polynomial of order 4 and a uniform kernel function to fit the data citepcattaneo2019practical.

Table 8 shows the results of analyzing Equation 2, with the outcome variable being the *dMSCI_Score* in the year $t+1$. The regression coefficient shows that the intercept of the fitted polynomial for the treated firms is, on the 0-10 scale, 0.52 (using polynomials

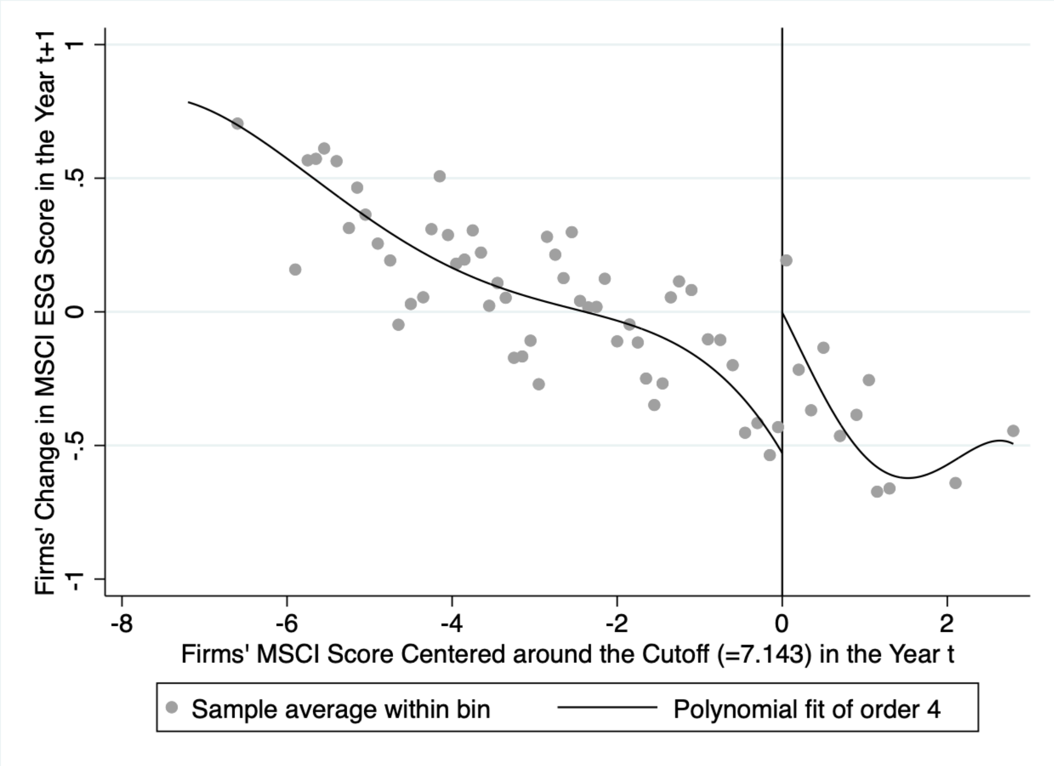



**Figure 7: Global Regression Discontinuity Plot of the Change in the Firms' MSCI ESG Score at the Cutoff (=7.143)**

of order 1), or 0.46 (using polynomials of order 2) higher than the intercept of the fitted polynomial for control firms. Figure 6 represents these results graphically.

These results show that the MSCI ESG score of treated firms declines less than the MSCI ESG score of control firms, in the year following the treatment. This implies that treated firms, that have benefited from the advantages of better ESG label, try to maintain their MSCI ESG scores, while control firms do not care much to do so.

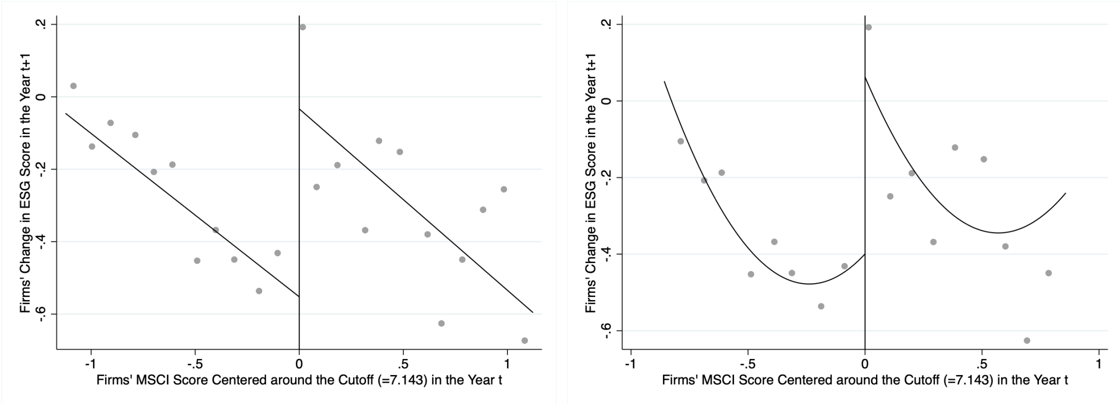
# 5 Robustness Checks

The main advantage of RD framework is that the process of assigning treatment is based on an observable feature of units. If units do not have the ability to precisely manipulate this feature, then the treatment assignment is as good as random, and the potential effect of the treatment on the outcome variable is isolated from other characteristics of the units. Therefore, the results of RD analysis are comparable to those of random assignment strategy. I show in Section 3.2 that in this study, the treated and control firms

## Table 8: The Difference in Change in ESG Score between Treated and Control Firms

| **Panel A: Using *Triangular* Kernel** | | | | |
|---|---|---|---|---|
| | (1)<br>bw = 1.123<br>(MSE) | (2)<br>bw = 0.858<br>(MSE) | (3)<br>bw = 0.761<br>(CER) | (4)<br>bw = 0.550<br>(CER) |
| *Treatment Dummy* | 0.519*** | 0.462* | 0.511*** | 0.510* |
| | (3.815) | (1.692) | (3.310) | (1.764) |
| Kernel | Triangular | Triangular | Triangular | Triangular |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [1246 470] | [958 347] | [870 309] | [679 248] |
| **Panel B: Using *Uniform* Kernel** | | | | |
| | (1)<br>bw = 0.740<br>(MSE) | (2)<br>bw = 0.962<br>(MSE) | (3)<br>bw = 0.0.501<br>(CER) | (4)<br>bw = 0.616<br>(CER) |
| *Treatment Dummy* | 0.546*** | 0.497** | 0.432** | 0.414* |
| | (3.242) | (2.224) | (2.435) | (1.732) |
| Kernel | Uniform | Uniform | Uniform | Uniform |
| Polynomial Order | 1 | 2 | 1 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm |
| No. Obs. [below above] | [870 309] | [1066 380] | [679 248] | [753 288] |

This table shows the coefficient $\tau$ (the intercept of the fitted polynomial above the cutoff minus the intercept of the fitted polynomial below the cutoff) in Equation (2) under different settings, where the outcome variable ($Y_{it+1}$) is *dMSCI_Score* for the firm $i$ in the year $t+1$. The running variable ($X_{it}$) is the MSCI ESG score for the firm $i$ in the year $t$. t-statistics based on standard errors clustered at the firm level are shown in parentheses. Panel A reports the results using Triangular kernel, and Panel B reports the results using Uniform kernel. Column (1) and (2) reports the results using MSE optimal bandwidth (bw). Columns (3) and (4) report the results using CE optimal bandwidth. Column (1) and (3) use polynomials of order 1 to fit the data, while columns (2) and (4) use polynomials of order 2. The last row in each panel report the effective number of observations below and above the cutoff that are used in the estimation. The firms' changes in MSCI ESG score are significantly higher (in absolute terms) for treated firms ($D_{it} = 1$) than for control firms ($D_{it} = 0$), in the year following the treatment ($t+1$). (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively)




**(a)** Polynomial Fit of Order 1 **(b)** Polynomial Fit of Order 2

**Figure 8: Firms' Change in ESG Score around the Cutoff.** This figure shows the result of running the Equation (2) on the MSE-optimal bandwidth with triangular kernel, where the outcome variable ($Y_{it+1}$) is *dMSCI_Score* for the firm $i$ in the year $t+1$. In **(a)** polynomials of order 1 are used to fit the data (the setting reported in column (1) in Panel A of Table 7), while in **(b)**, polynomials of order 2 are employed (the setting reported in column (2) in Panel A of Table 7). Dots in graphs show the average change in the MSCI ESG score for the firms in each bin.

are similar in their observable characteristics. This, and the complexity of MSCI ESG rating methodology, provide evidence that it is highly unlikely that firms can precisely manipulate their ratings to be placed in the treatment group.

In this section, I provide further tests to check the validity of my empirical results. First, I repeat the empirical results of analyzing Equation (2) for different outcome variables using arbitrarily chosen bandwidths to check the sensitivity of the results to the choice of bandwidth. However, it should be noted that I can not deviate too much from the MSE-optimal bandwidth, because using bandwidths that are much larger than the optimal bandwidth would increase the bias in the estimation of the regression coefficient, and using bandwidths that are much smaller than the optimal bandwidth would increase the variance in the estimation of the regression coefficient, both of which would make the results unreliable. Second, I repeat the analysis by choosing placebo cutoffs. the idea is that if the identified effect is actually due to the treatment, then we should not observe similar effects by choosing arbitrary cutoffs.

**Table 9: Sensitivity of Results to the Choice of Bandwidth**

| Bandwidth | (1) Ownership by ESG Ins. Inv. | (2) Perceived Cost of Capital | (3) Actual Return | (4) NetD-NetE | (5) Change in MSCI Score |
|---|---|---|---|---|---|
| 0.6 | | -0.683 | | -0.073* | |
| | | (-1.586) | | (-1.877) | |
| 0.7 | | -0.617** | -0.172* | -0.071* | |
| | | (-2.410) | (-1.746) | (-1.735) | |
| 0.8 | 0.099** | -0.466*** | -0.168* | -0.064** | 0.468 |
| | (2.553) | (-3.154) | (-1.956) | (-2.038) | (1.530) |
| 0.9 | 0.089** | -0.438*** | -0.148** | -0.055** | 0.460* |
| | (2.550) | (-2.951) | (-2.404) | (-2.276) | (1.657) |
| 1.0 | 0.089** | | -0.118*** | | 0.4771* |
| | (2.297) | | (-2.742) | | (1.736) |
| 1.1 | 0.087** | | | | 0.508* |
| | (2.296) | | | | (1.717) |
| 1.2 | | | | | |
| Kernel | Triangular | Triangular | Triangular | Triangular | Triangular |
| Pol. Order | 2 | 2 | 2 | 2 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm | Firm |
| MSE-opt bw | 0.997 | 0.782 | 0.916 | 0.823 | 0.858 |

This table shows the sensitivity of results in Section 4 to the choice of different bandwidths. Each column repeats the regression coefficients of Equation 2 for each of the outcome variables, choosing four different bandwidths around the MSE-optimal bandwidth, reported in the last row. Note that the MSE-optimal bandwidth is a function of outcome variable, among other factors; therefore, the optimal bandwidth and the four test bandwidths around it are different for each outcome variable. For saving space, I have only presented the result of using triangular kernel and polynomials of order 2 to fit the data. t-statistics based on standard errors clustered at the firm level are shown in parentheses. (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively

## 5.1 Sensitivity of Results to the Choice of Bandwidth

Table 9 reports the results of analyzing Equation (2) for different outcome variables by choosing different bandwidths around the MSE-optimal bandwidth. Note that I have limited the analysis to four bandwidths that are close to the MSE-optimal bandwidth to reduce the bias and variance in the estimation of the regression coefficient. Also note that the MSE-optimal bandwidth is a function of outcome variable, among other factors; therefore, the optimal bandwidth and the four test bandwidths around it are different for each outcome variable. I have only reported the results of using a triangular kernel and polynomials of order 2 to fit the data below and above the cutoff. These results show that in almost all settings, the findings of Section 4 have low sensitivity to the choice of bandwidth and remain statistically significant.

**Table 10: Placebo Cutoffs**

| Cutoff | (1)<br>Ownership by<br>ESG<br>Ins. Inv. | (2)<br>Perceived<br>Cost of<br>Capital | (3)<br>Actual Return | (4)<br>NetD-NetE | (5)<br>Change in<br>MSCI Score |
|---|---|---|---|---|---|
| 7.8 | -0.024 | -0.658 | -0.038 | -0.011 | -0.251 |
| | (-0.304) | (-1.615) | (-0.400) | (-0.199) | (-0.591) |
| 6.8 | 0.013 | -0.255 | -0.082* | 0.012 | 0.197 |
| | (0.781) | (-1.597) | (-1.648) | (0.644) | (1.038) |
| 5.8 | -0.039 | -0.001 | -0.013 | -0.032 | 0.455*** |
| | (-1.296) | (-0.235) | (-0.051) | (-1.536) | (3.584) |
| 4.8 | 0.008 | -0.017 | -0.066* | -0.005 | 0.077 |
| | (0.190) | (-0.010) | (-1.668) | (-0.369) | (0.927) |
| Kernel | Triangular | Triangular | Triangular | Triangular | Triangular |
| Pol. Order | 2 | 2 | 2 | 2 | 2 |
| SE Cluster | Firm | Firm | Firm | Firm | Firm |

This table shows the results of using placebo cutoffs, instead of the main cutoff. Each column repeats the analysis of regression equation 2 for each of the outcome variables. Each row represents an arbitrarily chosen cutoff. For saving space, I have only presented the result of using triangular kernel and polynomials of order 2 to fit the data. t-statistics based on standard errors clustered at the firm level are shown in parentheses. In almost all settings, the results are not statistically significant, implying that the detected effects in Section 4 have been due to the treatment, which is not present in other cutoffs. (*, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively)

## 5.2 Placebo Cutoffs

Table 10 shows the regression coefficient of Equation (2) for different outcome variables by choosing different arbitrary cutoffs. I have only reported the results of using a triangular kernel and polynomials of order 2 to fit the data below and above the cutoff. Lack of statistically significant results shows that the the detected effects in Section 4 have been actually due to the treatment, i. e., a sudden change of ESG letter rating at the cutoff 7.143.

# 6 Discussion

The results of this study contribute to the literature on the impact of ESG ratings on financial markets and firms' behaviour. Theoretical works on this topic suggest that ESG preferences of investors towards firms with better ESG ratings would reduce the cost of capital for these firms, because ESG investors are willing to replace part of their financial returns with non-pecuniary utility that they receive from investing in high-

ESG firms (Heinkel et al., 2001; Pedersen et al., 2021; Pástor et al., 2021). Lower cost of capital for high-ESG firms would, in theory, provide these firms with more investment opportunities, and therefore, low-ESG firms would be motivated to improve their ESG performance to take advantage of these benefits. The empirical evidence on this topic, however, have been mixed, and sometimes contradictory. One of the main empirical challenges in this literature is identifying the effect of a firm's ESG rating on the outcome variables, isolated from the effect of other characteristics of the firm.

This study provides a novel setting that addresses this difficulty. In an RD design, I focus on a small sample of firms around the cutoff where the firms' MSCI ESG label changes from "Average" to "Leader," or equivalently, their MSCI ESG letter rating changes from "A" to "AA." In this way, I can divide the sample into two groups: treated firms with a higher MSCI ESG label, and control firms with a lower ESG label. Importantly, show empirically that the firms in these two groups are similar on every other aspects, except for their MSCI ESG label. Since firms can not precisely manipulate their MSCI ESG ratings, the distribution of firms around this cutoff is as good as random, and any possible differences between the firms in these groups in the year following the treatment can be causally attributed to their different ESG labels. Therefore, I can capture the causal effect of ESG ratings on the outcome variables.

First, I confirm that ESG institutional investors are actually willing to invest more in the treated firms compared to control firms, in the year following the treatment. However, I do *not* verify if this willingness is actually a result of genuine ESG preferences (i. e., due to altruistic purposes), or a result of expecting better long-term financial performance (i. e., lower risk (Dumitrescu and Zakriya, 2021)) from high-ESG firms.

Second, I show that treated firms expect this higher demand, and in the year following the treatment, reduce their perceived cost of capital. Consistent with this expectation, treated firms earn lower actual returns compared to control firms, which can be interpreted as lower cost of equity for the former, compared to latter. Since I do not find any evidence that these two groups of firms have any differences in their cost of debt, it can be concluded that treated firms have lower cost of capital compared to control firms, in the year following the treatment.

Third, I find that consistent with the result regarding the lower cost of equity for treated firms, they issue more net equity than net debt, compared to control firms, in the year following the treatment. However, I do *not* find evidence that lower cost of capital provides treated firms with more investment opportunities. One possible explanation for this may be lack of difference in the actual discount rates between treated and control firms. While the managers of the treated firms have lower perceived cost of capital, they do not incorporate this in the actual discount rates that they use as the required rate of return for making investment decisions. Another reason may be the longer duration (more than one year) that it takes for the lower cost of capital to be reflected in the firms' real investment decisions. However, while I do not find any differences in the investment channel, treated firms can finance their current activities by issuing cheaper equity compared to control firms. I argue that, besides the investment channel, this "financing" channel may be another possible impact mechanism through which divestment strategy can motivate firms to improve their ESG performance.

Consistent with this argument, I find that the change in the MSCI ESG score is less negative for the treated firms compared to control firms in the year following the treatment. I show that in general, as the MSCI ESG score in the year $t$ increases, the change in this score in the year $t+1$ decreases. On average, this change is negative for the firms that have high ESG scores. However, there is a jump in the change in the ESG score for the firms that are below and above the cutoff where the ESG label changes from "Average" to "leader." In other words, while for both groups of firms, on average, the MSCI ESG score decreases in the year following the treatment, this decrease is smaller in magnitude for the treated firms. This implies that the benefits of better ESG label for treated firms can convince them to try to prevent their ESG scores from declining and downgrading to a lower level.

These results should be interpreted according to the specific setting of the study. While the results of an RD design provide high internal validity, they can *not* be generalized to the whole sample. The focus of this empirical analysis has been on a small neighborhood around the cutoff, which ignores a large part of the sample. Any observed effect should thus be interpreted in a local sense. Moreover, in the continuity-based approach to RD design in this study, the estimates are the discontinuity in the intercepts of the fitted polynomials below and above the cutoff. Therefore, these estimate do *not* show the difference in the average outcome for the treated and control firms. Finally, the choice of bandwidth, order of polynomials, and the kernel function can significantly affect the magnitude of the estimates in the continuity-based approach to RD design. The fact that my results are statistically significant, with consistent findings in a host of different settings, shows the validity of the conclusions, but does *not* justify the magnitude of the results.

# 7 Conclusion

This paper investigates the impact of ESG ratings on financial markets and firms' behaviour. I find that firms with better ESG labels attract more ESG institutional investors. Consequently, these firms have lower cost of equity, and can finance their current activities through issuing more net equity, that is now cheaper for them, than net debt. Therefore, these firms try to maintain their ESG score at the current level, to take advantage of this benefit. However, I do *not* find any evidence that firms with better ESG performance and lower cost of capital have better investment opportunities. Therefore, along with the theoretically supported "investment" channel, this "financing" channel may be another impact mechanism of divestment strategy, which has been until now overlooked in the literature in this field.

# References

Berg, F., Heeb, F., and Kölbel, J. F. (2022). The economic impact of ESG ratings. *Available at SSRN 4088545*.

Berk, J. and Van Binsbergen, J. H. (2021). The impact of impact investing. *Available at SSRN 3909166*.

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

Bond, P., Edmans, A., and Goldstein, I. (2012). The real effects of financial markets. *Annual Review of Financial Economics*, 4(1): 339–360.

Briere, M. and Ramelli, S. (2022). Green sentiment, stock returns, and corporate behavior. *Available at SSRN 3850923*.

Broccardo, E., Hart, O., and Zingales, L. (2022). Exit versus voice. *Journal of Political Economy*, 130(12): 3101–3145.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2016). Coverage error optimal confidence intervals for regression discontinuity designs. *Available at arXiv:1808.01398*.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, 113(522): 767–779.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *Econometrics Journal*, 23(2): 192–210.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *Stata Journal*, 17(2): 372–404.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4): 909–946.

Cao, J., Titman, S., Zhan, X., and Zhang, W. (2022). ESG preference, institutional trading, and stock return patterns. *Journal of Financial and Quantitative Analysis*, pages 1–35.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531): 1449–1455.

Cattaneo, M. D. and Titiunik, R. (2022). Regression discontinuity designs. *Annual Review of Economics*, 14: 821–851.

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

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

Dumitrescu, A., Gil-Bazo, J., and Zhou, F. (2022). Defining greenwashing. *Available at SSRN 4098411*.

Dumitrescu, A. and Zakriya, M. (2021). Stakeholders and the stock price crash risk: What matters in corporate social performance? *Journal of Corporate Finance*, 67: 101871.

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

El Ghoul, S., Guedhami, O., Kwok, C. C., and Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*, 35(9): 2388–2406.

Fama, E. F. and French, K. R. (2007). Disagreement, tastes, and asset prices. *Journal of Financial Economics*, 83(3): 667–689.

Gantchev, N., Giannetti, M., and Li, R. (2022). Does money talk? Divestitures and corporate environmental and social policies. *Review of Finance*, 26(6): 1469–1508.

Glück, M., Hübel, B., and Scholz, H. (2021). ESG rating events and stock market reactions. *Available at SSRN 3803254*.

Goldstein, I., Kopytov, A., Shen, L., and Xiang, H. (2022). On ESG investing: Heterogeneous preferences, information, and asset prices. *Availabe at SSRN 3823042*.

Gormsen, N. J. and Huber, K. (2023). Corporate discount rates. *Available at SSRN 4160186*.

Gormsen, N. J., Huber, K., and Oh, S. (2023). Climate capitalists. *Available at SSRN 4366445*.

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

Heath, D., Macciocchi, D., Michaely, R., and C. Ringgenberg, M. (2023). Does socially responsible investing change firm behavior? *Review of Finance*, 27(6): 2057–2083.

Heinkel, R., Kraus, A., and Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis*, 36(4): 431–449.

Hsu, P. H., Li, K., and Tsou, C. Y. (2023). The pollution premium. *Journal of Finance*, 78(3): 1343–1392.

Hwang, C. Y., Titman, S., and Wang, Y. (2022). Investor tastes, corporate behavior, and stock returns: An analysis of corporate social responsibility. *Management Science*, 68(10): 7131–7152.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3): 933–959.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2): 615–635.

Kisgen, D. J. (2006). Credit ratings and capital structure. *Journal of Finance*, 61(3): 1035–1072.

Kisgen, D. J. (2019). The impact of credit ratings on corporate behavior: Evidence from Moody's adjustments. *Journal of Corporate Finance*, 58: 567–582.

Krueger, P., Sautner, Z., and Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, 33(3): 1067–1111.

Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of Financial Economics*, 115(2): 304–329.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2): 281–355.

Lemmon, M. L., Roberts, M. R., and Zender, J. F. (2008). Back to the beginning: Persistence and the cross-section of corporate capital structure. *Journal of Finance*, 63(4): 1575–1608.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2): 698–714.

MSCI (2023). ESG ratings methodology. *Available at https://www.msci.com/esg-and-climate-methodologies*.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2): 550–571.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2022). Dissecting green returns. *Journal of Financial Economics*, 146(2): 403–424.

Pedersen, L. H., Fitzgibbons, S., and Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2): 572–597.

Rzeźnik, A., Hanley, K. W., and Pelizzon, L. (2022). Investor reliance on ESG ratings and stock price performance. *Available at SSRN 3801703*.