Three Essays on the Impact of Sustainable Investing

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktor der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

Florian Heeb

from Sennwald-Sax, SG

Approved in October 2022 at the request of:
Prof. Dr. Marc Chesney
Prof. Dr. Thorsten Hens

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 26.10.2022

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgements

Firstly, I would like to thank my PhD supervisor, Prof. Marc Chesney; for making this dissertation possible, for continuous guidance, and for encouraging me to venture into the field of financial economics as an environmental scientist. I would also like to thank Prof. Thorsten Hens for accepting to serve as an expert on my PhD committee and for valuable feedback throughout the different stages of my work.

I am grateful to Prof. Julian Kölbel for his friendship and his academic mentoring. Working with him for those five years was a privilege and made scientific work a true pleasure.

Further, I would like to thank the coauthors of the chapters included in this thesis—namely, Florian Berg, Prof. Timo Busch, Prof. Julian Kölbel, Prof. Falko Paetzold, and Prof. Stefan Zeisberger—for all their inspiration, hard work, and the many things I had the pleasure of learning from them.

Also, I am grateful to all the members of the Department of Finance who helped create the inspiring environment and run the administrative processes required for an enjoyable PhD, especially Prof. Steven Ongena, Sarah Wikus, and Eckhard Jäger. I would like to extend a special thanks to Prof. Falko Paetzold for creating such a vibrant hub for research on sustainable finance and for all his efforts toward making this research matter in the real world. And, of course, to all my current and past teammates at the Center for Sustainable Finance and Private Wealth—it was a pleasure working with them.

I am also grateful to many scholars who provided ideas for and feedback on my work: Jonathan Krakow, Prof. Emilio Marti, Prof. Stefano Ramelli, Julien Senn, and Prof. Paul Smeets and Vincent Wolf, to name just a few. Thanks also to Dave Brooks for tireless efforts to help a non-native navigate the depths of proper English language.

Finally, I would like to thank my family, my wife Esther, and my son Noël for being such patient and faithful companions on my PhD journey. Their love, joy, and support have been invaluable when navigating the ups and downs of scientific work. I look forward to all the future journeys we will embark on together.

Zurich, October 2022

Florian Heeb

Contents

I Introduction

I.1 Motivation

The world is waking up to a range of global challenges. To address these challenges, our economy needs to be transformed fundamentally, but how can that transformation be financed? With the Paris Agreement, most countries have agreed on limiting global warming to 1.5° Celsius above preindustrial temperature levels. The signatories agree that finance plays a vital role in efforts to reach this target. In Article 2 of the Paris Agreement, they commit to "*Making finance flows consistent with a pathway towards low greenhouse gas emissions and climate-resilient development.*" The scale of the financing required is unprecedented: according to a recent estimate, USD 3.5 trillion in additional investment is needed annually if the goals of Paris are to be reached (McKinsey Global Institute, 2022). And addressing other global challenges also requires large amounts of investment. The OECD estimates that USD 4.2 trillion of additional investment is needed per year if we are to achieve its 17 Sustainable Development Goals (SDGs), which include, among others, eradicating poverty, ending hunger, and gender equality (OECD, 2020).

At the same time, private investors are showing a growing demand for investment products with sustainability credentials. According to the Global Sustainable Investment Alliance, the total volumes of assets marketed as "sustainable" amounted to USD 35 trillion in 2020 (GSIA, 2021). Investors' demand for sustainable investment products is seen as a promising lever for solving global challenges (Benabou and Tirole, 2010; Broccardo, Hart, and Zingales, 2020; Hong, Wang, and Yang, 2021; Paetzold, Busch, and Chesney, 2015; Pástor, Stambaugh, and Taylor, 2021). This view is shared by many policy makers.[1]

However, it remains unclear under which circumstances the demand for investment products marketed as sustainable might (or might not) translate into the real-world change required to address global challenges. The three essays included as Chapters II–IV in this dissertation address this question from different perspectives. Chapter II explores

[1]For example, the EU's Sustainable Finance Disclosure Regulation (SFDR) states: "*As the Union is increasingly faced with the catastrophic and unpredictable consequences of climate change, resource depletion and other sustainability-related issues, urgent action is needed to mobilise capital not only through public policies but also by the financial services sector.*"

the different mechanisms by which investors can affect the real economy and provides a definition of investor impact (Kölbel, Heeb, Paetzold, and Busch, 2020). Chapter III puts the focus on ESG integration, the most prominent sustainable investing approach, investigating the impact of ESG ratings on financial markets and corporate behavior (Berg, Heeb, and Kölbel, 2022). Finally, Chapter IV examines how investors consider the impact of investment options in their decision-making (Heeb, Kölbel, Paetzold, and Zeisberger, 2022). The following section summarizes these three contributions.

I.2 Summary of Research Papers

I.2.1 How Can Investors Have Impact?

In Kölbel, Heeb, Paetzold, and Busch (2020) we investigate how investors can contribute to us reaching societal goals. First, we provide a formal definition of investor impact—so, of investors' effect on the real economy. Second, we review the existing evidence on how investors can have impact, distinguishing between three mechanisms: shareholder engagement, capital allocation, and indirect impacts.

We find empirical evidence that investors can trigger improvements in companies' ESG practices with shareholder engagement; yet these improvements seem to be limited to low-cost measures. The empirical evidence for investors' ability to affect corporate growth with their capital allocation is limited to young or small companies, or to immature financial markets. Theoretical models predict that investors can trigger improvements in ESG practices by withholding capital from companies that do not apply certain practices. We do not find empirical evidence, though, for such improvements. Further, from expert interviews and the literature we identify several indirect impact mechanisms via which investors might affect the behavior of consumers, regulators, or other investors. Studies, though, are scarce, and we do not find clear empirical evidence for any of these mechanisms.

Applying our findings to the current range of investment products marketed as sustainable, we conclude that only for a small fraction of these products is there evidence indicating that they enable investors to make a meaningful contribution to meeting societal challenges.

I.2.2 How Do ESG Ratings Affect the Economy?

In Berg, Heeb, and Kölbel (2022), we focus on how ESG ratings affect financial markets and the real economy. Theoretical models (as, e.g., in Pástor, Stambaugh, and Taylor (2021)) imply that by overweighting companies with a high level of ESG performance investors improve the cost of capital and thus the growth of such companies. Further, they incentivize improvements in corporate ESG practices. Our study specifically investigates the effects of changes in the MSCI ESG rating on mutual fund holdings, stock return, and corporate behavior.

We show that ESG rating downgrades reduce ownership by mutual funds with a dedicated ESG strategy, while upgrades increase it. This indicates that the ESG commitments of such funds are not merely cheap talk. We also find that stock markets react to ESG ratings: we observe a negative long-term response of stock returns to downgrades and a slower and weaker positive response to upgrades. We do not find evidence that ESG ratings affect corporate growth: we find no significant reaction of firms' capital expenditure to rating changes. We show that firms adjust their ESG practices following rating changes, but only in the governance dimension. While firms' governance improves after downgrades, it deteriorates after upgrades.

Taken together, our results show that ESG ratings have material effects on financial markets. However, our findings suggest that, so far, investors' consideration of ESG ratings has only a limited impact on the real economy.

I.2.3 How Do Investors Consider Impact in Their Decision-Making?

In Heeb, Kölbel, Paetzold, and Zeisberger (2022), we explore how investors consider the impact of investments in their decision-making. Current models of the effects of sustainable investing routinely assume that pro-social investors optimize between the cost and the impact of investment opportunities (e.g., Broccardo, Hart, and Zingales (2020); Oehmke and Opp (2019)). Yet this assumption has not been tested empirically. In a framed field experiment with private investors, we investigate how investors' willingness to pay (WTP) for sustainable investments responds to the impact of those investments.

We show that investors care about impact in the sense that they have a substantial WTP for investments with impact. However, they barely care about how *much* impact

investments have: investors do not pay significantly more for investments with more impact. This behavior does not seem to be due to a lack of experience, and the finding also holds for dedicated impact investors. Further, we show that when investors can compare several sustainable investments, they pay slightly more for investments that have a higher impact compared to other available options. Yet their WTP does not depend on how high that impact is in absolute terms.

Our findings suggest that investors' WTP for sustainable investments is primarily driven by an emotional rather than a calculative valuation of impact. Thus, rather than rewarding firms that effectively reduce negative externalities, pro-social investors may be rewarding financial intermediaries that sell products that offer warm glow.

I.3 Concluding Remarks

Taken together, these three essays suggest that sustainable investing has the potential to contribute to our efforts to address societal challenges—yet, current market practices barely realize that potential. Our findings suggest that sustainable investments are most suited to fostering the growth of young, impactful companies or to promoting the dissemination of good ESG practices. While this may help to reduce corporate externalities, it is unlikely to enable the large-scale infrastructure investments required, including—for example—for the decarbonization of our economy. Further, our findings suggest that the current market for sustainable investment is dominated by products that promise a rather modest impact and that investors may not necessarily ask for more. Thus, the tremendous market volumes referred to at the beginning of this introductory chapter may not translate into the change their sheer size may lead us to expect.

Our findings suggest two avenues of action for policy makers who want to avoid greenwashing and optimize the impact of investors' growing demand for sustainability. First, reliable information on the impact of investment products is needed. We show that the evidence required on the effectiveness of different investment approaches is insufficient in many cases—which also points to a possibly rewarding avenue for further research. Second, the positive emotions investors derive from selecting sustainable investments need to be linked with the real-world impact of these investments. This could be achieved, for example, with labels that directly align with socially desirable outcomes, or by monetizing the real-world impact of investments.

II How Can Investors Have Impact?

Can Sustainable Investing Save the World? Reviewing the Mechanisms of Investor Impact*

Julian F. Kölbel, Florian Heeb, Falko Paetzold, and Timo Busch

A version of this paper has been published as Kölbel, Julian F., Florian Heeb, Falko Paetzold, and Timo Busch, 2020, Can Sustainable Investing Save the World? Reviewing the Mechanisms of Investor Impact, *Organization & Environment*, 33(4), 554–574.

Abstract

This paper asks how sustainable investing (SI) contributes to societal goals, conducting a literature review on investor impact—that is, the change investors trigger in companies' environmental and social impact. We distinguish three impact mechanisms—shareholder engagement, capital allocation, and indirect impacts—concluding that the impact of shareholder engagement is well supported in the literature, the impact of capital allocation only partially, and indirect impacts lack empirical support. Our results suggest that investors who seek impact should pursue shareholder engagement throughout their portfolio, allocate capital to sustainable companies whose growth is limited by external financing conditions, and screen out companies based on the absence of specific ESG practices that can be adopted at reasonable cost. For rating agencies, we outline steps to develop investor impact metrics. For policy makers, we highlight that SI helps to diffuse good business practices, but is unlikely to drive a deeper transformation without additional policy measures.

*We gratefully acknowledge financial support from the Luc Hoffman Institute. We also acknowledge financial support from the European Union's EIT Climate KIC under the SGA2019. We thank Marc Chesney, James Gifford, Jonathan Krakow, Vincent Wolf, Alex Barkawi, Britta Rendlen, Emilio Marti, Raj Thamoteram, Tillmann Lang, Julia Meyer, David Wood, Florian Berg, Ola Mahmoud, Bert Scholtens, and Paul Smeets for valuable input and discussions. We would also like to thank participants of the Luc Hoffmann Institute workshop "Shareholder Activism for Sustainability" on May 28, 2018, the Yale Initiative on Sustainable Finance Symposium 2018, and the Global Research Alliance for Sustainable Finance and Investment 2020 Annual Conference.

II.1 Introduction

There are growing expectations that sustainable investing (SI)—that is, investing that takes environmental, social, and governance (ESG) information into account—will contribute to the achievement of societal goals. Historically, the Quakers divested to avoid supporting the slave trade, and colleges divested to challenge the South African apartheid regime (Broadhurst, Watson, and Marshall, 2003). Today too, many investors are attracted to SI due to their altruistic motives (Hartzmark and Sussman, 2019; Riedl and Smeets, 2017), expecting that SI will allow them to make a positive impact. Banks and asset managers are catering to these expectations by offering more and more investment products that emphasize sustainability, responsibility, and—increasingly—impact (GSIA, 2018). Policy makers too are discussing SI as a potential mechanism for mitigating climate change (IPCC, 2018) and for helping us to realize the United Nations' Sustainable Development Goals (SDGs) (Betti, Consolandi, and Eccles, 2018).

Yet in spite of these high expectations, little is known about the actual impact investors make through SI. We define investor impact as the change that investor activities achieve in company impact, and company impact as the change that company activities achieve in social and environmental parameters. These definitions are consistent with prior academic literature (Brest, Gilson, and Wolfson, 2018) as well as with the views of leading institutions in the field of impact evaluation (IFC, 2019).

The concept of investor impact is only beginning to take root in the SI industry. Currently, most SI funds either exclude firms operating in harmful industries or focus on companies that have in the past performed well on metrics of ESG performance. This is a static approach, and one that ignores that impact is fundamentally about change. Companies can and do change over time, and investors make an impact by triggering or accelerating such change. Due to a lack of suitable metrics for investor impact, however, very few investors analyze how their activities cause companies to change. As a result, the majority of the USD 30 billion that is deployed in SI today (GSIA, 2018) is invested in ways that promise only modest and perhaps even negligible investor impact.

To date, academic literature on SI has also neglected the concept of investor impact. Many studies rely in their analysis on ESG metrics, which can be interpreted as a proxy

for company impact.[1] However, the vast majority of studies use ESG metrics as an explanatory variable, very few analyzing ESG metrics as a dependent variable. As a result, little is known about what drives changes in company impact in general. This is true in particular for the literature on the financial performance of SI (Friede, Busch, and Bassen, 2015; Lins, Servaes, and Tamayo, 2017; Renneboog, Ter Horst, and Zhang, 2008), which investigates how ESG metrics influence investment performance, but not how different ways of investing influence ESG metrics. As a consequence, there is a gap regarding the mechanisms of investor impact in the literature on SI.

While there are studies that deal in some way with investor impact, they often do so in a context that is not related to SI. This paper therefore conducts a broad review of such literature to bring together findings regarding the mechanisms of and evidence for investor impact. We distinguish shareholder engagement, capital allocation, and indirect impacts as the three principal mechanisms of investor impact. For each mechanism, we evaluate the existing empirical evidence as reported in the literature and establish key determinants that increase or decrease investor impact. Shareholder engagement emerges as the most reliable mechanism for investors seeking impact, in the sense that it has been clearly demonstrated empirically. The impact of capital allocation is less reliable, since different parts of the mechanism have been studied empirically, but not yet in combination. Indirect impact mechanisms, which include stigmatization, endorsement, benchmarking, and demonstration, have hardly any empirical support in the literature so far.

The findings of this review have important implications for investors, ESG data providers, and policy makers. Investors who want to stimulate real-world impact based on evidence have three ways of pursuing this aim. First, roll out shareholder engagement throughout their portfolio, focusing on requests that have a good chance of success and yield substantial improvements in company impact. Second, allocate capital to companies that have positive company impact, under the condition that these firms are constrained in their growth by external financing conditions. This condition most likely applies to smaller firms operating in less mature financial markets. Third, screen out investments based on the absence of ESG practices, focusing on a few specific and transparently communicated practices that have a low "cost of reform" for companies.

[1] Studies that address the quality of ESG metrics (Chatterji, Durand, Levine, and Touboul, 2016; Vörösmarty, Osuna, Koehler, Klop, Spengler, Buonocore, Cak, Tessler, Corsi, Green, et al., 2018) highlight a number of important problems with those metrics, but do not criticize their lack of attention to investor impact.

This should be pursued in a large coalition of investors and encompass stocks and bonds. ESG data providers, meanwhile, should consider developing metrics of investor impact. Most existing SI fund ratings provide a snapshot of the company impact of the portfolio constituents. This leaves a gap for ratings that reflect the change in company impact that the fund is driving through its investment activity. Finally, policy makers should be aware that while SI is a powerful mechanism for diffusing good business practices, sustainable investing alone is unlikely to transform industries without additional policy measures.

II.2 Key Concepts and Scope

Our literature review aims to gather the available scientific evidence for the different mechanisms of investor impact. To set the scope of the review, we provide a detailed explanation of the concept of investor impact and describe the mechanisms of investor impact, which are illustrated in Figure II.1.

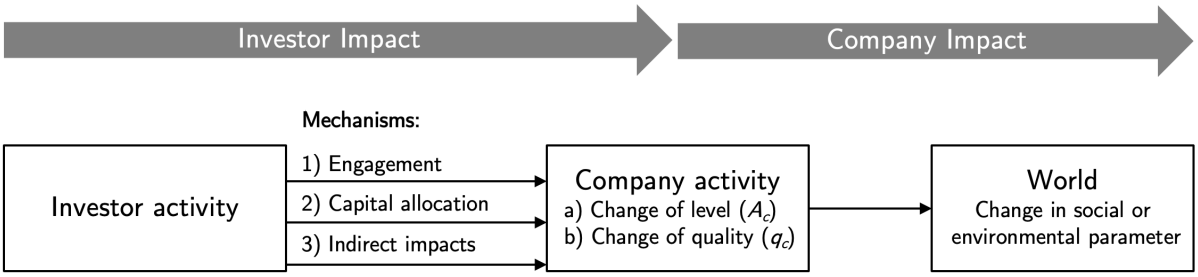


Figure II.1. Investor Impact. This figure illustrates our definition of investor impact. We define investor impact as the change an investor's activity achieves in company impact. And we define company impact as the change a company's activities achieve in a social or environmental parameter.

II.2.1 Investor Impact

The notion of impact in an investment context originally occurs in development finance, where funds are directed toward programs with the intention of improving livelihoods. The World Bank characterizes impact as "causal effects of a program on an outcome of interest" (Gertler, Martinez, Premand, Rawlings, and Vermeersch, 2016, p. 8). There is a

rich literature concerned with impact evaluation, mostly with applications to development finance, philanthropy, and foreign aid (Bamberger and Mabry, 2019). In this literature, impact is consistently described as having three defining characteristics: (1) it describes a change against a baseline, (2) it relates to a clearly defined parameter, and (3) it implies causality in the sense that the change would not have occurred in the absence of the activity. The last requirement is also referred to as additionality (see, e.g., Greiner and Michaelowa (2003)). On this basis, we define impact as *change in a specific social or environmental parameter that is caused by an activity.*

In the context of SI, it is useful to distinguish between the impact of investors and the impact of companies. Investors do not have a direct impact on social and environmental parameters. Instead, investors have an impact on the companies they invest in, which in turn have an impact on social and environmental parameters (Brest and Born, 2013; Brest, Gilson, and Wolfson, 2018). Thus, we define investor impact as the *change that investor activity achieves in company impact*, and we refer to company impact as the *change that a company's activities achieve in a social or environmental parameter*. This definition is in line with the recently released principles for impact management from the International Finance Corporation (IFC), which stipulate that investors should establish a narrative that outlines how the investor contributed to the achievement of company impact (IFC, 2019).

Based on this qualitative definition, we also provide a formal definition of investor impact. For this, we denote company impact as the product of A_c and q_c. Here, A_c is a measure of the level of a company's economic activities (e.g., its output of goods and services or its revenue). And q_c is the effect a company's activities have on a social or environmental parameter of interest, per unit of economic activity. We also refer to q_c as the "quality" of the company's activities. Further, we differentiate between two concurrent states of the world. In the first (observable) state, which we denote as state 1, an investor's activity of interest has taken place. In the second (hypothetical) state, which we denote as state 0, an investor's activity has not occurred. This corresponds to the baseline state or the counterfactual. Thus, we formally define investor impact as follows:

$$\textit{Investor impact} \;=\; \Delta(A_c \,*\, q_c) \;\stackrel{\text{def}}{=}\; A_{c,1} \,*\, q_{c,1} \;-\; A_{c,0} \,*\, q_{c,0}. \tag{II.1}$$

We can differentiate between two special cases of how investors' activities can have impact, as illustrated in Figure II.2. First, certain activities of investors may affect the level of a firm's economic activities but not the quality of the firm's activities. In this case, the investor impact is equal to $\Delta A_c * q_{c,0}$. We refer to this case as the "growth" channel. Second, certain activities of investors may not affect the level of a firm's economic activities but alter the quality of its activities. We refer to this case as the "reform" channel. In this case, the investor impact equals $A_{c,0} * \Delta q_c$.

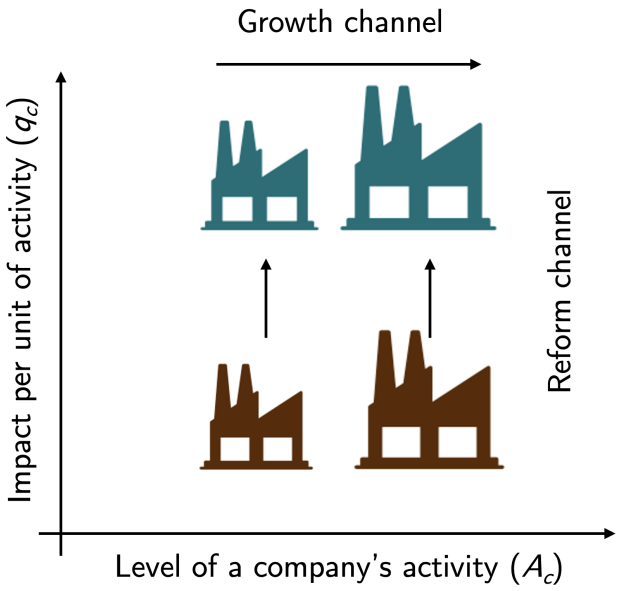


Figure II.2. Growth and Reform Channels. This figure illustrates two fundamentally different ways in which investors can affect companies' impact: either by affecting firms' growth and thus their level of economic activity (A_c) or by triggering reforms in firms' social or environmental practices, thus affecting the quality of firms' activities (q_c). Figuratively speaking, investors may have an impact by accelerating the growth of "green" companies, or by causing companies to become "greener."

To give an illustrative example applying our definition, consider a company that manufactures solar panels. We will look at impacts on global carbon dioxide emissions per year as the parameter of interest. The level of the company's activity A_c is measured in solar panels produced, and the quality of the company's activity q_c is given by the average carbon emission savings achieved with each panel produced. Let us assume the company sells 10,000 panels per year, each of which, over its life cycle, precludes 3 tons of carbon emissions. Then the company impact is 30,000 tons of precluded carbon emissions per year.

Now consider a USD 10 million investment in this solar panel manufacturer. Let us assume that this investment allows the company to increase its level of activities, in this case the production of solar panels, by 10%. As a result, the company's impact will increase by 10% yielding an investor impact of 3,000 tons of precluded carbon emissions. Alternatively, the investor behind this investment may use the influence that comes with his or her equity stake to improve the quality of the company's activity. For example, the investor may convince the company's management to recycle old solar panels, increasing the carbon emissions precluded per panel by 10%. Again, the company's impact will increase by 10%, yielding an investor impact of 3,000 tons of precluded carbon emissions—simply by alternative means.

To illustrate how this example could be applied in practice, one might consider using a company's ESG practices as a proxy for q_c and a company's size as a proxy for A_c. Positive investor impact could be achieved by either improving the company's ESG practices or by increasing the company's market share, provided that its ESG practices are above average.

II.2.2 Mechanisms of Investor Impact

There are various mechanisms of investor impact—so, mechanisms via which investors can influence the level of firms' economic activities, A_c, or their quality, q_c. In order to ensure that our review covers all the relevant mechanisms, we conducted a series of interviews with experts from the sustainable investing industry. Specifically, we asked industry experts for anecdotal evidence of when their activities or the activities of their organizations have triggered change at companies. In addition, we shared and obtained feedback on earlier versions of this paper from asset managers, asset owners, and regulators in Switzerland and the U.S. during conferences and workshops. We focused on mechanisms that are available to investors alone, excluding mechanisms that are also available to other actors. For example, investors may have an influence by lobbying regulators. However, examining the impacts of lobbying activity would broaden the scope of this paper, taking it far beyond investor impact alone.

Through these exchanges with industry experts and regulators we identified three fundamental mechanisms of investor impact, as shown in Figure II.1. Shareholder engagement refers to shareholder activities that are intended to change companies' ESG practices, often referred to as "voice" (Hirschman, 1970). These include the right to vote

on shareholder proposals during annual general meetings, discussions during informal meetings with management, and criticizing corporate practices in news outlets, as well as threats of selling a company's assets (see, e.g., Admati and Pfleiderer (2009); McCahery, Sautner, and Starks (2016), and Amel-Zadeh and Serafeim (2018)). Capital allocation refers to the investor activity of allocating capital to particular financial assets. Investors may either buy a company's financial assets, implicitly backing the company with their capital, or sell a company's financial assets, denying the company such backing. The latter is commonly referred to as "exit" (Hirschman, 1970). Indirect impacts include a range of impact mechanisms where investor activities do not directly affect company activities, but where instead the activity of investors influences a third party, which in turn affects company activities. Stigmatization refers to an investor tainting a company's image in public; endorsement refers to an investor endorsing and promoting a company's sustainability performance; benchmarking refers to rating agencies measuring and benchmarking a company's ESG performance; and demonstration refers to investors encouraging other investors to follow their lead.

II.2.3 Scope of the Literature Review

The scope of the review includes scholarly work that addresses the identified mechanisms of investor impact. For each mechanism, we queried academic databases with keywords describing the investor impact mechanisms, yielding an initial body of literature. We then extended the range of keywords by searching for central concepts and keywords drawn from the body of literature already identified. For example, the concept of "stock price elasticity" was identified as an important theoretical basis for capital allocation, leading us to include a body of literature dealing with stock price elasticity in our review. This approach ensured that we could identify all contributions that are important for the identified mechanisms, even if they use different terms to describe the mechanisms, or deal only with particular aspects of the mechanisms.

Using this approach, we identified a total of 64 relevant contributions from a range of different disciplines. Capital allocation is dealt with mostly in the financial economics literature, specifically asset pricing and corporate finance. Shareholder engagement is dealt with mostly in the corporate governance literature, as well as in management science. Indirect impacts are dealt with primarily in business ethics, management science, and

sociology. We analyze this body of literature to assess the empirical evidence for each mechanism as well as known determinants affecting the effectiveness of the mechanisms.

II.3 Literature Review

Table II.1 provides a comparison of the level of empirical evidence for the mechanisms of investor impact, as well as the known determinants that are likely to influence the mechanisms' effectiveness. Determinants that have a positive influence on effectiveness are denoted with a plus sign, those with a negative influence with a minus sign.

II.3.1 Shareholder Engagement

Active engagement of shareholders may cause companies to improve the quality, q_c, of their activities. There are five empirical studies that analyze the extent to which companies comply with shareholder engagement requests (Barko, Cremers, and Renneboog, 2017; Dimson, Karakaş, and Li, 2015, 2021; Dyck, Lins, Roth, and Wagner, 2019; Hoepner, Oikonomou, Sautner, Starks, and Zhou, 2018). The key results of these studies are summarized in Table II.2. Taking different approaches and relying on different data, each study reports the success rate of a given number of shareholder engagement requests as part of its analysis. The results show that while shareholder engagement requests do not always succeed, there is a reasonable probability that they do, ranging from 18 percent to 60 percent.

Going into more detail, Barko, Cremers, and Renneboog (2017) and Dyck, Lins, Roth, and Wagner (2019) show that shareholder proposals are associated with subsequent increases in the ESG ratings of targeted companies, providing evidence that shareholder engagement can lead to changes in company activities that are detectable in data that is provided by sources other than engagement service providers themselves. Together, these studies provide strong evidence that shareholder engagement is an effective mechanism through which investors can trigger reforms that improve the quality of company activities.

The success rate of shareholder engagements depends on a host of determinants related to characteristics of the engagement request, the company engaged with, the investor engaging, and the specific process of engagement (Goranova and Ryan, 2014). The studies

Table II.1
Evidence for and determinants of investor impact

This table summarizes our findings regarding the level of empirical evidence for the different mechanisms of investor impact, as well as the key determinants governing the effectiveness of these mechanisms. For the determinants, factors that positively affect the effectiveness of a mechanism are marked with (+), and factors that negatively affect the effectiveness of a mechanism with (−).

Mechanism	**Empirical Evidence**	**Key Determinants**
Shareholder engagement	Entire mechanism has been analyzed. Evidence of effect on quality of company activity.	• Investor influence (+) • Company's level of ESG experience (+) • Cost of requested reform (−)
Capital allocation		
Incentivizing improvements	Key parts of the mechanism have been analyzed separately. Indications of effect on quality of company activity.	• Market share of investors applying a screening approach (+) • Substitutability of affected assets (−) • Cost of requested reform (−)
Affecting growth	Key parts of the mechanism have been analyzed separately. Indications of effect on level of company activity.	• Improvement in financing conditions (+) • Size of company (−) • Age of company (−) • Maturity of financial markets (−)
Indirect impacts		
Stigmatization	No evidence.	
Endorsement	Some parts of the mechanism have been analyzed in isolation. Insufficient evidence of effect on level or quality of company activity.	• ESG reputation prior to endorsement (−)
Benchmarking	Some parts of the mechanism have been analyzed in isolation. Insufficient evidence of effect on level or quality of company activity.	• Consistency of ESG benchmarks (+)
Demonstration	No evidence.	

Table II.2
Success rates of shareholder engagement requests

This table summarizes the findings of five empirical studies that investigate the effectiveness of shareholder engagement. For each study, the number of individual shareholder requests are listed, as well as the sampling period over which these requests were observed. Further, the table reports the success rate of the requests—i.e., how many of the observed requests were closed as considered successful.

Reference	# of requests	Sample period	Success rate
Dimson et al. (2015)	2,152	1999–2009	18%
Hoepner et al. (2016)	682	2005–2014	28%
Barko et al. (2017)	847	2005–2014	60%
Dimson et al. (2018)	1,671	2007–2017	42%
Dyck et al. (2019)	147	2004–2013	33%

cited above highlight three specific determinants that have an important influence on the average rate of success.

The first determinant is the cost of the reform that is associated with complying with the engagement request. A consistent finding of the reviewed studies is that requests in the environmental domain tend to have lower success rates compared to requests in the social domain, and that requests in the corporate governance domain have the highest rate of success. Dimson, Karakaş, and Li (2015) attribute this to the fact that reforms in the environmental domain are likely to be costlier than those in the governance domain. More explicitly, Barko, Cremers, and Renneboog (2017) show that requests that require some form of costly reorganization have lower success rates compared to requests that entail lower costs. Taken together, these findings indicate that the chances of success decrease as the costs of the requested reform rise.

The second determinant is investor influence. There is evidence that engagement requests are more likely to succeed when the shareholder engaging holds a larger share of the targeted company (Dimson, Karakaş, and Li, 2015, 2021). However, it is not only the presence of larger holdings that causes investor influence to increase. Dimson, Karakaş, and Li (2021) find that a group of investors engaging has more influence when the engagement is spearheaded by an investor who is from the same country as the company being engaged with, suggesting that linguistic and cultural elements may play a role as well. Additionally, the chances of success rise when asset managers that are large and internationally renowned are part of the group of investors engaging.

The third determinant is the company's level of ESG experience. The success rate of engagement is higher with companies that have previously complied with engagement requests (Barko, Cremers, and Renneboog, 2017; Dimson, Karakaş, and Li, 2015). Furthermore, companies that had high ESG ratings prior to the engagement are more likely to comply with engagement requests (Barko, Cremers, and Renneboog, 2017).

II.3.2 Capital Allocation

While the impact of capital allocation on company activities may seem intuitive at first glance, it touches upon a rather fundamental question—namely, to what extent the decisions of investors influence the course of the real economy (see, e.g., Morck, Shleifer, Vishny, Shapiro, and Poterba (1990)). From the reviewed literature, we identify two mechanisms governing how the capital allocation of sustainable investors may influence company impact: by creating incentives to improve ESG practices and therefore the quality, q_c, of company activities, and by affecting growth and therefore the level of company activities, A_c. In the following, we review the available literature for each of these two mechanisms.

Incentivizing Improvements

Sustainable investors may shift asset prices by applying screening approaches. There are several different screening approaches in practice, including negative screening (excluding certain harmful industries), norm-based screening (excluding companies that do not adhere to widely accepted norms of business conduct), and best-in-class screening (allocating capital to companies that have the best ESG performance relative to their industry peers). ESG integration, where investors consider ESG metrics as part of the investment analysis, can be regarded as a rather complex screening approach, which eventually results in some companies being over- or underweighted or excluded from the investment portfolio. All approaches have in common that they result in a portfolio allocation that differs from the market portfolio. Through this deviation, investors may create incentives for companies that do not fulfill inclusion criteria to enact reforms. Thus, investors may be able to trigger changes in ESG practices through screening approaches. However, there is no empirical evidence that explicitly links sustainable investors' screening approaches to changes in ESG practices. There is some evidence that screening approaches affect

asset prices, and theoretical models that predict an effect on ESG practices. There remains, however, considerable uncertainty as to whether the model assumptions hold in practice.

Several theoretical studies have modeled the consequences of screening approaches for asset prices. In their equilibrium model, Heinkel, Kraus, and Zechner (2001) argue that sustainability preferences of investors can influence asset prices. This is in line with findings of Fama and French (2007) as well as those of Luo and Balvers (2017), both pairs of authors showing—based on standard asset pricing models—that preference-neutral investors require a premium for balancing out the portfolio choices of investors who share a particular nonfinancial preference because this forces the preference-neutral investors to deviate from the market portfolio. Further, Heinkel, Kraus, and Zechner (2001) predict that if the decrease in the stock prices of firms that do not conform to the requirements of sustainable investors' screening approaches is significant enough, these firms will start to implement the reforms demanded by sustainable investors. This is in accordance with Edmans, Goldstein, and Jiang (2012), who argue that when managerial incentives are tied to stock market value, managers will be sensitive to non-fundamental shifts in the share price of their corporation. Regarding the proportion of investors that needs to apply a screening approach in order to trigger corporate improvements, Heinkel, Kraus, and Zechner (2001) provide a numerical example in which at least 20 percent of investors need to apply a common screening approach to create the incentive for a company to implement reforms that cost 5 percent of its annual cash flow.

Similarly, the equilibrium model of Gollier and Pouget (2014) shows that by "voting with their feet"—that is, by consistently divesting from companies that do not fulfill certain criteria—investors can lower asset prices for these companies and with that incentivize firms to invest in order to fulfill these criteria. Based on the findings of the Stern Review (Stern, 2007), Gollier and Pouget (2014) estimate that 8 percent of investors applying the same screening approach is required to incentivize the investments in the new technologies required to mitigate climate change. However, both the quantitative estimates of Gollier and Pouget (2014) and those of Heinkel, Kraus, and Zechner (2001) must be considered with caution as they depend strongly on stylized assumptions. As a result, it remains unclear what share of SI investors is required to cause movements in asset prices sufficiently large to incentivize meaningful improvements.

A series of studies provide empirical evidence that screening approaches of sustainable investors can affect asset prices in stock and bond markets, as well as in markets for

venture capital and private equity. The reviewed studies differ substantially, however, in terms of the reported effect size.

Two empirical studies that investigate sustainability preferences in stock markets come to opposing conclusions regarding the effect on share prices. Hong and Kacperczyk (2009) examine the effect of investors excluding "sin stocks", such as tobacco, alcohol, and gambling, from their portfolios. They show that sin stocks have depressed prices and exhibit outperformance of 2.5 percent per year, relative to comparable stocks. This result implies that the moral aversion of investors for sin-stock companies has decreased the stock prices of these companies. At the same time, a related study focusing on the effects of divestment in the context of the South Africa boycotts of the 1980s concludes that these divestments had no discernible effects on asset prices (Teoh, Welch, and Wazzan, 1999).

Recent studies on green bonds—that is, bonds that are issued to finance projects with environmental benefits—indicate that the sustainability preferences of investors can influence bond prices. Baker, Bergstresser, Serafeim, and Wurgler (2018) find that at issue, the yields of green bonds are on average 0.06 percent below the yields of comparable bonds. They present supporting evidence that the observed differences are caused by nonfinancial preferences of investors. Similarly, Zerbib (2019) shows that sustainability preferences of investors result in a negative yield premium of 0.02 percent for green bonds. Hachenberg and Schiereck (2018) too confirm that green bonds are traded with a negative yield premium, while Tang and Zhang (2020) do not find evidence of such a premium.

Studies on the role of demand in public equity markets confirm that shifts in investor demand can influence stock prices, even when fundamental values remain unchanged. A large set of studies makes use of the fact that, due to passive investors, the inclusion or exclusion of companies in or from popular indexes, such as the S&P 500 index, triggers substantial investments in or divestments from these firms (Beneish and Whaley, 1996; Chang, Hong, and Liskovich, 2015; Kaul, Mehrotra, and Morck, 2000; Lynch and Mendenhall, 1997; Shleifer, 1986; Wurgler and Zhuravskaya, 2002). All in all, these studies find that the observed sudden changes in demand do affect stock prices. Studies that make use of order books (Ahern, 2014), announcements of equity issues (Loderer, Cooney, and Van Drunen, 1991), or auction repurchases (Bagwell, 1992) come to similar conclusions.

There is, however, no consensus on how strongly changes in demand affect share prices. A useful measure for the influence of demand on prices is the price elasticity of

demand.[2] Highly negative elasticity values indicate little influence of changes in demand, whereas less negative values indicate a stronger influence of demand on prices. The results presented by Loderer, Cooney, and Van Drunen (1991), Kaul, Mehrotra, and Morck (2000), and Wurgler and Zhuravskaya (2002), as well as Ahern (2014), indicate elasticities of around -5 to -10. The studies by Bagwell (1992), Chang, Hong, and Liskovich (2015), and Shleifer (1986) indicate lower elasticities, of between -1 and -1.5.

Only a few studies investigate demand effects in private markets such as markets for private equity and venture capital. Gompers and Lerner (2000) show that a doubling of inflows of capital increases the valuation of new investments of venture capital funds by between 7% and 21%, corresponding to an elasticity of -5 to -14. This is in line with the findings of Diller and Kaserer (2009), who find that demand effects influence private equity funds' returns.

Taken together, the literature provides evidence that the capital allocation of sustainable investors can affect asset prices. However, it leaves open two important questions. First, there is no agreement on the size of the effect sustainable investors have on asset prices, making it difficult to judge whether the effect is material. Second, while there is evidence that the capital allocation of sustainable investors has affected asset prices in some cases, there is so far no evidence that such changes in asset prices have translated into changes in ESG practices. Nevertheless, the literature reviewed above highlights three specific determinants that increase the likelihood that sustainable investors' screening approaches lead to improvements in the quality of companies' activities.

Firstly, the effect of an investor's screening approach is likely to be higher if a large proportion of investors apply the same approach. The equilibrium models of Heinkel, Kraus, and Zechner (2001), Fama and French (2007), Gollier and Pouget (2014), and Luo and Balvers (2017) indicate that the total effect of screening approaches on asset prices, as well as the marginal effect per additional dollar involved, increases with the fraction of wealth commanded by investors that apply the same screening approach. Hence, the effect of an individual investor's decisions depends on how many other investors apply the same screening approach.

[2]Price elasticity of demand is defined here as $\Delta Q/\Delta P$, where Q is the quantity of the demanded good and P its price. As for stocks, supply curves are vertical and Q can be interpreted as excess demand (Wurgler and Zhuravskaya, 2002). Hence, a price elasticity of -10 implies that a 1% increase in prices leads to an 10% decrease in demand. Vice versa, an increase in demand of 10% would be associated with a 1% increase in prices. The elasticity of a flat demand curve would be negative infinity, and changes in demand would not affect prices at all.

Secondly, the effect of investors' screening approaches is likely to be higher for companies whose assets are not easily substitutable. The models of Heinkel, Kraus, and Zechner (2001) and Fama and French (2007) show that the capital allocation of sustainable investors has a stronger effect on the prices of assets whose returns are only weakly correlated with the market portfolio—that is, assets that are not easily substitutable. Counterbalancing sustainable investors' demand for these assets requires a higher deviation from an optimally diversified portfolio from neutral investors than is the case for stocks that have very close substitutes. Accordingly, both Wurgler and Zhuravskaya (2002) and Ahern (2014) find empirical evidence that stocks with low substitutability exhibit a lower price elasticity.

Thirdly, a screening approach is more likely to cause companies to improve their ESG practices if the costs for a company to implement the reforms required to conform to the requirements embodied in the screening are low. The models of Heinkel, Kraus, and Zechner (2001) and Gollier and Pouget (2014) point out that whether changes in asset prices induced by SI provide an incentive for companies to improve their ESG practices depends on the cost of the necessary reforms.

Affecting Growth

The capital allocation of sustainable investors may also affect the growth of companies by changing the financing conditions these companies face. In this way, sustainable investors may be able to alter the levels of activity of particularly sustainable or unsustainable companies. The literature studies two ways in which investors can change financing conditions for companies and highlights that the impact depends on further characteristics, such as company size, company age, and market maturity.

A direct way in which sustainable investors may enhance a company's financing conditions is by providing capital on concessionary terms. Subsidizing companies that are deemed beneficial for development by providing them with financing that is more attractive than that available at market conditions is widely practiced by development finance institutions and other public sector actors and has been shown to enhance corporate investment (see, e.g., Cravo and Piza (2016); Kersten, Harms, Liket, and Maas (2017); Schreiner and Yaron (2001)). Brest and Born (2013) and Brest, Gilson, and Wolfson (2018), as well as Chowdhry, Davies, and Waters (2019), argue that private investors too can promote sustainable companies by providing them with capital on con-

cessionary terms—that is, with better conditions than these companies would obtain from preference-neutral investors.[3]

A more indirect way for sustainable investors to alter financing conditions is by affecting the prices of a company's financial assets for the entire market. As discussed in the previous section, the capital allocation of sustainable investors can affect asset prices under specific conditions, most notably the condition that sustainable investors represent a substantial market share. In this way, capital allocation may not only create managerial incentives to change but also change the cost at which affected companies raise capital from other investors (see, e.g., Fischer and Merton (1984)). In an equilibrium model Beltratti (2005) shows that if investors underweight unsustainable companies, this can increase the cost of capital for these companies, reduce their investment activity, and thus decrease their market share.[4] However, while effects of capital allocation on asset prices and the cost of capital are supported in the empirical literature, associated changes in growth are not.

Regardless of how investors alter a company's financing conditions, the literature points to several company characteristics that determine whether a change in financing conditions translates into accelerated growth of company activities. A non-fundamental movement in stock prices, such as the one created by the demand of sustainable investors, only translates into corporate investment activity when the company depends on external capital to finance these investments (Baker, Stein, and Wurgler, 2003).

For seasoned publicly listed companies, stock prices seem not to have a substantial effect on corporate investment activity (Blanchard, Rhee, and Summers, 1993; Morck, Shleifer, Vishny, Shapiro, and Poterba, 1990). Accordingly, Hadlock and Pierce (2010) find that financing constraints decrease with increasing size and age of companies. While a number of studies show that large companies with good ESG ratings enjoy a lower cost of

[3] Brest and Born (2013) as well as Brest, Gilson, and Wolfson (2018) point out that investors can make concessions in different ways. Brest and Born (2013) list six ways in which sustainable investors can provide capital to companies at conditions these companies would not otherwise enjoy: (1) making investments at below market conditions, (2) providing loan guarantees, (3) taking subordinated debt or equity positions, (4) accepting longer terms before exit, (5) providing flexibility in adapting investments, and (6) identification of investment opportunities that the market fails to notice, which, as Brest, Gilson, and Wolfson (2018) argue, is associated with concessions in the form of increased due diligence and monitoring costs.

[4] In fact, changes in the cost of capital caused by sustainable investors' capital allocation can also be viewed as a concession. As illustrated by the model of Beltratti (2005), increased costs of capital for unsustainable companies are associated with decreased returns for sustainable investors.

capital (e.g., Chava (2014)), it is ambiguous whether this is due to investor demand—and, thus, investor impact—or to the superior risk characteristics of those companies.

In contrast, a series of empirical studies show that small firms and young firms as well as firms operating in less mature financial markets with weak institutions are more likely to be restricted in terms of their growth by the cost of external financing (Beck, Demirgüç-Kunt, Laeven, and Maksimovic, 2006; Beck, Demirgüç-Kunt, and Vojislav, 2008; Bloom, Mahajan, McKenzie, and Roberts, 2010; Rajan and Zingales, 1996). Especially in developing countries, many small and medium-sized companies lack any access to external financing (Beck, Demirgüç-Kunt, Laeven, and Maksimovic, 2006). The finding that many small firms are restricted by the cost of capital or even access to capital is consistent with the finding that most small companies use retained earnings, insider finance, and trade credit to finance their investments (Berger and Udell, 1998; Carpenter and Petersen, 2002). Financing constraints seem to have a particularly strong inhibiting effect on entrepreneurial activities. Evans and Jovanovic (1989), as well as Holtz-Eakin, Joulfaian, and Rosen (1994), show that wealthy individuals are much more likely to become successful entrepreneurs.

Taken together, there is only partial empirical evidence that capital allocation by sustainable investors can enhance the growth of sustainable companies. Nevertheless, the literature points to determinants on which such an impact likely depends. The first determinant is whether the investor changes the company's financing conditions, either by lowering the cost of capital or improving access to finance. Additional determinants arise from the fact that a change in financing conditions only affects the level of company activity when the company's growth is constrained by external financing conditions. Company size and age and the maturity of the financial market in which the company's financial assets are traded all tend to reduce the degree to which growth is constrained in this way. Thus, capital allocation is more likely to affect growth for young, small firms in immature markets than for large, established firms in mature financial markets.

II.3.3 Indirect Impacts

Next we focus on a series of indirect mechanisms of investor impact. Via these mechanisms, investors do not directly affect companies' level of economic activities, A_c, or the quality, q_c, of these activities. However, investors' activities may affect other stakehold-

ers' behavior, including that of consumers, regulators, or other investors, which in turn may affect corporate behavior.

Stigmatization

Investors can stigmatize a company by divesting that company's assets or categorically excluding it from their portfolio. Apart from the impact through capital allocation this might have, the action can also impact other relevant stakeholders of the company. For example, people might be deterred from working at a company that is excluded by investors. Literature on the impact of such stigmatization, however, is thin. In a detailed assessment of the carbon divestment movement, Ansar, Caldecott, and Tilbury (2013) postulate that one of its most important impacts might be the stigmatization of the fossil fuel industry. For the anti-apartheid divestment campaign, there is anecdotal evidence that the campaign helped to elevate the issue of apartheid on the political agenda. Desmond Tutu, South African archbishop and an important figure in the struggle against the apartheid regime, commented that the disinvestment campaign in the U.S. added punch to the political struggle (Knight, 1990). However, we were not able to find studies that analyze to what extent exclusion decisions made by sustainable investors have led to stigmatization.

Endorsement

Investors can endorse companies for their social or environmental performance by including them in their portfolio or sustainability index. Such endorsement may help to increase the visibility and improve the reputation of a company, indirectly helping it to gain customers or motivate employees. We were not able to identify studies that analyze to what extent company reputation and sales were improved as a consequence of investor endorsement.

There are two studies, however, that investigate whether companies that were included in a sustainability index decided subsequently to communicate their inclusion to stakeholders (Carlos and Lewis, 2018; Searcy and Elkhawas, 2012). The fact that companies communicate index inclusion suggests that such inclusion helps to improve reputation. The studies show, however, that nearly half of the companies that were included in the Dow Jones Sustainability Index chose not to communicate their inclusion publicly. Carlos

and Lewis (2018) find that companies are more likely to remain silent about their index membership when they have a strong reputation for ESG performance already. Thus, one important determinant of the endorsement effect seems to be a company's prior ESG reputation.

Benchmarking

SI is feeding a growing industry of ESG rating agencies (Eccles and Stroehle, 2018). These rating agencies develop standards, create ESG benchmarks, and request increasing amounts of data from companies. The growth of this industry may encourage companies to report on their ESG practices in order to satisfy these increasing demands for data. Measuring and reporting may, then, also induce companies to improve their performance, for example because companies are benchmarked against peers, or simply because measuring ESG performance indicators also enables companies to manage their ESG performance.

The literature provides no evidence of investors' indirect impact exercised via their support for ESG rating agencies. This because even though it is fairly obvious that ESG rating agencies exist due to a demand from sustainable investors, it is not clear whether additional investors buying or using ESG ratings will further strengthen the impact of these agencies' benchmarking activities.

A number of studies have, however, investigated the direct impact of ESG ratings and ESG reporting standards on companies' social and environmental performance. Regarding standards, one study concludes that the introduction of the voluntary ISO 14000 standard for environmental management has led firms to improve their environmental outcomes (Melnyk, Sroufe, and Calantone, 2003). Another study, however, concludes that the adoption of this standard had no discernible effect on environmental outcomes (Hertin, Berkhout, Wagner, and Tyteca, 2008). Thus, the mere existence of ESG standards may not suffice to improve outcomes, even though it must be borne in mind that ISO 14000 is just one of many different standards in the ESG domain.

Studying ESG ratings, Chatterji and Toffel (2010) provide evidence that companies improve environmental performance in response to receiving a low ranking in an environmental benchmark. They find this to be especially the case when the cost of the necessary reforms is low, and when the company operates in a highly regulated industry. A problem with this effect, however, is that there are remarkable differences between the ESG

benchmarks compiled by different agencies (Chatterji, Durand, Levine, and Touboul, 2016). Due to these differences, Chatterji, Durand, Levine, and Touboul (2016) conclude that "SRI ratings will have a limited impact on driving rated firms toward any particular shared behaviors." One determinant of the effectiveness of the impact of benchmarking is thus the consistency of ESG benchmarks—the more consistent ESG benchmarks are, the greater is their effect on company activities.

Demonstration

Besides affecting companies' impact through their own activity, SI investors may also encourage other investors to do the same. We identify two mechanisms via which sustainable investors may be able to do so.

Firstly, sustainable investors may help to establish SI as a social norm (see, e.g., Gilbert, Fiske, and Lindzey (1998)). Research on charitable giving has shown that potential donors are more likely to give if they learn that others give as well (DellaVigna, List, and Malmendier, 2012; Frey and Meier, 2004; Shang and Croson, 2009). Whether this mechanism applies in the context of SI, however, has not been investigated.

Secondly, investors pioneering novel investment projects, for example financing pioneering renewable energy projects in a developing country, may increase the subsequent flow of capital into such projects. For example, Egli, Steffen, and Schmidt (2018) show that learning and associated efficiency gains within the renewable energy finance industry have reduced the cost of capital for renewable energy projects over time. Geddes, Schmidt, and Steffen (2018) argue that enabling financial sector learning is an important way in which development banks encourage private investments in low-carbon energy generation. Thus, early investors in novel approaches or markets may have an impact not only through their own investment, but also through subsequent investments that they facilitate. Since in these cases investors are one player in an evolving ecosystem involving technology providers, service providers, and regulators, it is very difficult to separate investor impact from contemporaneous factors.

II.4 Discussion

Table II.1 provides an overview of the reviewed mechanisms of investor impact. Shareholder engagement emerges as most reliable in the sense that an effect of investor activity

on company activity has been demonstrated empirically. Several studies show that shareholder engagement can lead to measurable improvements in companies' ESG practices. The impact of shareholder engagement increases with the influence of the investor engaging and the ESG experience of the company engaged with. It decreases as the cost of the requested reforms rises, meaning that shareholder engagement is more likely to trigger incremental improvements rather than transformative change.

Capital allocation has not been studied in its entirety, but important parts of the mechanism are empirically demonstrated, both for incentivizing improvements and for affecting growth. Capital allocation emerges from the literature review as a somewhat less reliable mechanism since we found no study that establishes a direct link between capital allocation by SI investors and a change in company activities. However, key parts of the mechanism have been studied and the results indicate that capital allocation could bring about investor impact in two different ways. First, screening approaches may incentivize companies to adapt their practices. There is evidence that screening can affect asset prices; there is, however, no evidence to date that such a change in asset prices has indeed led companies to improve the quality of their activities. The likelihood that screening approaches have such an effect increases with the market share of the investors applying the same screening approach, and decreases with the substitutability of the securities that are excluded. A further determinant is the cost of the reform that is required if a company is to evade the screen. Second, investing in sustainable companies may increase the level of company activity. There is evidence that improved financing conditions can accelerate a firm's growth, but only when financing is a limiting factor. The likelihood of this impact increases with the improvement in financing conditions that the investor provides to the company compared to the status quo. It decreases with the age and size of the company, as well as with the maturity of the financial market in which the company's financial assets are traded.

Indirect impacts are mostly unproven given a lack of empirical studies that indicate their effectiveness. While there is anecdotal evidence for indirect impacts, none of the indirect impact mechanisms have been analyzed comprehensively—in the sense that investor activities have been related to a change in company activities. There is no empirical evidence for the effects of stigmatization and demonstration. There is some empirical evidence, however, for endorsement and benchmarking, yet it covers only part of the mechanisms, so that the extent of investor impact remains a matter for speculation. In terms of determinants, there are indications that an investor's endorsement is more

valuable when the endorsed company has a poor prior ESG reputation. Benchmarking would likely be more effective were different ESG benchmarks consistent—that is, if they identified the same laggards and leaders. It is important to note that a lack of evidence for indirect impacts does not imply that indirect impacts are irrelevant. Indirect impacts could be important under the right circumstances, but so far the academic literature does not provide evidence for their effectiveness.

Applying these findings to today's USD 30 billion market for SI (GSIA, 2018) suggests that the bulk of SI assets are invested in ways that promise rather modest and perhaps even negligible investor impact. Shareholder engagement, identified as the most reliable mechanism, is practiced for only 18 percent[5] of global SI assets, and for a mere 10 percent in the US.

About 50 percent of assets are invested via screening approaches, and 32 percent rely on ESG integration. However, there is a lot of diversity in screening approaches and inconsistency between different ESG ratings. This diversity means that even though the combined market share of these approaches is substantial, the effective market shares behind specific approaches are small. This dilutes any effect on asset prices and with it the incentives for companies to implement reforms. In addition, some of the most popular screening approaches exclude industries rather than practices, meaning that companies in affected industries are barely incentivized to improve at all as even if they were to improve, they would not have the opportunity to conform to the investment screen.

According to GSIA (2018), a mere 1 percent of global SI assets is invested in so-called impact investing, where accelerating the growth of sustainable companies is a key objective. Such investments often include concessions to the investee, and are placed in companies that have limited access to financing. Other SI approaches may also have an impact on the growth of sustainable companies; this is, however, less likely due to the fact that over 80 percent of SI assets are invested in publicly listed firms, where financing is usually a less important constraint for growth.

In addition, SI may have indirect impacts. For example, it may be that the fossil fuel divestment campaign stirs political and societal debate around fossil fuel consumption, and perhaps ultimately leads to less fossil fuel consumption. Similarly, it may be that ESG ratings drive companies to implement more ESG practices. Most promising, perhaps, demonstrating the feasibility of investments in novel approaches or technologies may

[5]This and the percentages that follow are based on GSIA (2018).

trigger subsequent investments. So far though, there is no empirical evidence that investor activity is actually driving such developments.

II.4.1 Implications

Taken together, our results suggest that the investor impact of SI as it is practiced today is rather modest. At the same time, they hold a number of implications for investors, rating agencies, and policy makers with regard to how investor impact can be increased.

For Investors

First, investors who want to stimulate real-world impact can roll out shareholder engagement throughout their portfolios. Ideally, investors should focus on requests that have a good chance of success and yield substantial improvements in company impact. In addition, investors can pool their shareholder rights with like-minded investors to increase their influence, and outsource the engagement mandate to specialized firms.

Second, investors can allocate capital to companies with a positive company impact that are constrained in terms of their growth by external financing conditions. While financing constraints are less likely for large, well-established companies, many small and young companies are constrained by external financing conditions, especially in less mature financial markets—in developing countries for example. By easing the financing constraints of such companies, investors can support the growth of sustainable businesses.

Third, investors can screen out investments based on the absence of ESG practices, focusing on a few specific and transparently communicated practices. The most promising practices can be implemented by companies at a low cost and result in substantial improvements in company impact. The incentives for companies to adopt the demanded ESG practices increase with the share of investors applying the same screening approach. Hence, screening approaches should be pursued in a large coalition of investors and encompass stocks and bonds. The way in which the Institutional Investor Coalition on Climate Change communicates its members' expectations for specific sectors may constitute a step in this direction.

Fourth, investors who are convinced that they can have indirect impacts should attempt to demonstrate them. SI funds could provide examples or look at intermediate

proxies that make indirect impacts more tangible. For instance, investors could measure the level and the tone of media attention in response to an exclusion announcement as a proxy for stigmatization. Similarly, fund managers who launch an innovative product could track the uptake of their innovation as a proxy for demonstration.

For Rating Agencies

Rating agencies and ESG data providers could play an important role in changing the SI industry by developing investor impact metrics. Lately, several ESG data providers have released company impact metrics, which is an important step toward us understanding better how company activity can support the achievement of the SDGs. However, these metrics are often used to indicate the impact an investor has when investing in a certain fund, which is misleading. For example, a fund that holds companies that have implemented best practices regarding their greenhouse gas emissions would receive a top rating based on company impact metrics. However, a fund that holds companies with mediocre practices and drives them to upgrade these practices would receive only a mediocre rating.

Investor impact metrics should reflect the temporal change in company impact that investors can expect to cause. There is some activity in this direction,[6] and the need for investor impact metrics is also recognized by regulators. This paper provides an overview of the mechanisms and determinants of investor impact and thus provides a blueprint of how investor impact metrics might look.

For Policy Makers

Policy makers should be aware that without additional policy measures SI is unlikely to result in the dramatic transformation that is required, for example, for the decarbonization of the economy. SI seems to be well suited to diffusing ESG practices throughout industries, ensuring that all low-hanging fruit are harvested. SI can also support innovation by promoting the growth of young and small sustainable businesses. Yet more fundamental changes also require policies that directly change the viability of economic activities, such as taxes on pollution, limiting emission allowances or minimum standards (see, e.g., Chesney, Gheyssens, Pana, and Taschini (2016)). Rather than making such

[6]See, for example, https://realimpacttracker.com and https://impactmanagementproject.com/.

policies unnecessary, SI may be a suitable complement in that it incentivizes companies to adopt ESG practices and explore business models made viable by these policies.

There can be a positive feedback loop by which regulators enact policies that make certain ESG practices superior in terms of financial performance and investors encourage ESG practices that anticipate future regulation. Many SI proponents argue that material ESG aspects, where doing good coincides with superior returns, will drive positive change. However, materiality ultimately rests on the argument that ESG practices result in superior financial performance, which often depends on policy. For example, investment in water-saving technologies is not a "material ESG practice" in a situation where water tariffs are too low to justify the investment. But it can become material once regulators begin to raise tariffs to appropriate levels.

II.4.2 Limitations

We acknowledge that investors may have a variety of motivations to engage in SI, and that the desire to make an impact may not be important for all SI investors. For example, investors may have the desire to be morally aligned with their portfolio. If so, excluding certain industries that are perceived as "dirty" may be perfectly consistent with an investor's motivation, even if it has no impact on the excluded industries (see, e.g., Haidt (2007)). Other investors may engage in SI out of a financial motivation—expecting less risk for example. If so, integrating certain ESG factors into investment decisions is perfectly consistent with that motivation (see, e.g. Friede, Busch, and Bassen (2015)). In both of these cases, investor impact is irrelevant to investor motivation. However, we argue that from a societal perspective investor impact is the essential feature of SI. If SI does not make a difference, policy makers would have no reason to foster it, and academics would have little reason to study it.

This review is limited to published academic results, which may not fully capture all relevant aspects of investor impact. The academic literature is biased toward publicly listed corporations and stock markets, due to data availability. Accordingly, this literature review is also somewhat biased toward public stock markets. And there are, potentially, further relevant impact mechanisms in specific financial markets—such as corporate debt, private equity, bank lending, and real estate—that are not reflected here. Also, aspects that are difficult to measure or are currently unfolding may not be represented appropriately. For example, there are increasing numbers of alliances of asset owners promoting

sustainable finance, including the Investor Coalition on Climate Change. The effects of these alliances have not yet been researched. Also, novel ideas such as blended finance, where private and public investors combine funds, may bring about significant investor impact, but have not yet been researched. Thus, a key limitation of this study is that there may be important investor impact mechanisms that have not yet been covered by academic research.

II.4.3 Future Research

This paper concludes that shareholder engagement is a promising way to ensure investor impact. An important question that remains, however, is how to quantify the impact of engagement activities in a comparable way. Existing studies have quantified the success rate of engagement requests, but it is also necessary to quantify how substantial an engagement request is (Barko, Cremers, and Renneboog, 2017). One substantial request may have a greater impact than several superficial requests. Combining the aforementioned success rate with a measure of how substantial a request is could yield a comparable measure. Such a standard for reporting the impacts of engagement activities could make shareholder engagements comparable and also more visible and marketable.

Regarding the impact of capital allocation, one critical knowledge gap is that there is currently no empirical study that relates capital allocation decisions made by sustainable investors to corporate growth or to improvements in corporate practices. Hong and Kacperczyk (2009) point out that while their study demonstrates an effect on the share prices of tobacco companies, it does not investigate the effects on the activities of tobacco companies. Studies that not only relate SI to asset prices but also investigate the response of affected companies in terms of management and investment decisions would advance our understanding of investor impact decisively. Such studies would be a first essential step toward developing a metric for the impact of capital allocation.

Regarding indirect impacts, there is a need for studies that either investigate the entire causal chain of indirect impacts or inform critical pieces that are currently missing. It would be valuable to conduct, for example, a case study of the Fossil Free divestment campaign, to establish the consequences of the campaign in terms of media attention, investor behavior, and corporate decisions, as well as to relate it to broader economic, political, and cultural dynamics around the fossil fuel industry. Such a study would close

important gaps in the scientific understanding of the investor impact of stigmatization and would provide guidance as to how and when to pursue it.

II.5 Conclusion

SI is increasingly thought of as a mechanism for achieving societal goals such as the United Nation's Sustainable Development Goals. We observe, however, that in both research and practice the notion of investor impact is neglected, and conduct a literature review on the mechanisms of investor impact. We conclude that bringing investor impact to bear through shareholder engagement is a relatively reliable mechanism, that capital allocation is likely to have an impact on companies whose growth is limited by external financing conditions or if investors screen out companies based on the absence of specific ESG practices, and that indirect impacts remain unproven regarding their effectiveness. Our results suggest that the current practice of SI has only a modest investor impact, and call for the development of investor impact metrics that reflect the contribution of SI to societal goals.

III How Do ESG Ratings Affect the Economy?

The Economic Impact of ESG Rating Changes*

Florian Berg, Florian Heeb, Julian F. Kölbel

A version of this paper has been released on SSRN and is available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4088545.

Abstract

This study examines the impact of ESG rating changes on mutual fund holdings, stock returns, corporate investment, and corporate ESG practices, using panel event studies. Looking specifically at changes in the MSCI ESG rating, we document that rating downgrades reduce ownership by mutual funds with a dedicated ESG strategy, while upgrades increase it. We find a negative long-term response of stock returns to downgrades and a slower and weaker positive response to upgrades. Regarding firm responses, we find no significant effect of up- or downgrades on capital expenditure. We find that firms adjust their ESG practices following rating changes, but only in the governance dimension. These results suggest that ESG rating changes matter in financial markets, but so far are having only a limited impact on the real economy.

III.1 Introduction

Investors are increasingly integrating information on environmental, social, and governance (ESG) issues into their investment decisions. Over 4,000 asset managers and asset owners, representing over USD 100 trillion in assets, have signed the Principles for

*We thank Marc Chesney, Thorsten Hens, Andrew King, Philipp Krüger, Stefano Ramelli, Roberto Rigobon, Emil Verner and the participants at the Geneva Finance Research Institute Seminar.

Responsible Investment (PRI), the first of these principles requiring investors to incorporate ESG issues into investment analysis and decision-making processes. According to the Global Sustainable Investment Alliance (GSIA, 2021), the volume of investment products explicitly integrating ESG information exceeded USD 25 trillion in 2020 and is destined to grow further. ESG ratings, which provide an aggregated assessment of firms' sustainability performance based on a multitude of ESG indicators, serve as a primary information source for investors implementing ESG integration.

Theoretical models predict that investor demand for ESG investments could have profound effects on financial markets and corporate behavior and may create positive social impact (Broccardo, Hart, and Zingales, 2020; Heinkel, Kraus, and Zechner, 2001; Hong, Wang, and Yang, 2021; Landier and Lovo, 2020; Oehmke and Opp, 2019; Pástor, Stambaugh, and Taylor, 2021). In broad strokes, these models make three major predictions. First, investors with a preference for "green" companies—so, companies with high ESG performance—tilt their investment holdings toward them and away from "gray" companies, which have low ESG performance. Second, if there are enough such investors, this tilt drives up the asset prices of green companies, decreasing expected returns and consequently the cost of capital for green companies. Vice versa, the cost of capital for gray companies increases. Third, this change in asset prices should trigger changes in the real economy through two key channels: On the one hand, via a growth channel, where shifts in the cost of capital influence real investments—green firms would invest more and grow faster, gray firms would invest less and grow at a slower pace. On the other hand, via a reform channel, where firms improve their ESG practices with the goal of attracting green investors and increasing their stock market valuation. Both channels would ultimately lead to a positive social impact on the real economy. Thus far, however, it remains unclear to what extent these predictions hold empirically.

In this study, we investigate the impact of ESG rating changes on financial markets and corporate behavior. For our sample of 3,665 listed U.S. companies, we observe 4,679 ESG rating changes between 2013 and 2020. We focus on the MSCI ESG rating, one of the most influential of its kind.[1] Given that ESG ratings are an essential information source for ESG investing, we expect ESG rating changes to have a measurable impact on investment choices. We employ a panel event study methodology (Clarke and Tapia-Schythe, 2021; Freyaldenhoven, Hansen, Pérez, and Shapiro, 2021; Schmidheiny and Siegloch, 2019),

[1] A report by Opimas states that MSCI has the largest market share, with 31%: http://www.opimas.com/research/742/detail/.

based on three considerations. First, this allows us to make causal inferences on the effects of ESG rating changes, given that the identifying assumption of parallel trends for treated and untreated firms holds. Second, we can separately analyze the effect of up- and downgrades, which may not be symmetric. Third, rather than making restrictive assumptions on how and when the effects of ESG rating changes unfold, we can estimate and visualize dynamic treatment effects.

First, we show that mutual fund holdings react to ESG rating changes. We investigate ESG ownership—so, the fraction of a firm's shares that are owned by domestic U.S. mutual equity funds with an explicit ESG objective. In September 2020, these funds collectively represented USD 71 billion in assets under management and owned 0.03% of the average company in our sample. ESG ownership reacts significantly to both ESG rating upgrades and downgrades. Two years after a downgrade, ESG ownership is on average 13.1% lower; two years after an upgrade, it is 17.1% higher than one month before the rating change. The response to rating changes is slow and gradual: ESG mutual funds adjust their holdings gradually over the two years following ESG rating changes.

Second, we show that ESG rating changes have long-lasting effects on stock returns. We detect significant negative effects on buy-and-hold returns, for holding periods of up to 24 months, after a downgrade, reaching a maximum of -3.78% for a holding period of 19 months. After upgrades, we detect a weaker and slower positive effect on buy-and-hold returns, reaching a maximum of 2.62% for a holding period of 22 months. Similar to our results on ESG ownership, the effect of ESG rating changes on buy-and-hold returns unfolds gradually over a prolonged time period.

Third, we explore how changes in ESG ratings affect corporate behavior, both via a growth and via a reform channel. We do not find evidence in support of the growth channel. ESG rating changes have no discernible effect on the subsequent level of firms' capital expenditure over a two-year time horizon. We find some evidence for the reform channel. We exploit the fact that MSCI ratings separately measure a firm's management of ESG issues and a firm's exposure to these issues. For each of the environmental, social, and governance dimensions, we construct a score solely reflecting firms' management practices. For the environmental and social management score, we do not find any significant reaction to ESG rating changes. For the governance score, we find that firms react to both upgrades and downgrades: following a downgrade, firms improve their governance practices; following an upgrade, firms tend to let their governance practices deteriorate.

Our study provides an empirical assessment of the economic impact of ESG rating changes. Our results are compatible with model predictions in that they confirm that changes in ESG performance affect the holdings of ESG-aware investors and stock returns. We find that, at least for ESG mutual funds, ESG commitments are not merely "cheap talk" and that ESG ratings do affect stock markets. However, the real economic impact of ESG investing seems limited, given our results. First, the cost-of-capital advantage (disadvantage) for green (gray) firms does not seem to affect firms' real investment. Thus, we do not find evidence for a growth channel that might allow green firms to outgrow gray firms. Second, we only find evidence for the reform channel in the governance dimension, not in the social or the environmental dimension. Thus, the reform channel seems to function only for a subset of ESG issues. One speculative explanation for this is that reform of governance practices costs less and is less controversial than reforms in the environmental or the social domain.

Our paper contributes to the literature examining the effects of preferences for sustainability on financial markets. While there is a rich body of theoretical work on these effects (Broccardo, Hart, and Zingales, 2020; Gollier and Pouget, 2014; Heinkel, Kraus, and Zechner, 2001; Hong, Wang, and Yang, 2021; Landier and Lovo, 2020; Oehmke and Opp, 2019; Pedersen, Fitzgibbons, and Pomorski, 2021; Pástor, Stambaugh, and Taylor, 2021), our study adds to a relatively recent literature strand that tests theoretical predictions. Specifically, it contributes to three strands of the empirical literature.

First, our study contributes to empirical work investigating how preferences for sustainability affect investors' holdings and corporate ownership structures. Based on holdings data, which the authors connect with behavioral experiments, Riedl and Smeets (2017) show that pro-social preferences and social signaling explain why a considerable proportion of private investors hold sustainable mutual funds, financial motives being, comparatively, less influential. In line with this, Hartzmark and Sussman (2019) show that mutual funds labeled as sustainable attract inflows. At the same time, both Gibson, Glossner, Krueger, Matos, and Steffen (2021) and Kim and Yoon (2020) find that, at least in the US, portfolio-level ESG scores of PRI signatories hardly deviate from those of non-signatories, casting doubt on how strongly investors' sustainability preferences translate into investment decisions. Our study adds to this literature by showing that the holdings of mutual funds with an explicit ESG mandate react significantly to changes in ESG scores. Thus, we document that there is a pool of money that responds to ESG rating

changes, but that this pool of money is several magnitudes smaller than the combined assets of PRI signatories.

Second, our study contributes to the literature investigating the effects that taste and diverging beliefs have on asset prices. A broad literature on investor sentiment shows that price effects induced by non-rational beliefs about future cash flows and investment risks may not be fully eliminated by arbitrage (Baker and Wurgler, 2007; Barberis, Shleifer, and Vishny, 1998; Shleifer and Vishny, 1997). Related studies investigating the effects of inclusions, exclusions, or reweighting in popular stock market indices show that non-fundamentally driven changes in the demand for certain companies can have persistent stock price effects (Chang, Hong, and Liskovich, 2015; Kaul, Mehrotra, and Morck, 2000; Wurgler and Zhuravskaya, 2002). Very recently, Greenwood, Laarits, and Wurgler (2022) have shown that retail investors' preference for certain stocks exert strong price effects following fiscal stimuli. Thus, preferences for ESG performance may also have stock price effects. Indeed, several recent studies find that changes in ESG ratings affect asset prices on time scales of a few days to a few months. Glück, Hübel, and Scholz (2021) show that downgrades in the environmental and social dimensions of an MSCI ESG rating lead to negative abnormal stock returns within 11 days of a rating change. Shanaev and Ghimire (2022) also document an effect of MSCI rating downgrades over a one-month period. Rzeznik, Hanley, and Pelizzon (2022) show that a profound methodology change in Sustainalytics ESG ratings led to transitory price pressure for firms whose ratings changed the most, pressure that subsided five months after the event. We add to this literature by showing that ESG rating changes seem to have relatively persistent effects on share prices, which is arguably an important condition for knock-on effects on firms' capital expenditure or efforts to reform ESG practices.

Third, we contribute to empirical studies investigating the effect of ESG investing on corporate behavior in the real economy. Addressing the growth channel, Berk and van Binsbergen (2021) argue that the effect of investors' ESG preferences on the cost of capital is negligible and that the inclusion or exclusion of a firm from a leading ESG index has no detectable effect on that firm's investment decisions. In contrast, Briere and Ramelli (2021) find that in quarters with higher non-fundamental demand for green assets, green firms increase both their capital investments and cash holdings. Our study provides an additional indication that improved ESG performance and the associated interest from ESG investors does not have a measurable effect on the level of corporate investment.

Regarding the reform channel, our study is closely related to that of Gantchev, Giannetti, and Li (2022), who show that firms improve their environmental and social practices following negative news coverage of these practices. While the authors document responses in the environmental dimension that are economically large, we detect no response in the environmental dimension. We only detect a response in the governance dimension, and one that is relatively small. These diverging findings can be due to many factors, including sampling (they study an international sample, as opposed to our U.S. sample), methodological differences, and data sources. While we also provide some support for the effectiveness of the reform channel, our conclusion is more cautious. Based on our results, we conclude that although ESG rating changes have a visible impact in financial markets and one that corresponds to theoretical predictions, these changes' impact on the real economy is barely detectable, at least so far.

III.2 MSCI ESG Rating Changes

In our study, we investigate the economic impact of changes in the MSCI ESG rating. MSCI is a leading provider of ESG ratings in the U.S. market.[2] The MSCI ESG rating assesses companies' management of financially relevant sustainability risks and opportunities along three dimensions: environment (E), social aspects (S), and governance (G). Beneath those three dimensions, MSCI computes scores for 37 "key issues" (e.g., Carbon Emissions, Health and Safety, or Corruption and Instability), basing these scores on a wide variety of indicators sourced from corporate disclosure, internal modeling, and news reports.[3] MSCI evaluates firms' performance with regard to these issues on the one hand based on the firm's exposure to risk (exposure score), and on the other based on the firm's capability to manage risk (management score). For example, a company's exposure to emissions regulation would enter the exposure score, while its capability to drive down and manage emissions would enter the management score. The exposure score and the management score are combined in the issue score. Performance across all issues is aggregated in a raw ESG score, weighted based on MSCI's view of the financial materiality of the different issues. This raw ESG score is then benchmarked against the raw ESG scores of industry peers, resulting in a peer group–adjusted final ESG score. This score ranges from zero to ten, where zero indicates poor management of ESG risks.

[2]As concluded, e.g., in a report by Opimas: http://www.opimas.com/research/742/detail/.
[3]For details, see https://www.msci.com/documents/1296102/21901542/ESG-Ratings-Methodology-Exec-Summary.pdf.

Based on this final score, MSCI assigns an ESG rating that sorts companies into seven categories ranging from CCC to AAA, similar to the scales used by credit rating agencies. These seven rating categories are equally spaced with regard to the underlying ESG score; there are six predefined thresholds defining a company's ESG rating. MSCI updates its ESG scores continuously, but does not follow a strict reassessment schedule. When a company's ESG score passes a threshold, its ESG rating changes. MSCI ESG ratings can change for several reasons. First, when a company's ESG practices improve or deteriorate in the view of MSCI analysts. Second, when a company's exposure to specific ESG issues is believed to have increased or decreased. Third, when MSCI's view of the financial relevance of one issue relative to another changes. Fourth, since the final score is benchmarked within industries, when the ESG performance of industry peers changes. Of course, ratings can also change due to any combination of these four factors.

We observe, in a sample of 3,665 listed U.S. corporations, 4,679 rating changes that take place between February 2013 and September 2020. Figure III.1 shows the distribution of underlying changes in the numerical ESG score for both up- and downgrades. We observe 2,545 ESG rating upgrades and 2,133 downgrades. Table III.1 shows the descriptive statistics of all the firm-level variables that we use in our analysis.

III.2.1 Challenges of ESG Ratings

While ESG ratings are widely used in the financial industry, they are not without their critics. Several studies show that ESG ratings from different data providers diverge widely (Berg, Kölbel, and Rigobon, 2022; Chatterji, Durand, Levine, and Touboul, 2016; Gibson Brandon, Krueger, and Schmidt, 2021). According to Berg, Kölbel, and Rigobon (2022), this divergence can be explained mainly by two factors: differences in measurement (explaining 58 percent of the observed divergence) and difference in scope (explaining 38 percent of the divergence).

Differences in measurement question the quality of data used to determine ESG ratings. While it is difficult to objectively determine the "true" value for many underlying data points, including for firms' biodiversity impacts, there is considerable divergence even for data points where objective measurement standards are available, corporate greenhouse gas emissions being an example (Busch, Johnson, and Pioch, 2022). Poor data quality may impede ESG ratings' usefulness in measuring firms' ESG practices or their exposure to ESG risks. As an illustrative example, the two most commonly used

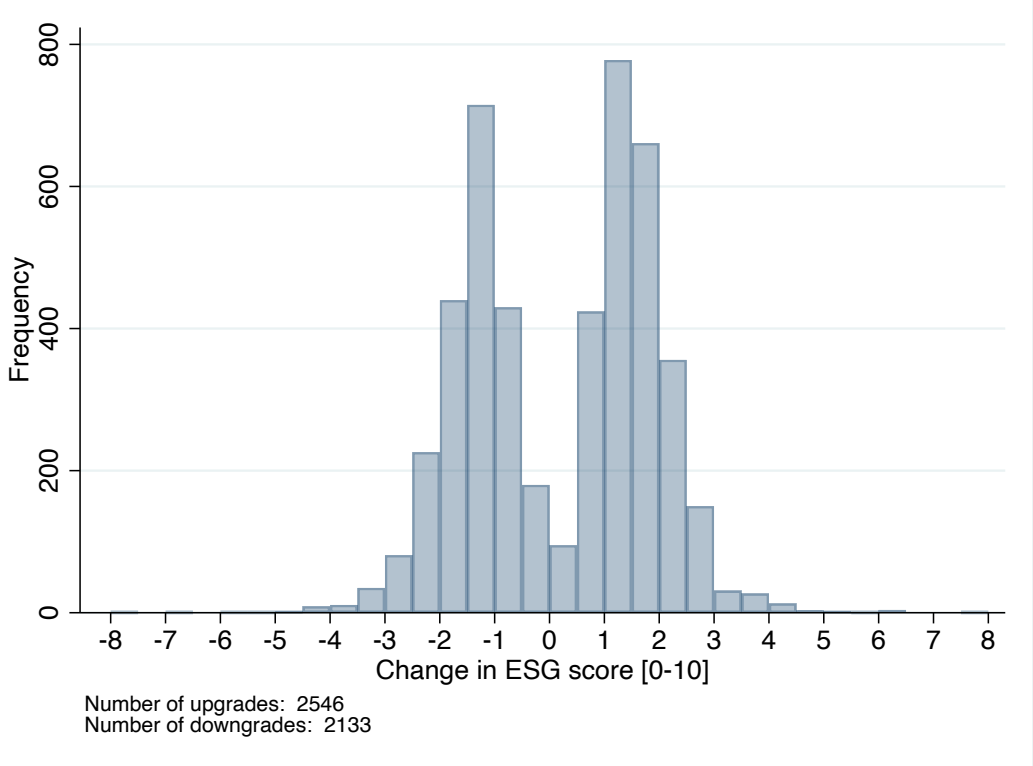


Figure III.1. The Size of ESG Rating Changes. This figure shows the distribution of the changes in the underlying ESG score for our sample of ESG rating changes. The ESG score changes are binned with a bin width of 0.5 score points. In total, we observe 2,546 ESG rating upgrades and 2,133 ESG rating downgrades.

ESG ratings, the MSCI ESG Rating and Sustainalytics' ESG Risk Rating, have both failed to capture the governance shortcomings that led to the collapse of Wirecard, a German digital payment processor.[4]

Differences in scope highlight that different ESG ratings are designed to measure different things and should be interpreted in the light of what they intend to capture. For example, the MSCI ESG Rating measures companies' ESG risks strictly relative to industry peers, while other rating agencies apply a sector-agnostic assessment. This has the consequence that, for example, Royal Dutch Shell currently receives a AA ESG rating (the second highest category), as it is considered a sustainability leader within the "integrated oil and gas" industry. At the same time, Shimano, a Japanese producer of

[4]As analyzed in the article "Anatomy of a Scandal: Wirecard Tests ESG," in the *Financial Times* (July 1, 2020).

Table III.1
Descriptive statistics for firm characteristics

This table shows descriptive statistics of the firm-level characteristics we use in our studies. Our sample consists of 3,665 listed U.S. firms with available MSCI ESG rating data. Observations cover the period from February 2013 to September 2020. *ESG ownership* is calculated as the fraction of a company's outstanding shares owned by ESG mutual funds. *ESG weight* is the fraction that a company's shares represent in the portfolio value of a synthetic ESG mutual fund that aggregates the holdings of all ESG mutual funds. We obtain monthly stock returns from the CRSP Monthly Stock database. *ESG score* is MSCI's assessment of how well companies manage sustainability risks relative to industry peers, on a scale from 1 to 10. *Environmental management score*, *social management score*, and *governance management score* measure a company's management practices, calculated as described in Section III.6.1. We calculate *momentum* as the firms' average stock return between twelve and two months prior to an observation. We estimate *market beta* from regressions of monthly returns in excess of the 1-month treasury bill rate on the excess market return. *Capex* is firms' capital investments divided by the one-quarter lagged book value of their property, plants, and equipment in percentage points. *Leverage* is firms' long-term debt plus debt in current liabilities, divided by total assets in percentage points. *Size* is the log of firms' market capitalization. *Book-to-market* is firms' book value of equity divided by the market valuation. *Profitability* is income before extraordinary items over total assets.

	p5	p25	mean	p50	p75	p95	sd	count
Firm-level variables (monthly observations)								
ESG ownership (%)	0.004	0.015	0.063	0.030	0.072	0.234	0.089	200993
ESG weight (%)	0.000	0.001	0.031	0.003	0.015	0.171	0.088	200993
Monthly return	-0.149	-0.045	0.007	0.008	0.059	0.161	0.098	178317
ESG score [0–10]	1.500	2.858	4.222	4.100	5.500	7.600	1.911	204995
Environmental management score [0–10]	0.000	1.700	3.351	2.717	4.300	8.000	3.083	150933
Social management score [0–10]	1.250	2.417	3.345	3.200	4.200	5.800	1.368	150921
Governance management score [0–10]	2.900	4.200	5.336	5.200	6.350	8.500	1.676	150933
Momentum	-0.043	-0.007	0.008	0.009	0.024	0.053	0.032	175060
Market beta	0.204	0.666	1.051	1.004	1.369	2.068	0.594	181847
Firm-level variables (quarterly observations)								
Capex	0.701	2.455	6.417	4.280	7.343	16.412	19.195	63862
Leverage	0.000	9.014	27.486	24.847	40.167	67.351	23.499	62640
Size	5.627	6.780	7.925	7.773	8.903	10.847	1.606	67716
Book-to-market	0.033	0.221	4.127	0.415	0.691	1.292	581.438	67669
Profitability	-5.051	0.106	0.963	0.755	1.824	4.359	79.992	67680

bicycle components, receives a subpar BB rating, as its ESG risks are thought to be elevated compared to peers in the "leisure products" sector, few of which own comparable production facilities. Also, the MSCI ESG Rating exclusively measures financial risks for investors, while some other ESG ratings extend to firms' effects on a broader set of stakeholders. This focus is illustrated by a recent upgrade of the environmental rating of fast-food chain McDonald's, one of the world's largest greenhouse gas emitters (considering supply chain emissions).[5] The rating was not upgraded because of improvements in the firm's management of its greenhouse gas emissions but because MSCI's analysts concluded that, given the current regulatory development, these emissions do not pose

[5]As reported in detail in the article "ESG Mirage," on the news platform Bloomberg (December 10, 2021).

an imminent financial risk to investors. Further, certain ESG issues may not be covered altogether by ESG ratings. For example, within the financial sector, ESG ratings barely cover the toxicity of financial practices or the potential of financial products for generating systemic risks (Seele and Chesney, 2017).

Thus, while the MSCI ESG Rating rating may provide useful information for investors that want to minimize ESG risks with a fixed sector allocation, it is questionable whether it is a good proxy for corporate externalities. Nevertheless, we focus our analysis on the MSCI ESG Rating as it is the most commonly used ESG rating among U.S. investors. Therefore, it can be expected to have the most substantial impact on financial markets if such impacts are prevalent.

III.3 Do ESG Rating Changes Affect Stock Ownership?

In this section, we explore how ESG rating changes affect the holdings of dedicated ESG mutual funds. To do so, we run a panel event study with a measure of ESG ownership, showing that companies' ownership by domestic U.S. mutual equity funds with an explicit ESG objective reacts to both ESG rating upgrades and downgrades.

III.3.1 ESG Ownership

We define the variable *ESG ownership* as the fraction of a company's outstanding shares that is owned by funds with an explicit ESG or sustainability strategy. We identify ESG mutual funds by screening mutual funds' names and strategies for a set of ESG-related keywords.[6] We restrict our search to domestic U.S. equity funds that follow a capitalization, growth, growth-&-income, or income-based strategy. For each company covered by our sample of ESG rating changes, we compute the aggregate number of shares owned by the identified ESG mutual funds based on the CRSP U.S. Mutual Fund Holdings database. We retrieve the total number of shares outstanding for each company in our

[6]Specifically, we define ESG mutual funds as funds that use any of the following letter sequences in their title or strategy description: "SRI," "social," "ESG," "green," "sustain," "environ," "impact," "responsible," "clean," and "renewable."

sample from the CRSP Monthly Stock database. We calculate *ESG ownership* by dividing the number of a company's shares owned by ESG mutual funds by the company's total number of shares outstanding. To reduce the effect of outliers, we trim the variable *ESG ownership* at the 1st and 99th percentiles for each month. *ESG ownership* provides a measure of how strongly ESG mutual funds concentrate their holdings in specific companies. For example, the identified ESG mutual funds jointly owned 0.15 percent of Tesla's stocks in September 2020.

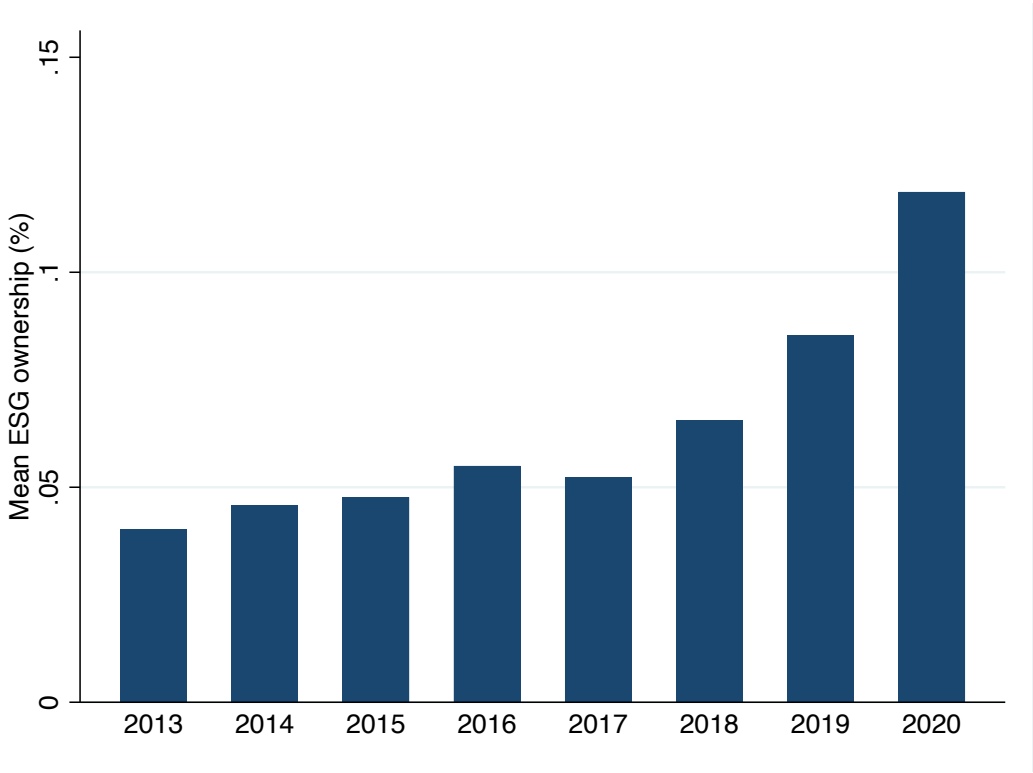


Figure III.2. ESG Ownership Over Time. This figure shows the temporal development of firms' average level of *ESG ownership*, from 2013 to 2017, in percentage points. *ESG ownership* is calculated as the fraction of a company's outstanding shares owned by ESG mutual funds.

The average proportion of shares owned by ESG mutual funds increases in the course of our sampling period. Figure III.2 shows the temporal development of the average level of *ESG ownership* within our sample. While the percentage of companies' shares owned by ESG mutual funds was rather constant between 2013 and 2017, the figure experienced substantial growth between 2018 and 2020. The overall share of stocks

owned by dedicated ESG mutual funds remains, however, limited: in September 2020, at the end of our sampling period, the share of the market capitalization of companies within our sample owned by ESG mutual funds amounts to USD 71 billion, compared to a total market capitalization of these firms of USD 36 trillion.

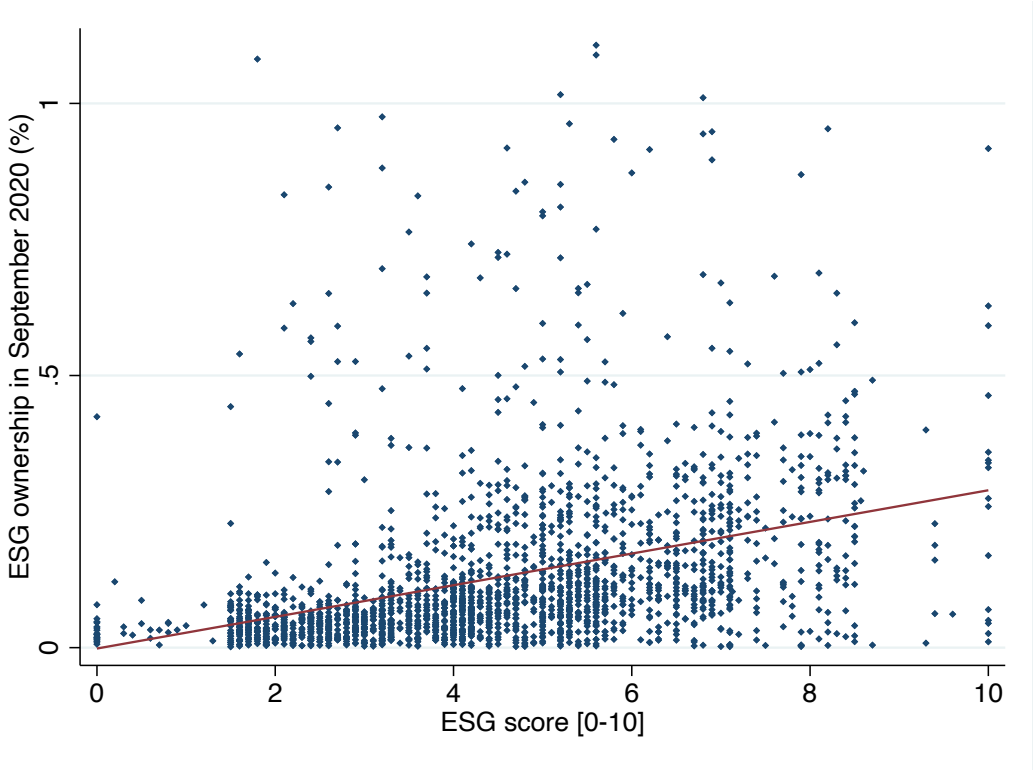


Figure III.3. ESG Ownership and ESG Scores. This figure plots the distribution of firm-level *ESG ownership* against firms' ESG scores in the cross-section, for September 2020. The red line shows a linear correlation between the two variables at this date.

In the cross-section, *ESG ownership* tends to be higher for companies with a high ESG score; there is, however, substantial variation on the level of individual companies. Figure III.3 plots companies' *ESG ownership* in September 2020 against their MSCI ESG score. There is a significant positive correlation between the two variables ($p < 0.001$). However, we cannot make any causal inference from this correlation. It is, for example, possible that large companies, on average, have better ESG ratings and are over-represented in ESG mutual funds. ESG mutual funds may also invest early in companies that later obtain favorable ESG ratings. To investigate whether mutual funds do indeed respond to ESG ratings, we analyze the effect of ESG rating changes in a panel event study design.

III.3.2 Analytical Methodology

We analyze the effect of ESG rating changes on ESG ownership using a panel event study model (Clarke and Tapia-Schythe, 2021; Freyaldenhoven, Hansen, Pérez, and Shapiro, 2021; Schmidheiny and Siegloch, 2019). A panel event study is a specific form of a staggered difference-in-difference design, which allows the estimation of dynamic treatment effects—that is to say, we do not have to make any restrictive assumptions regarding how and when the anticipated effects unfold; instead, we obtain non-parametric estimates of the treatment effect of the event on the outcome in any pre- and post-event period. Treatment effects of events are calculated against the counterfactual of "untreated" firms, which do not experience a rating change at the same time. The identifying assumption is that treated units would have behaved like untreated units had they not been treated. An important plausibility check for this assumption is a parallel trend of the outcome variable prior to the event for both treated and untreated firms. In addition to this, the panel event study specification accommodates firm and time fixed effects.[7]

We jointly estimate the effect of both MSCI ESG rating upgrades and downgrades on *ESG ownership*. Accordingly, we define dummy variables indicating the occurrence of ESG rating upgrades, (u_{it}), and downgrades, (d_{it}), at a specific company i at a specific month t:

$$u_{it} \;=\; 1[t \in \{\upsilon_{i,1}, ..., \upsilon_{i,n}\}] \tag{III.1}$$

$$d_{it} \;=\; 1[t \in \{\delta_{i1}, ..., \delta_{in}\}]. \tag{III.2}$$

Here, δ_{i1} to δ_{in} denote the time periods (months, in this case) in which the company, i, receives an ESG downgrade, and $\upsilon_{i,1}$ to $\upsilon_{i,n}$ the periods in which it receives an upgrade.

To estimate the effect of ESG rating changes on *ESG ownership* we use the following specification:

$$y_{it} \;=\; \sum_{j=\underline{j}-1}^{j=\bar{j}+1} \beta_j b_{it}^j \;+\; \sum_{j=\underline{j}-1}^{j=\bar{j}+1} \gamma_j c_{it}^j \;+\; \mu_i \;+\; \theta_t \;+\; \varepsilon_{it}. \tag{III.3}$$

[7] Given that ESG rating changes are relatively rare events, panel regressions with ESG ratings as a regressor usually do not perform well with firm fixed effects.

Here, y_{it} denotes the level of *ESG ownership* as described above; μ_i and θ_t are firm and month fixed effects, respectively. The unobserved error term is denoted by ε_{it}. Leads and lags for rating changes within treatment windows, ranging from $\underline{j}$ periods prior to the event to $\overline{j}$ periods after the event, are denoted by the variables b_{it}^j (for upgrades) and c_{it}^j (for downgrades). These are binary variables indicating whether a certain firm, i, is j periods away from a rating change in the time period t. As suggested by Schmidheiny and Siegloch (2019), we bin treatment leads and lags that exceed $\underline{j}$ periods before or $\overline{j}$ periods after an event to alleviate potential underidentification issues. We define b_{it}^j and c_{it}^j as

$$b_{it}^j = \begin{cases} \sum_{s=t-\underline{j}+1}^{\overline{t}} u_{is} & \text{if} \quad j = \underline{j} - 1 \\ u_{i,t-j} & \text{if} \quad \underline{j} \leq j \leq \overline{j} \\ \sum_{s=\underline{t}}^{t-\overline{j}-1} u_{is} & \text{if} \quad j = \overline{j} + 1 \end{cases} \tag{III.4}$$

$$c_{it}^j = \begin{cases} \sum_{s=t-\underline{j}+1}^{\overline{t}} d_{is} & \text{if} \quad j = \underline{j} - 1 \\ d_{i,t-j} & \text{if} \quad \underline{j} \leq j \leq \overline{j} \\ \sum_{s=\underline{t}}^{t-\overline{j}-1} d_{is} & \text{if} \quad j = \overline{j} + 1. \end{cases} \tag{III.5}$$

Here, $\underline{t}$ is the first time period in our panel, and $\overline{t}$ the last one. To investigate the effect of ESG rating changes on *ESG ownership* we consider a treatment window ranging from $\underline{j} = 12$ months prior to a rating change to $\overline{j} = 24$ months after a rating change. Following the standard approach, we omit $b_{it}^{j=-1}$ and $c_{it}^{j=-1}$ from the regression, which normalizes the remaining lead and lag coefficients to the level of our dependent variable one time period before the event. Thus, the coefficients of interest, $\beta_{j=0}$ to $\beta_{j=24}$ as well as $\gamma_{j=0}$ to $\gamma_{j=24}$, reflect the effect ESG rating upgrades and downgrades have on *ESG ownership* within the two years following an event, relative to its level one period before the event. Setting the post-event treatment window to two years allows us to capture relatively long-term effects of rating changes while keeping the treatment window in reasonable proportion to our sampling period.

Estimating coefficients for pre-event treatment leads, $\beta_{j=-12}$ to $\beta_{j=-2}$, as well as $\gamma_{j=-12}$ to $\gamma_{j=-2}$, allows us to assess the presence of pre-event trends in our dependent variable. While the absence of a pre-event trend does not prove that this assumption is justified, it makes such an assumption more plausible. We cluster standard errors both at the firm and at the month level to allow for autocorrelation within both dimensions (Petersen, 2009).

III.3.3 Results, Stock Ownership

ESG rating upgrades and downgrades have a long-term effect on *ESG ownership*, as shown in Figure III.4. *ESG ownership* increases significantly after ESG rating upgrades and decreases significantly after ESG rating downgrades. The increase in *ESG ownership* is significant from the first month after an ESG rating upgrade. ESG rating downgrades show a significant effect after seven months. On average, two years after an upgrade ESG ownership is 17.1 percent higher than one month before the upgrade. Vice versa, ESG ownership is 13.1 percent lower two years after a downgrade compared to one month before the downgrade.

The adjustment in ESG ownership we observe following ESG rating changes happens relatively slowly, but it is persistent. Rather than adjusting immediately after a rating change, ESG ownership increases (respectively decreases) more or less gradually within the first two years following a rating change. First, this is evident visually in Figure III.4. Second, the coefficient for the post-event-windows bin for upgrades ($b_{it}^{j=25}$, which sums all upgrade lags beyond 24 months) is significantly positive, and the coefficient for the post-event-windows bin for downgrades ($c_{it}^{j=25}$) is significantly negative.

We do not observe any pre-event trends in *ESG ownership*, which is an important basis for interpreting the relationship between the ESG rating changes and the changes in *ESG ownership* as being a causal one. For both ESG rating upgrades and downgrades, we do not detect any significant differences in the level of *ESG ownership* between the month prior to an ESG rating change and the 11 months before. Also, we do not observe any evident trends in *ESG ownership* in the year before a rating change. While this does not prove the identification assumption that treated units would have followed a similar trend as untreated firms in the period after an event, it makes it unlikely that the observed effect is caused by factors other than ESG rating changes or concurrent confounding events. The lack of differences and trends prior to the event also helps to rule out that the observed effects are driven by non-concurrent changes in time-variant confounds. For example, if an improvement in management quality results in both increased ESG performance and increased *ESG ownership*, this should show up as a pre-event trend, as it likely would take a prolonged time until the improved management results in a rating upgrade. The only candidate for a plausible concurrent confounding event we can conceive of are media scandals regarding ESG-related issues at the same time as the ESG rating change. If firms are exposed in the press for poor ESG practices, this may trigger an ESG rating

downgrade and reduce ESG ownership at the same time. One observation that speaks against this possibility is that the effects of up- and downgrades are nearly symmetric. While negative news exposure with regard to ESG issues may be a reasonable explanation for downgrades, it is unlikely to trigger upgrades in the same way. Also, even if concurrent ESG-related media scandals were partly responsible for the observed effects, it would still mean that ESG mutual funds adjust their holdings based on novel information, which is reflected in the ESG ratings change.

To further explore why ESG mutual funds adjust their holdings in response to ESG rating changes, we run a series of regressions, finding that the nature of ESG rating changes does not significantly affect ESG mutual funds' reaction. Table III.2 shows the regression coefficients for the 12-month post-event lag for ESG rating upgrades and downgrades. Focusing on the 12-month post-event lag allows us to assess how rating changes affect *ESG ownership* within the one year following such changes. Whereas specification (1) corresponds to the initial model as described in III.3.2, specifications (2) to (5) add interactions between the 12-month post-event lag and indicator variables with regards to specific characteristics of the rating changes and their timing.

First, specification (2) explores whether the magnitude of the change in the numeric score that underlies the rating categories plays a role. Ratings can change due to small changes that barely cross the threshold to the next category, but they can also change due to large changes that go far beyond the threshold. The variable *High ESG score change* is a dummy variable indicating whether the change in the underlying ESG score that triggers an ESG rating upgrade is greater or equal to the median of all score changes leading to upgrades. For downgrades, it indicates whether the absolute change in the underlying ESG score is greater or equal to the median of all score changes leading to downgrades. There is no significant correlation between the interaction of the 12-month post-event lags with *High ESG score change* and *ESG ownership*. This suggests that ESG mutual funds are more sensitive to changes in the rating categories than to changes in the underlying scores.

Second, we show, in specification (3), that the ESG rating changes that shift companies into MSCI's ESG "leader" or "laggard" categories do not have a significantly different effect on *ESG ownership* than do other rating changes. MSCI classifies companies with an ESG rating of CCC or B as ESG "laggards" and ones with an ESG rating of AAA or AA as "leaders," arguably adding an even stronger signal on top of the letter ratings. Accordingly, the dummy variable *Leader* indicates rating changes where a company is

upgraded into the "leader" category, and the dummy variable *Laggard* ones where a company is downgraded into the "laggard" category. We do not find a significant correlation between the interaction of the 12-month post-event lag with *Leader* or *Laggard* and *ESG ownership*. This suggests that although a rating change has significant consequences for *ESG ownership*, a transition to the "leaders" or "laggards" category does not trigger a qualitatively different reaction from ESG mutual funds.

Third, we show that ESG rating changes that are associated with a higher level of changes in companies' ESG management practices do not significantly differ in their effect on *ESG ownership* from other rating changes. As discussed in Section III.2, MSCI ratings can change for various reasons. To see whether there is a specific effect when companies adjust their management practices (as opposed to changes in risk exposure, the materiality assessment, or the peer benchmark), we run specification (4). The dummy variable *High ESG practice change* indicates ESG rating upgrades that feature a median-or-above change in the average of our E, S, and G management practice scores as described in detail in Section III.6.1. It also indicates ESG rating downgrades that feature a median-or-above absolute change in this score. Again, there is no significant correlation between the interaction of the 12-month post-event lags with *High ESG practice change* and *ESG ownership*, suggesting that ESG funds do not put particular emphasis on why ESG ratings change.

Finally, we show that the effect that ESG rating changes have on *ESG ownership* does not significantly change over time. The dummy variable *Post 2016* indicates rating changes in the years after 2016, dividing the time period covered by our panel into two periods of similar length. Yet the interaction of the 12-month post-event lags with *Post 2016* does not significantly correlate with *ESG ownership*.

Summing up, we find that changes in companies' ESG ratings have a substantial effect on how many of their shares are owned by mutual funds with an explicit ESG strategy. This indicates that ESG mutual funds do indeed react to changing ESG ratings: They increase their stock holdings following upgrades and decrease them following downgrades. These changes happen relatively slowly, and they accumulate over a period of two years. However, the nature of an ESG upgrade or downgrade does not seem to play an important role when it comes to how strongly ESG mutual funds adjust their portfolios. We do not find significant effects of how large the underlying change in companies' ESG scores is, whether a company enters the ESG "Leader" or "Laggard" category, or whether change

is associated with a high level of change in corporate practices. Neither does the effect increase over time.

III.3.4 Robustness

We corroborate the finding that ESG rating changes affect the holdings of ESG mutual funds in two robustness checks. First we confirm this finding using an alternative measure of ESG mutual funds' ownership of companies. Second, we run a placebo test, which shows that we do not find any significant effects for randomized event dates.

We define *ESG weight* as the fraction that a company's shares represent in the portfolio value of a synthetic ESG mutual fund that aggregates the holdings of all ESG mutual funds in our sample. As shown in Figure III.2, *ESG ownership* increases over time, especially after 2017, reflecting the strong growth of the market for ESG investments. Although we control for time trends by including month fixed effects, this may raise some concerns with respect to our measurement of ESG mutual funds' ownership, especially regarding the upward trend following upgrades. *ESG weight* provides an alternative measurement of ESG mutual funds' ownership that is not directly affected by the growing volume of assets owned by ESG mutual funds. We use the same keyword search as described in Section III.3.2 to identify U.S. domestic equity funds with an explicit ESG strategy based on the CRSP U.S. Mutual Fund Holdings database. Again, we compute the total number of companies' shares held by ESG mutual funds using this database for each month. In addition, we calculate the corresponding share of companies' market capitalization owned by these funds using stock price data from the CRSP Monthly Stock database. Further, we calculate the total market capitalization owned by all identified ESG mutual funds combined, based on the two databases, for each month. We define *ESG weight* as a company's market capitalization that is held by ESG mutual funds in a given month, divided by the total market capitalization held by ESG mutual funds in the same month.

The effect ESG rating changes have on *ESG weight* is very similar to that observed for *ESG ownership*. Figure III.A.1 summarizes the results of a panel event study investigating the effect ESG rating changes have on *ESG weight*, applying the same specifications as for the study of *ESG ownership*. The results confirm our findings that ESG rating upgrades lead to a persistent increase in companies' weight within ESG mutual funds and that downgrades decrease this weight. Whereas the relative effect sizes are of a similar

magnitude as for *ESG ownership*, the confidence intervals of the estimates are somewhat larger.

Next we run a placebo test. We generate random placebo upgrade and downgrade events. To do so, we set the probability of an upgrade taking place for a given company–month such that it generates roughly 2,500 events in total. We repeat the same procedure for downgrades. This gives us a number of randomized upgrades and downgrades close to the number of ESG rating changes that we observe in our sample. We apply the same panel event study method as described in Section III.3.2. The results we obtain in this manner are summarized in Figure III.A.2. We do not find any significant effects of the placebo events on *ESG ownership*. This indicates that the effects we observe are unlikely to be artifacts of our method. Further, it provides reassurance that the observed patterns are not simply a property of the sample.

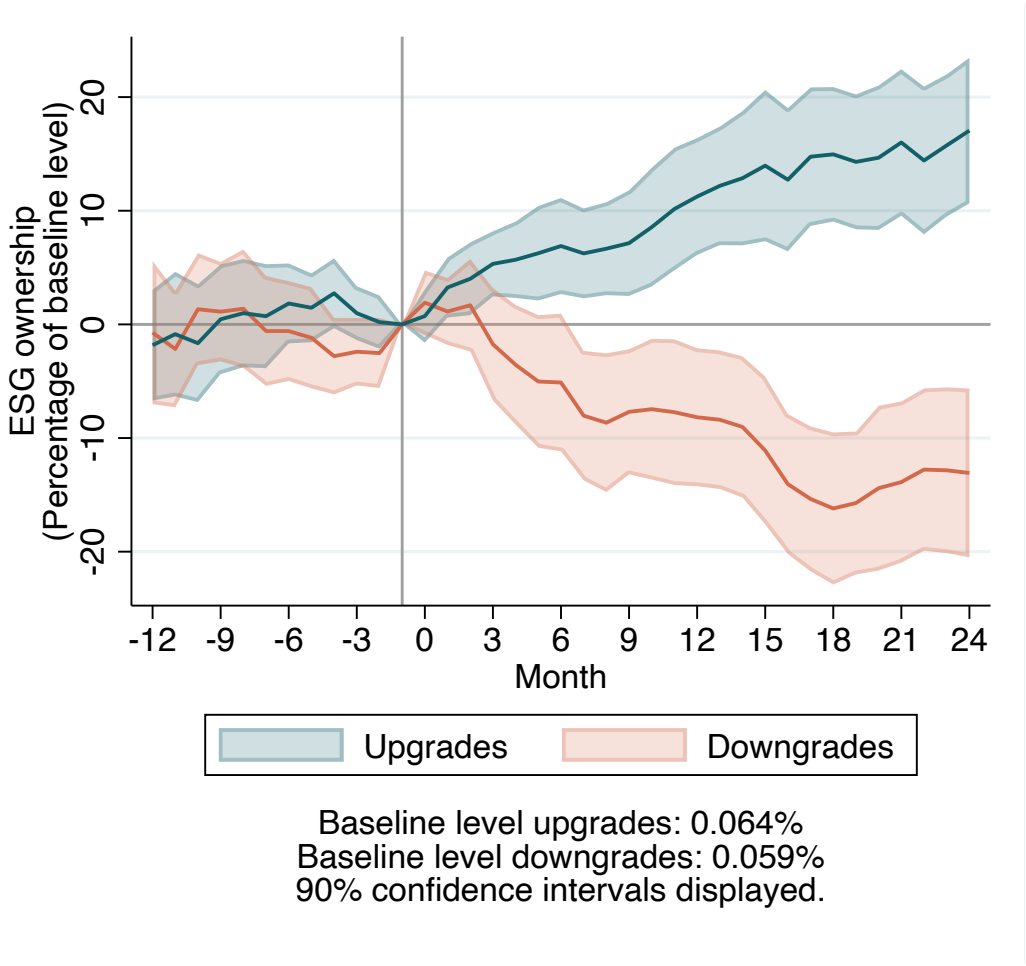


Figure III.4. The Reaction of ESG Ownership to ESG Rating Changes. This figure shows the results of a panel event study with firms' level of *ESG ownership* as the dependent variable and up- and downgrades in the MSCI ESG rating as events. The observation period is February 2013 to September 2020. The figure shows regression coefficients for all treatment leads and lags from 12 months prior to the event to 24 months after. Coefficients are normalized to the baseline level, i.e., the average level of *ESG ownership* one month before a rating change. Coefficients are shown in percentage points of this baseline level. Confidence intervals are based on standard errors clustered at the firm and the month level.

Table III.2
ESG ownership and ESG rating change characteristics

This table shows the results of a panel event study with *ESG ownership* as the dependent variable and MSCI ESG rating up- and downgrades between February 2013 and September 2020 as events. All specifications include treatment leads and lags for the occurrence of ESG rating upgrades and downgrades, i.e., dummy variables indicating that a rating change takes place a specific number of months before (lags) or after (leads) a given observation. Coefficients are normalized to the level of *ESG ownership* one month before a rating change. All specifications include time and firm fixed effects. Specification (1) shows the coefficients for the 12-month lags of ESG downgrades and upgrades, estimating the reaction *ESG ownership* shows to ESG rating changes. Specification (2) interacts these lags with *High ESG score change*, a dummy variable indicating whether the change in the underlying ESG score that triggers an ESG rating change is greater or equal to the median of all score changes leading to changes. Specification (3) includes interactions of the 12-month lags with *Leader* or *Laggard*, dummy variables indicating whether a firm is upgraded into the "Leader" category of MSCI (AAA and AA) or downgraded into the *Laggard* category (CCC and B). Specification (4) includes interactions of the 12-month lags with *High ESG practice change*, which indicates whether an ESG rating change coincides with a median-or-above change in the average of our ESG management practice scores, as described in Section III.6.1. Specification (5) includes interactions of the 12-month lags with the dummy variable *Post 2016*, indicating that a rating change takes place after 2016. t statistics are based on standard errors clustered at the firm and month level shown in parentheses. Stars mark coefficient estimates that are significantly different from zero (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

	(1) *ESG ownership*	(2) *ESG ownership*	(3) *ESG ownership*	(4) *ESG ownership*	(5) *ESG ownership*
Downgrade (12-month lag)	-0.0000483**	-0.0000499**	-0.0000445	-0.0000551**	-0.0000488**
	(-2.21)	(-2.01)	(-1.58)	(-2.34)	(-2.18)
Upgrade (12-month lag)	0.0000715***	0.0000626***	0.0000640***	0.0000790***	0.0000654***
	(3.60)	(2.71)	(3.18)	(3.54)	(3.00)
Downgrade (12-month lag) x *High ESG score change*		0.00000299			
		(0.10)			
Upgrade (12-month lag) x *High ESG score change*		0.0000174			
		(0.56)			
Downgrade (12-month lag) x *Laggard*			-0.00000978		
			(-0.23)		
Upgrade (12-month lag) x *Leader*			0.0000807		
			(1.26)		
Downgrade (12-month lag) x *High ESG practice change*				0.0000145	
				(0.61)	
Upgrade (12-month lag) x *High ESG practice change*				-0.0000177	
				(-0.77)	
Downgrade (12-month lag) x *Post 2016*					0.000000920
					(0.02)
Upgrade (12-month lag) x *Post 2016*					0.0000170
					(0.57)
Firm FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Pre-event leads	Yes	Yes	Yes	Yes	Yes
Post-event lags	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.538	0.538	0.538	0.538	0.538
N	200950	200950	200950	200950	200950

III.4 Do ESG Rating Changes Affect Stock Returns?

Next we explore how ESG rating changes affect the stock returns of affected companies. We run a series of panel event studies with buy-and-hold returns (BHRs) with different holding periods as the dependent variables, showing that companies' stock returns react to ESG rating changes, particularly to downgrades.

III.4.1 Analytical Methodology

For each observation in our panel, we calculate BHRs for different forward-looking holding periods τ. The holding periods range from $\tau = 0$, which covers only the month of the event, to $\tau = 24$, which covers the event month and the following 24 months. We calculate the BHRs based on monthly stock returns obtained from the CRSP Monthly Stock database, which we trim at the 1st and 99th percentiles for every month.

As opposed to *ESG ownership*, the BHR is a cumulative measure, covering several time periods, and the underlying return data reflects changes in value rather than the level of stock prices. Thus, we have to use a different model specification for BHRs. Instead of estimating a single model including pre- and post-event leads and lags, we estimate an individual model for each holding period τ:

$$BHR_{\tau,it} = \sum_{j=-\tau}^{j=\tau} \beta_{j\tau} b_{it}^{j} + \sum_{j=-\tau}^{j=\tau} \gamma_{j\tau} c_{it}^{j} + \mu_{i\tau} + \theta_{t\tau} + X_{it}'\psi_{\tau} + \varepsilon_{it\tau}. \tag{III.6}$$

Here, $\mu_{i\tau}$ and $\theta_{t\tau}$ are firm and month fixed effects, $\varepsilon_{it\tau}$ denotes the unobserved error term, and X_{it} is a vector of firm-level time-varying controls with coefficients ψ_{τ}, as described in detail below. For this specification, we define b_{it}^{j} as $u_{i,t-j}$ and c_{it}^{j} as $d_{i,t-j}$. Again, $u_{i,t-j}$ is a dummy variable indicating the occurrence of a ESG rating upgrade at a specific company i at a specific month $t-j$, while $d_{i,t-j}$ indicates a rating downgrade, as defined in Equations III.1 and III.2.

The variables of interest are $b_{it}^{j=0}$ and $c_{it}^{j=0}$. These variables indicate whether an upgrade or a downgrade takes place for a firm i in month t. We also include pre- and post-event lags, $b_{it}^{j\neq 0}$ and $b_{it}^{j\neq 0}$, to control for the fact that τ-month BHRs for observations less than τ months before or after an event overlap with the holding period of interest. Not

including controls for these observations may bias our estimates downward.[8] Thus, the coefficients $\beta_{j\tau}$ and $\gamma_{j\tau}$ estimate "abnormal" BHRs for a holding period τ that companies experience after a rating upgrade or downgrade, relative to all other τ-month BHRs of observations that are at least τ months away from a rating change, controlling for month and firm fixed effects, as well as for time-variant firm-level characteristics.

Similar to Bolton and Kacperczyk (2021a,b), as well as Briere and Ramelli (2021), we control for the following time-variant firm characteristics, based on quarterly accounting data obtained from the Compustat North America Fundamentals Quarterly database: *leverage* (long-term debt plus debt in current liabilities, divided by total assets, in percentage points), *size* (log(market capitalization)), *book-to-market* (book value of equity divided by market valuation), and *profitability* (income before extraordinary items over total assets). In addition, for each observation we estimate *market beta* from regressions of monthly returns in excess of the 1-month treasury bill rate on the excess market return. For this, we use an estimation window reaching back 48 months prior to the observation. Further, we calculate *momentum* as the firms' average stock return between twelve and two months prior to the observation, following Bessembinder, Cooper, and Zhang (2019). Again, we cluster standard errors both at the firm and at the month levels. We trim these variables at the 1st and 99th percentiles for each month and lag them by one month.

To evaluate the presence of pre-event trends in the BHRs, we run the model as specified in Equation III.6 with all ESG rating upgrade and downgrade event dates moved forward in time by 12 months, for holding periods from $\tau = 0$ to $\tau = 11$. This specification provides us with estimates for $\beta_{j\tau}$ and $\gamma_{j\tau}$ for "placebo" events taking place one year prior to the real events, with holding periods up to one month prior to the real events. This allows us to detect the presence of abnormal BHRs in the year prior to the event.

III.4.2 Results, Stock Returns

We find that ESG rating downgrades affect stock returns. Figure III.5 shows the results of panel event studies investigating the effects of rating changes on buy-and-hold returns (BHRs) with different holding periods. ESG rating downgrades have a pronounced negative effect on BHRs. This effect is significant for holding periods of up to two years following an ESG rating downgrade. We observe the strongest negative abnormal BHRs

[8] Our findings are robust to excluding these leads and lags, as shown in III.A.4; however, the estimated coefficients for $\beta_{j\tau}$ and $\gamma_{j\tau}$ are substantially lower.

19 months after a downgrade, amounting to -3.78%. This corresponds to an annualized abnormal return of -2.37%. Compared to the average annualized return within our sample, of 8.98%, this is economically meaningful. We also find a positive reaction of BHRs to ESG rating upgrades; the effect, however, is weaker and takes longer to materialize than for downgrades. After an (insignificant) initial decrease in BHRs with holding periods of around six months following upgrades, the effect becomes positive for holding periods of over ten months. We detect significantly positive abnormal BHRs for holding periods of 17, 20, and 22 months, reaching a maximum abnormal BHR of 2.62% (1.42% p.a.) on average after 22 months.

We do not identify any evident pre-trends in BHRs before ESG rating changes, which supports a causal interpretation of the stock return reactions we observe. Looking at the BHRs for the placebo events 12 months in advance of the real rating changes, we do not observe any significant abnormal BHRs for holding periods of up to one month before the real events. While this indicates that the observed abnormal BHRs are caused by the ESG rating changes, again we cannot fully rule out the possibility of concurrent confounding events. However, as discussed above, the only candidate apparent to us—namely, concurrent media scandals regarding ESG issues—does not fundamentally change the interpretation of our results.

Additional regression analyses focused on the 12-month horizon are shown in Table III.3, showing that the abnormal stock returns following ESG rating downgrades are significantly higher in the years after 2016, while the nature of rating changes does not significantly affect these abnormal returns. Specification (1) corresponds to the initial model, as described in Equation III.6, where we find a significant negative effect of downgrades and no significant effect of upgrades. We add interactions between upgrade and downgrade indicators ($b_{it}^{j=0}$ and $c_{it}^{j=0}$) and indicator variables regarding different characteristics of the rating changes and their timing. Specification (2) shows that the size of a rating change does not significantly affect the effect rating upgrades and downgrades have on 12-month BHRs (*High ESG score change*). As shown by specification (3), rating changes that take companies into the "leader" or "laggard" categories do not significantly differ in their effect size from the remaining rating changes (*Leader* or *Laggard*). Also, the effect size of rating changes that are associated with higher levels of change in companies' ESG management practices does not significantly differ from ones with a lower level (*High ESG practice change*), as shown by specification (4). Finally, specification (5)

shows that while there is a significant negative effect of downgrades on 12-month BHRs from 2013 to 2016, this effect is significantly stronger from 2017 to 2020.

Summing up, we find that ESG rating downgrades lead to significantly reduced stock returns for a period of up to two years following a rating change and that this reduction is economically meaningful. For ESG rating upgrades, we find a weaker and slower significant positive effect on stock returns. Again, the nature of ESG upgrades or downgrades does not seem to greatly affect how stock markets react to such rating changes. Yet we find that the negative effect of downgrades on stock returns has become larger over time.

III.4.3 Robustness

We corroborate our finding on the effects ESG rating changes have on stock returns by running a placebo test, in which we do not detect any significant effects on stock returns for randomized event dates. Again, we generate random placebo upgrade and downgrade events. To do so, we set the probability of an upgrade taking place for a given company such that it generates roughly 2,500 events in total. We repeat the same procedure for downgrades. We apply the model specification for BHRs as described in Section III.4.1. Figure III.A.3 summarizes the results we obtain in the placebo test. We do not find any significant effects of the placebo events on BHRs, indicating that the observed effects are not an artifact of the analytical method or an inherent property of our sample.

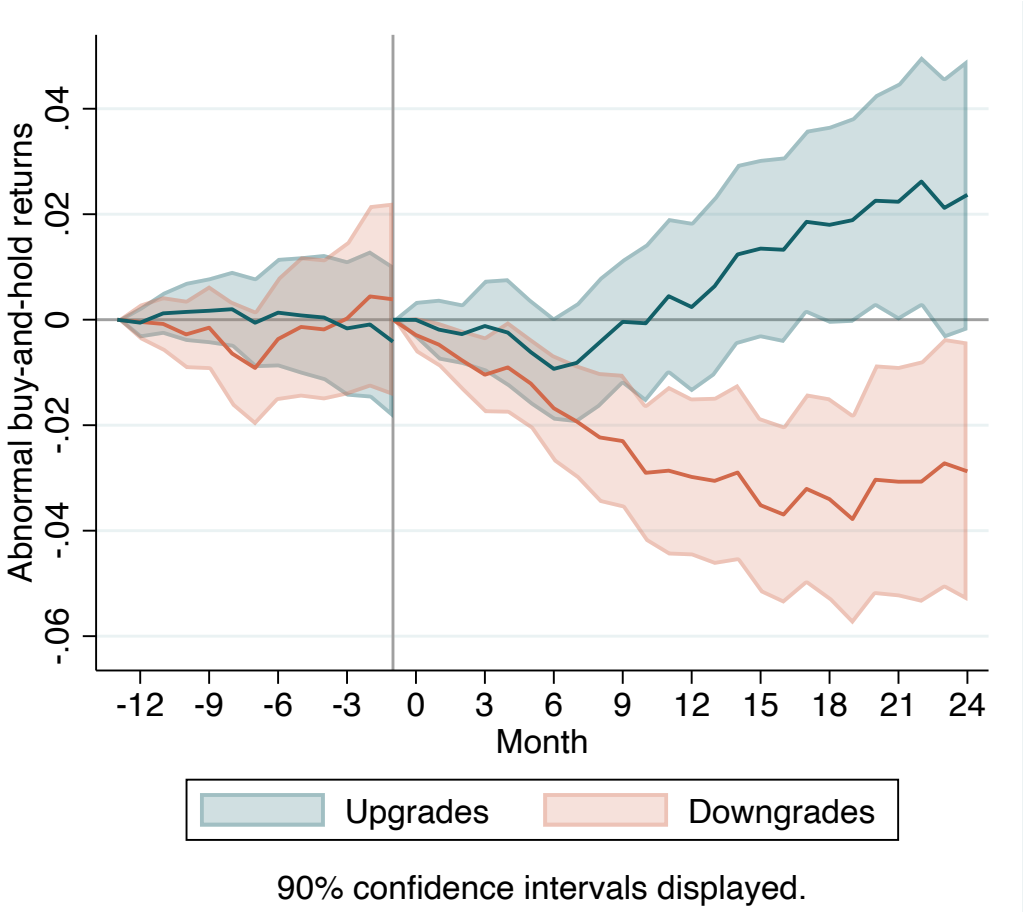


Figure III.5. The Reaction of Stock Returns to ESG Rating Changes. This figure shows the results of a series of panel event studies with BHRs for different holding periods as the dependent variables and up- and downgrades in the MSCI ESG rating between February 2013 and September 2020 as events. For the months greater or equal to zero, the graph displays the abnormal BHRs for holding periods of up to 24 months (corresponding to $\beta_{j\tau}$ and $\gamma_{j\tau}$ in Equation III.6). For the months smaller than zero, the graph displays abnormal BHRs following placebo events 12 months prior to real events, for holding periods of up to one month prior to the real events. All panel event studies include treatment leads and lags covering a time period corresponding to the holding period before and after ESG rating upgrades and downgrades, as well as firm and month fixed effects and lagged time-variant firm-level controls. Confidence intervals are based on standard errors clustered at the firm and at the month level.

Table III.3
Stock returns and ESG rating change characteristics

This table shows the results of a panel event study with 12-month BHRs as the dependent variable and MSCI ESG rating changes between February 2013 and September 2020 as events. Specification (1) shows coefficients for *Upgrade* and *Downgrade*, dummy variables indicating that a corresponding rating change takes place in a given month. Specification (2) includes interactions of these indicators with *High ESG score change*, a dummy variable indicating whether the change in the underlying ESG score that triggers an ESG rating upgrade is greater or equal to the median of all score changes leading to upgrades. Specification (3) includes interactions with *Leader* or *Laggard*, dummy variables indicating whether a firm is upgraded into the "Leader" category of MSCI or downgraded into the "Laggard" category. Specification (4) includes interactions with *High ESG practice change*, which indicates whether an ESG rating change features a median-or-above change in the average of our ESG management practice scores, as described in Section III.6.1. Specification (5) includes with the dummy variable *Post 2016*, indicating rating changes that take place after 2016. All specifications include treatment leads and lags 12 months before and after ESG rating changes, as well as firm and month fixed effects and lagged time-variant firm-level controls. t statistics based on standard errors clustered at the firm and month level are shown in parentheses. Stars mark coefficient estimates that are significantly different from zero (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

	(1)	(2)	(3)	(4)	(5)
	Buy-and-hold return t+12	*Buy-and-hold return t+12*	*Buy-and-hold return t+12*	*Buy-and-hold return t+12*	*Buy-and-hold return t+12*
Downgrade	-0.0237***	-0.0236**	-0.0229***	-0.0247***	-0.0148**
	(-3.25)	(-2.16)	(-2.67)	(-2.65)	(-2.12)
Upgrade	0.00467	0.00514	0.00380	0.0118	-0.000328
	(0.53)	(0.56)	(0.38)	(1.08)	(-0.03)
Downgrade x *High ESG score change*		-0.0000905			
		(-0.01)			
Upgrade x *High ESG score change*		-0.000949			
		(0.07)			
Downgrade x *Laggard*			-0.00199		
			(0.19)		
Upgrade x *Leader*			0.00874		
			(0.38)		
Downgrade x *High ESG practice change*				0.00194	
				(0.82)	
Upgrade x *High ESG practice change*				-0.0167	
				(-1.37)	
Downgrade x *Post 2016*					-0.0480***
					(-2.72)
Upgrade x *Post 2016*					0.0142
					(0.94)
Market beta	-0.0230*	-0.0230*	-0.0230*	-0.0230*	-0.0230*
	(-1.90)	(-1.90)	(-1.91)	(-1.91)	(-1.91)
Leverage	-0.000588	-0.000588	-0.000588	-0.000588	-0.000588
	(-1.31)	(-1.31)	(-1.31)	(-1.31)	(-1.31)
Log(market cap)	-0.245***	-0.245***	-0.245***	-0.245***	-0.245***
	(-10.24)	(-10.24)	(-10.24)	(-10.24)	(-10.24)
Book-to-market	0.0898***	0.0898***	0.0898***	0.0898***	0.0898***
	(2.88)	(2.88)	(2.88)	(2.88)	(2.88)
Profitability	0.000962	0.000962	0.000962	0.000962	0.000962
	(0.65)	(0.65)	(0.65)	(0.65)	(0.65)
Momentum	-0.555***	-0.555***	-0.555***	-0.555***	-0.555***
	(-3.72)	(-3.72)	(-3.72)	(-3.72)	(-3.72)
Firm FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Post-event lags (L1-L12)	Yes	Yes	Yes	Yes	Yes
R-squared	0.446	0.446	0.446	0.446	0.446
N	101799	101799	101799	101799	101799

III.5 Does Corporate Investment React to ESG Rating Changes?

As a next step, we explore the "growth" channel by investigating whether ESG rating changes affect corporate investment decisions. Assuming that ESG rating changes affect returns, it is conceivable that firms experience this as a change in their cost of capital and subsequently adjust their corporate investment activity. To investigate this channel, we run a panel event study with firms' capital expenditure as the dependent variable. In this case, we do not detect any significant effects of ESG rating changes on capital expenditure.

III.5.1 Analytical Methodology

We define the variable *capex* as firms' capital investments divided by the one-quarter lagged book value of their property, plants, and equipment, expressed in percentage points. We obtain the data from the Compustat North America Fundamentals Quarterly database. We trim the obtained observations at the 1st and 99th percentiles for each month.

We run a panel event study investigating the effect of ESG rating changes on *capex*, using the model specification described in Equation III.3. As *capex* is based on accounting data, we have to rely on quarterly observations. Accordingly we set the treatment window to a range from $\underline{j} = 4$ quarters prior to a rating change to $\overline{j} = 8$ quarters after a rating change. In addition, we include *market beta*, *profitability*, *book-to-market*, and *size* as time-variant firm-level controls.

III.5.2 Results, Corporate Investment

Figure III.6 shows the effects of rating changes on firms' real investment. We do not find any significant reactions of firms' capital expenditure to either ESG rating upgrades or downgrades for the two years following rating changes. Looking at the pre-event period, we do not find any significant differences in *capex* in the year prior to ESG rating downgrades. We detect increased levels of *capex* in the two quarters prior to a rating upgrade, compared to the quarter before the event, but not in the third one. While a

positive pre-event trend may mask a potentially negative effect ESG rating upgrades have on *capex*, we cannot discern a clear pre-event trend from the obtained results.

In sum, these results indicate that, while the holdings of ESG mutual funds, as well as stock returns, react to ESG rating changes, companies do not change their level of capital investment.

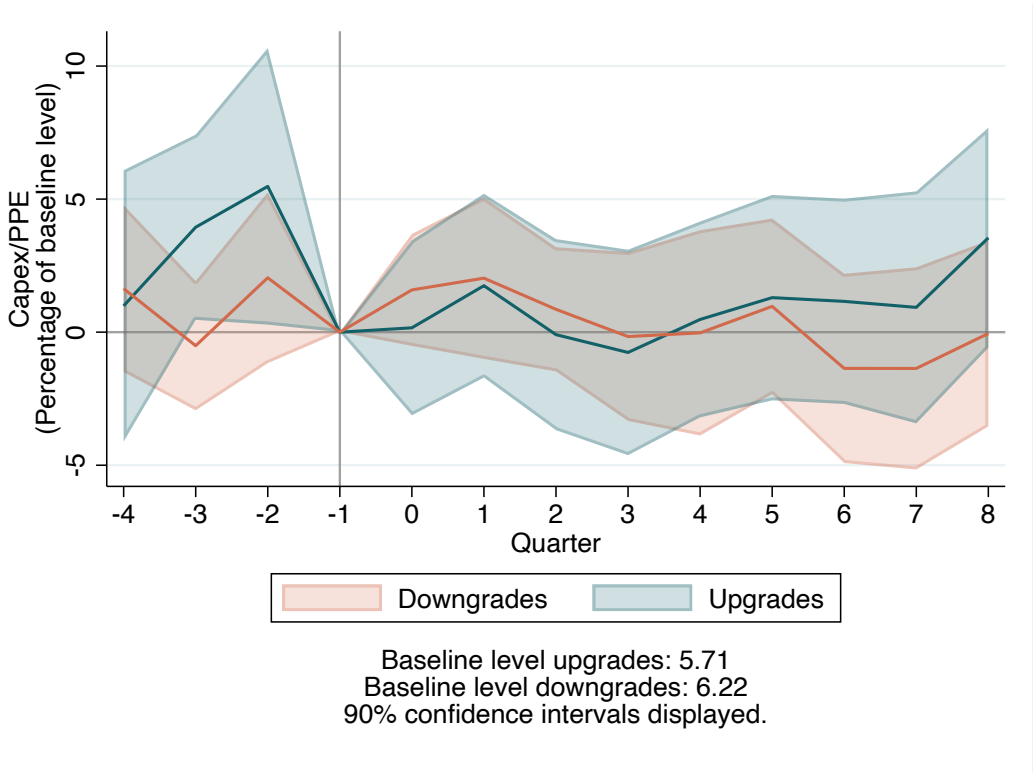


Figure III.6. The Reaction of Corporate Investment to ESG Rating Changes. This figure shows the results of a panel event study with firms' *capex* as the dependent variable and up- and downgrades in the MSCI ESG rating between February 2013 and September 2020 as events. The figure shows the regression coefficients for all treatment leads and lags from 4 quarters prior to the event to 8 quarters after. Coefficients are normalized to the baseline level, i.e., the average level of *capex* one quarter before a rating change. Coefficients are shown in percentage points of this baseline level. Confidence intervals are based on standard errors clustered at the firm and the quarter level.

III.6 Do Firms' ESG Practices React to ESG Rating Changes?

Finally, we investigate the "reform" channel by analyzing how firms' ESG practices react to ESG rating changes. Given that ESG rating changes affect the holdings of ESG mutual funds and stock returns, we can assume that managers have an incentive to optimize their firm's ESG performance. The rationale is that managers can increase their firm's valuation by expending some optimal level of effort on ESG practices, which is a key prediction of the models by Heinkel, Kraus, and Zechner (2001) and Pástor, Stambaugh, and Taylor (2021).

Under these assumptions, ESG rating changes may represent shocks that update a manager's view of the optimal level of ESG effort. After a downgrade, all else being equal, a higher level of effort may be justified to reverse the rating change. After an upgrade, a lower level may be justified given that the next threshold is hard to reach. Several of our findings support the view of ESG rating changes acting as shocks to management: First, as we document for the MSCI ESG rating, investors seem to respond primarily to changes in firms' aggregated ESG ratings and not directly to changes in firms' management practices. Second, even if ESG practices affect firms' ESG ratings, managers can only partially foresee rating changes because such changes also depend on MSCI's assessment of firms' exposure to ESG issues, the financial relevance of these issues, and industry peers' performance. Third, ESG rating changes are relatively rare events; the average company in our sample experiences 0.695 upgrades and 0.582 downgrades.[9]

To investigate how corporate ESG practices react to ESG rating changes, we derive specific measurements for firms' environmental, social, and governance practices and run a panel event study investigating how the results of these measurements develop before and after ESG rating changes.

III.6.1 ESG Management Scores

As discussed in Section III.2, the MSCI ESG rating is based on several components, combining an assessment of risk exposure and management capability. In this section,

[9]The maximum number of rating changes per firm is 7; for upgrades the maximum is 5, and for downgrades the maximum is 4.

we rely on the management capability component to determine whether firms react to rating changes by adjusting their management practices. To measure companies' ESG practices, we rely on management scores provided by MSCI that measure companies' management of 37 specific ESG issues within the social, environmental, and governance dimensions. These scores reflect the quality of a firm's management practices concerning a specific issue on a scale of 0 to 10, where higher scores indicate better practices. We calculate *environmental management score* and *social management score* as the unweighted average of all available management scores within the respective dimension. Within the governance dimension, MSCI does not measure exposure to issues grouped under the "Corporate Governance Theme." All metrics within this theme measure corporate practices relative to peers, in percentiles. We calculate the management score for the governance dimension as the average of the aggregated corporate governance score, which is scaled from 0 to 10, and the management scores for all remaining issues in the governance dimension as *governance management score*. These three measurements react directly to any changes in MSCI's assessment of a firm's ESG management practices but are not influenced by changes in MSCI's assessment of a firm's exposure to certain issues or the financial materiality of specific risks. The governance score is partially influenced by peer effects.

III.6.2 Analytical Methodology

We run three panel event studies investigating the effect ESG rating changes have on *environmental management score*, *social management score*, and *governance management score*, using the model specification described in Equation III.3. We set the treatment window to range from $\underline{j} = 12$ months prior to a rating change to $\overline{j} = 24$ months after a rating change.

The underlying measurements we use to derive the ESG management scores are also reflected in the MSCI ESG rating. Thus, the ESG rating changes we observe may include an adjusted assessment of firms' ESG management practices. This can lead to a jump in ESG management scores from one month before a rating change ($j = -1$) to the month of the rating change ($j = 0$). To explore whether companies adjust their ESG practices after a rating change, we omit $b_{it}^{j=0}$ and $c_{it}^{j=0}$ from the regression. This normalizes coefficients to the level of ESG management scores immediately after a rating change, enabling us to determine whether the management scores changed significantly in the following months.

To investigate pre-trends in ESG management scores, we run a separate regression, using the same specifications but omitting $b_{it}^{j=-1}$ and $c_{it}^{j=-1}$. This normalizes coefficients to the level of management scores immediately before a rating change, which enables us to determine whether there are any significant changes in ESG management scores in the year prior to an ESG rating change.

As the measurements we use to derive the ESG management scores are also considered in the ESG ratings, the relationship between ESG management scores and ESG rating changes is at least partly endogenous. A causal interpretation of observed changes in the practice scores may therefore not be warranted. Nevertheless, we can observe how companies' ESG management practices develop before ESG rating changes and whether companies adjust their ESG management practices following rating changes. Also, our analysis relies on the assumption that ESG management scores react to actual changes in firms' practices in a timely manner. While MSCI's policy is to update the underlying metrics on an ongoing basis, we cannot exclude that some adjustments of corporate practices are captured with a delay.[10] Given that we observe the development of the management scores over two years, however, it seems unlikely that such delays substantially affect our findings.

III.6.3 Results, ESG Practices

We do not find evidence that companies substantially adjust their environmental management practices after ESG rating changes. Figure III.7 shows the development of environmental management practices before and after ESG rating changes. Coefficients for the 24 months following a rating change ($j > 0$) are normalized to the level of *environmental management scores* immediately after a rating change and thus indicate how this score reacts to rating changes. We find a slightly but significantly increased average level of *environmental management scores* 13 months after an upgrade, but no significant reaction for any of the other 24 months. For ESG rating downgrades, we do not find any significant change in companies' *environmental management score* following a rating change. Coefficients for the 11 months before a rating change ($j < -1$) are normalized to the level of *environmental management scores* immediately before a rating change. Here

[10] MSCI's methodology summary states that "Companies are monitored on a systematic and ongoing basis, including daily monitoring of controversies and governance events. New information is reflected in reports on a weekly basis and significant changes to scores trigger analyst review and re-rating. Companies typically receive an in-depth review on an annual basis."

we observe a slightly negative trend in *environmental management scores* before downgrades, with levels being significantly higher one year before a rating change compared to the month before a downgrade. This indicates that the *environmental management scores* of firms that are close to receiving an ESG rating downgrade deteriorate in the year before the change. While we do not find a significant reduction after rating changes, we can exclude that there is a trend reversal—that is to say, that firms significantly improve their *environmental management scores* following a downgrade.

Also, we do not find clear evidence showing that firms substantially alter their social management practices after ESG rating changes. Figure III.8 shows how social management practices change before and after ESG rating changes. We do not find any significant changes in the *social management score* in the two years following ESG rating upgrades. We find a clear negative pre-trend in *social management scores* before upgrades, potentially indicating that the development of firms' social management practices stabilizes after ESG rating upgrades. Further, we observe that *social management scores* fall significantly starting six months after ESG rating downgrades. However, we also observe a clear negative pre-trend in *social management scores* before downgrades; thus, we cannot infer that the companies actively adjust their social management practices following downgrades. We can, in any case, exclude that downgrades are followed by any significant improvements in *social management scores*.

Concerning the governance dimension, our results indicate that companies adjust their governance efforts after ESG rating changes. Figure III.9 shows the development of firms' *governance management score* around ESG rating changes. We find that the *governance management score* significantly deteriorates after ESG rating upgrades, while *governance management scores* in the year prior to an upgrade do not significantly differ from the level immediately before an upgrade. In contrast, companies' *governance management score* significantly increases in the two years after ESG rating downgrades. In the year before ESG rating downgrades, we observe a clear downward-facing trend in the *governance management score*. Thus, we observe a clear trend reversal in firms' governance practices after downgrades, supporting a causal interpretation of the observed improvement.

In sum, we find that companies barely adjust their environmental and social practices following ESG rating changes. However, companies seem to adjust their governance after ESG rating changes, improving it after downgrades and letting it deteriorate to some extent after upgrades.

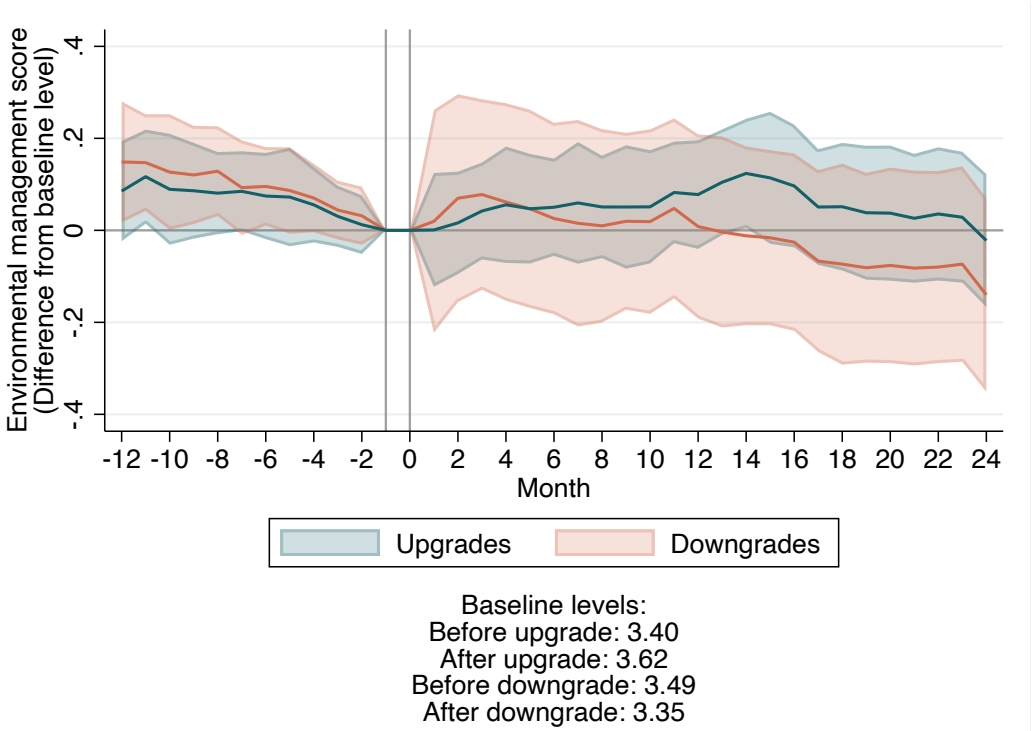


Figure III.7. The Reaction of Environmental Management Practices to ESG Rating Changes. This figure shows the results of a panel event study with firms' *environmental management score* as the dependent variable and up- and downgrades in the MSCI ESG rating between February 2013 and September 2020 as events. The figure shows the regression coefficients for treatment leads and lags from 12 months prior to the event to 24 months after. Coefficients for months before an ESG rating change are normalized to the level of *environmental management scores* immediately before the change (month -1). Coefficients for months after an ESG rating change are normalized to the level of *environmental management scores* immediately after the change (month 0). We display 90% confidence intervals, based on standard errors clustered at the firm and the month level.

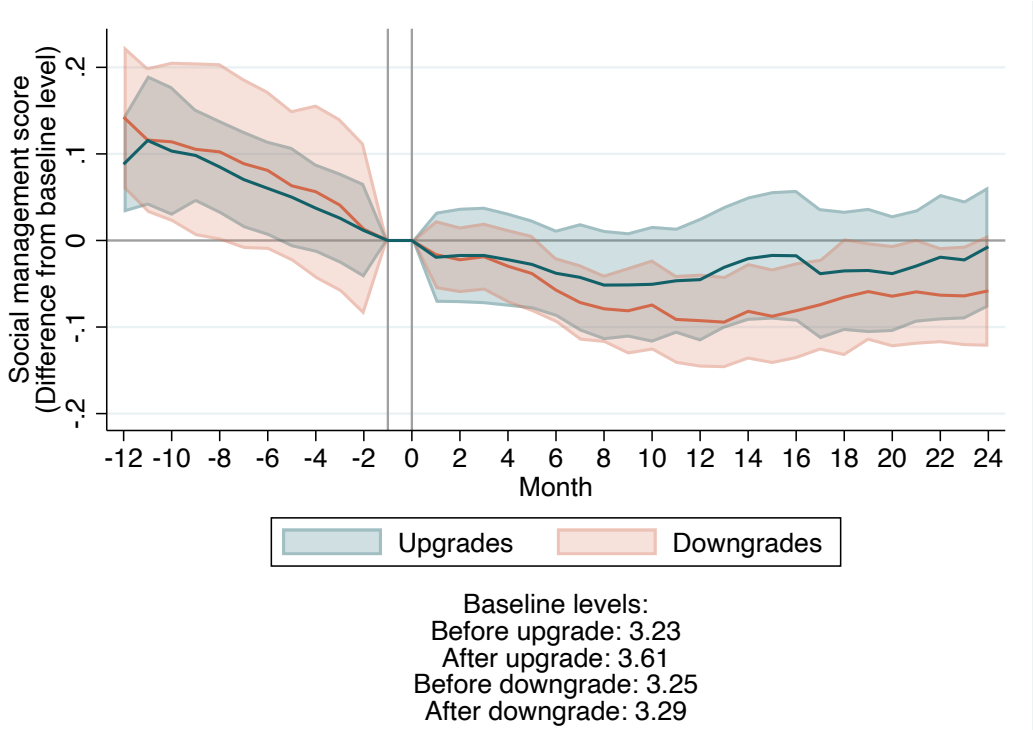


Figure III.8. The Reaction of Social Management Practices to ESG Rating Changes. This figure shows the results of a panel event study with firms' *social management score* as the dependent variable and up- and downgrades in the MSCI ESG rating between February 2013 and September 2020 as events. The figure shows the regression coefficients for treatment leads and lags from 12 months prior to the event to 24 months after. Coefficients for months before an ESG rating change are normalized to the level of *social management scores* immediately before the change (month -1). Coefficients for months after an ESG rating change are normalized to the level of *social management scores* immediately after the change (month 0). We display 90% confidence intervals, based on standard errors clustered at the firm and the month level.

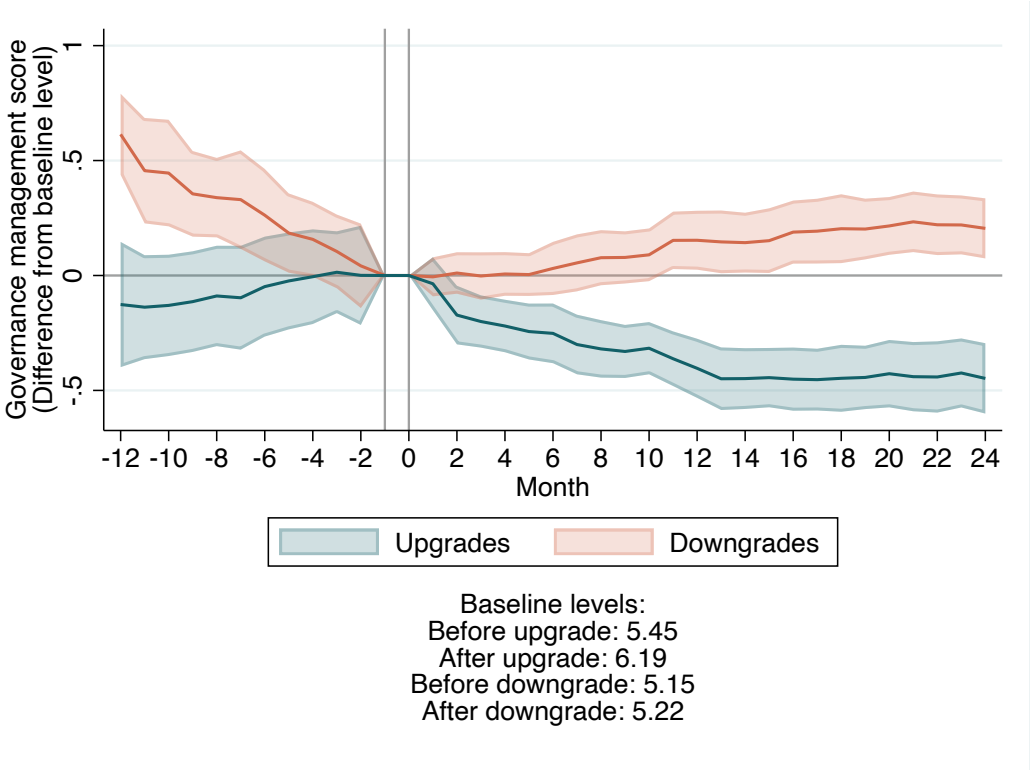


Figure III.9. The Reaction of Governance Practices to ESG Rating Changes. This figure shows the results of a panel event study with firms' *governance management score* as the dependent variable and up- and downgrades in the MSCI ESG rating between February 2013 and September 2020 as events. The figure shows the regression coefficients for treatment leads and lags from 12 months prior to the event to 24 months after. Coefficients for months before an ESG rating change are normalized to the level of *governance management scores* immediately before the change (month -1). Coefficients for months after an ESG rating change are normalized to the level of *governance management scores* immediately after the change (month 0). We display 90% confidence intervals, based on standard errors clustered at the firm and the month level.

III.7 Discussion

Our results shed light on the question of whether ESG ratings are economically relevant. This is an important question as regulators discuss whether and how to regulate ESG ratings. It also provides important empirical insights into the question of whether and how ESG investing has market relevance, which has so far been mostly studied at the theoretical level.

Our results provide some empirical support to theoretical models that predict an impact of ESG performance on stock prices. In general, such models assume that a fraction of the market derives nonpecuniary utility from holding "green" companies (Heinkel, Kraus, and Zechner, 2001; Landier and Lovo, 2020; Oehmke and Opp, 2019; Pedersen, Fitzgibbons, and Pomorski, 2021; Pástor, Stambaugh, and Taylor, 2021). However, in their study of PRI signatories, Gibson, Glossner, Krueger, Matos, and Steffen (2021) find that U.S. signatories do not hold portfolios that are significantly different from those of non-signatories with regard to these portfolios' ESG performance. This questions whether there actually are investors with genuine ESG preferences, and suggests that investors' ESG commitments may be primarily "cheap talk." In contrast, our results on the holdings of dedicated ESG mutual funds provide evidence that there is a part of the market that expresses ESG preferences and consequently adjusts holdings to changes in ESG ratings.

We acknowledge that we cannot discern whether ESG mutual funds adjust their holdings due to a belief that ESG ratings are a signal for future cash flows or due to a taste for holding companies with good ESG ratings. Both interpretations are consistent with the observation that ESG rating changes trigger an adjustment of holdings, and it is likely that both drivers are present in reality. Also, the fact that information on adjusted ESG ratings is integrated relatively slowly over the course of two years suggests that the "taste" channel is at least part of the explanation. The fact that we do not detect any influence of changes in underlying ESG scores or management practices adds to this notion. If ESG ratings contained novel cash flow–relevant information, we would expect more rapid holdings adjustments, and a stronger reaction to greater underlying changes in firms' ESG characteristics. A slow, unspecific adjustment is consistent with ESG mutual funds periodically (e.g., quarterly) adjusting holdings to maintain a certain portfolio-wide ESG performance. With such a process in place, funds' positions in downgraded companies would be slowly reduced, and their positions in upgraded companies would be increased.

An important caveat to our support for theoretical models of the effect of ESG investing is the market relevance of ESG mutual funds. Pástor, Stambaugh, and Taylor (2021) show that pricing effects of ESG investing should increase in two key parameters: first, the strength of investors' preference for ESG performance, and second, the share of ESG investors in the market. While we present evidence for the influence of ESG preferences in investment decisions of dedicated ESG mutual funds, these funds owned less than 0.2 percent of the total assets in our sample in September 2020. Thus, the overall numbers of investors with a genuine taste for ESG performance could be small. Along these lines, Ceccarelli, Glossner, Homanen, and Schmidt (2021) show that while PRI signatories manage more than USD 100 trillion globally, there is only a small subset of PRI signatories who contribute time and effort to engagement activities.

Nonetheless, we document an impact of ESG rating changes on stock returns. This is in line with theoretical models, in the sense that investors with a taste for ESG performance will bid up stock prices once a rating improves (and drive prices down when a rating deteriorates). Interestingly, the price response to ESG rating changes is slow, suggesting that there could be opportunities for arbitrage. While the temporal dynamic of the stock price reactions to ESG rating changes aligns well with the effect such changes have on the holdings of ESG mutual funds, we cannot determine whether the stock price effect is driven by the observed holdings changes. The observed price impact may also be caused by additional segments of the market that respond, perhaps less strongly, to ESG rating changes. In sum, our findings lend support to model predictions that firms' ESG performance has stock price implications.

Our results, however, provide only very limited evidence for the theoretically predicted impact of ESG investing on the real economy, which is perhaps the most important and interesting aspect of ESG investing. We find no evidence for the growth channel, where "green" firms grow faster due to a reduced cost of capital and "gray" firms grow more slowly. We do not observe any short- or long-term changes in firms' capital expenditure following ESG rating changes. Potentially the advantage in the cost of capital is too small to be economically relevant, as argued by Berk and van Binsbergen (2021).

Regarding the reform channel, where firms become "greener" to increase their valuation, our findings suggest that ESG investing encourages the picking of low-hanging fruit. We cannot detect any improvement in environmental and social management practices following downgrades. However, we do find that firms improve governance practices following downgrades. Our interpretation of this finding is that managers are aware of

valuation losses caused by ESG rating downgrades and thus try to improve their rating after such events. Improvements in the governance domain may offer the cheapest way of improving firms' ESG ratings. Also, such improvements may easily find the support of institutional investors, even if these do not have a particular preference for ESG performance. In contrast, improvements in the environmental and social domains may be more costly to achieve and more controversial. Also, firms may have fewer incentives to improve environmental and social practices, as, due to MSCI's risk focus, scores in these two dimensions may often depend on risk exposure rather than on management practices. We also find that governance practices deteriorate after rating upgrades, suggesting that management's attention to governance issues may decline once a targeted ESG rating upgrade is achieved.

III.8 Conclusion

This paper investigates the economic impact of ESG rating changes. We find that ESG rating upgrades lead to an increase in firms' ownership by mutual funds with an explicit ESG strategy and increased buy-and-hold returns over a window of up to two years. We find the opposite effect for downgrades. Regarding impact on the real economy, we find no evidence indicating that firms' growth is affected by ESG rating changes. We do find that firms improve their governance practices in response to downgrades, but not their social or environmental practices. Taken together, our results suggest that ESG rating changes matter for a relatively small subset of mutual funds with a clear ESG mandate and have pricing effects but little real economic impact. This picture could change in the future if more investors become more serious about ESG integration.

III.A Appendix

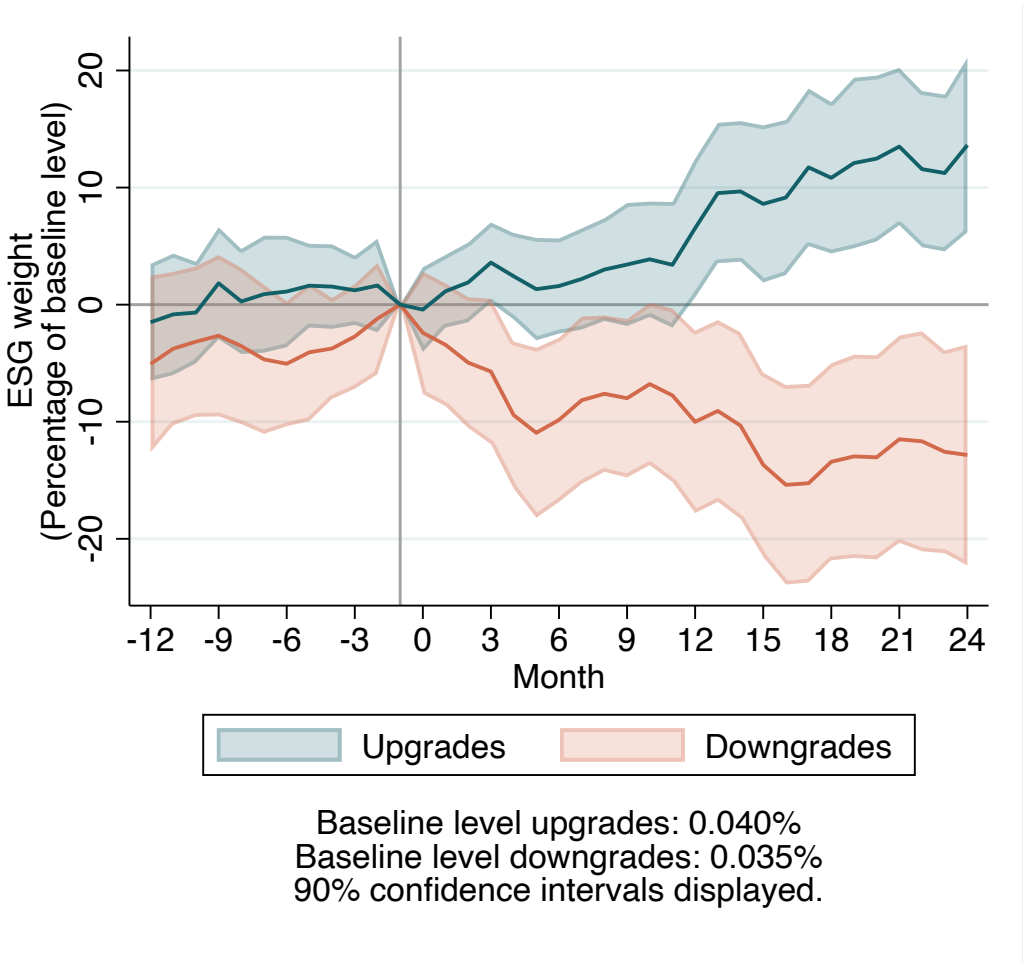


Figure III.A.1. The Reaction of *ESG weight* to ESG Rating Changes. This figure shows the results of a panel event study with firms' level of *ESG weight* as the dependent variable and up- and downgrades in the MSCI ESG rating as events. The observation period is February 2013 to September 2020. The figure shows regression coefficients for all treatment leads and lags from 12 months prior to the event to 24 months after. Coefficients are normalized to the baseline level, i.e., the average level of *ESG weight* one month before a rating change. Coefficients are shown in percentage points of this baseline level. Confidence intervals are based on standard errors clustered at the firm and the month level.

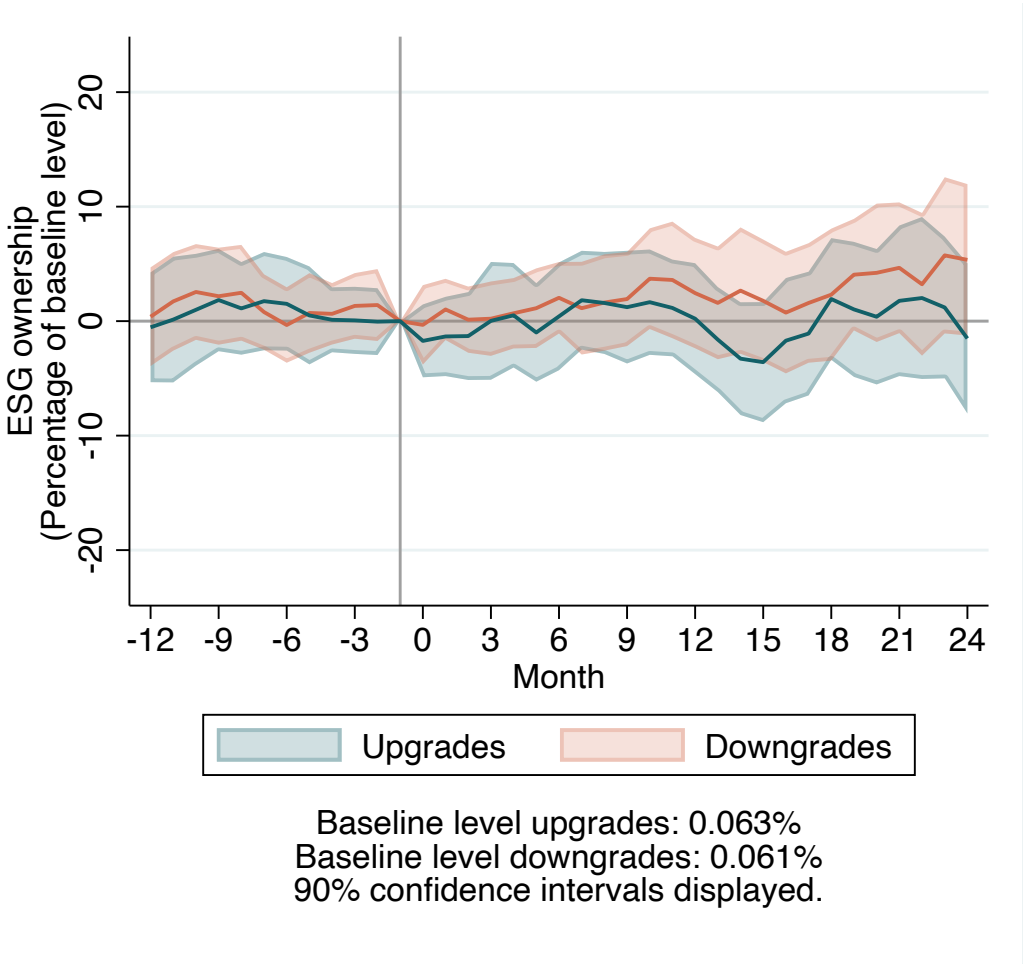


Figure III.A.2. Placebo Panel Event Study, *ESG ownership*. This figure shows the results of a panel event study with firms' level of *ESG ownership* as the dependent variable and roughly 2,500 simulated placebo upgrade and 2,500 placebo downgrade events. The observation period is February 2013 to September 2020. The figure shows regression coefficients for all treatment leads and lags from 12 months prior to the placebo events to 24 months after. Coefficients are normalized to the baseline level, i.e., the average level of *ESG ownership* one month before a placebo event. Coefficients are shown in percentage points of this baseline level. Confidence intervals are based on standard errors clustered at the firm and the month level.

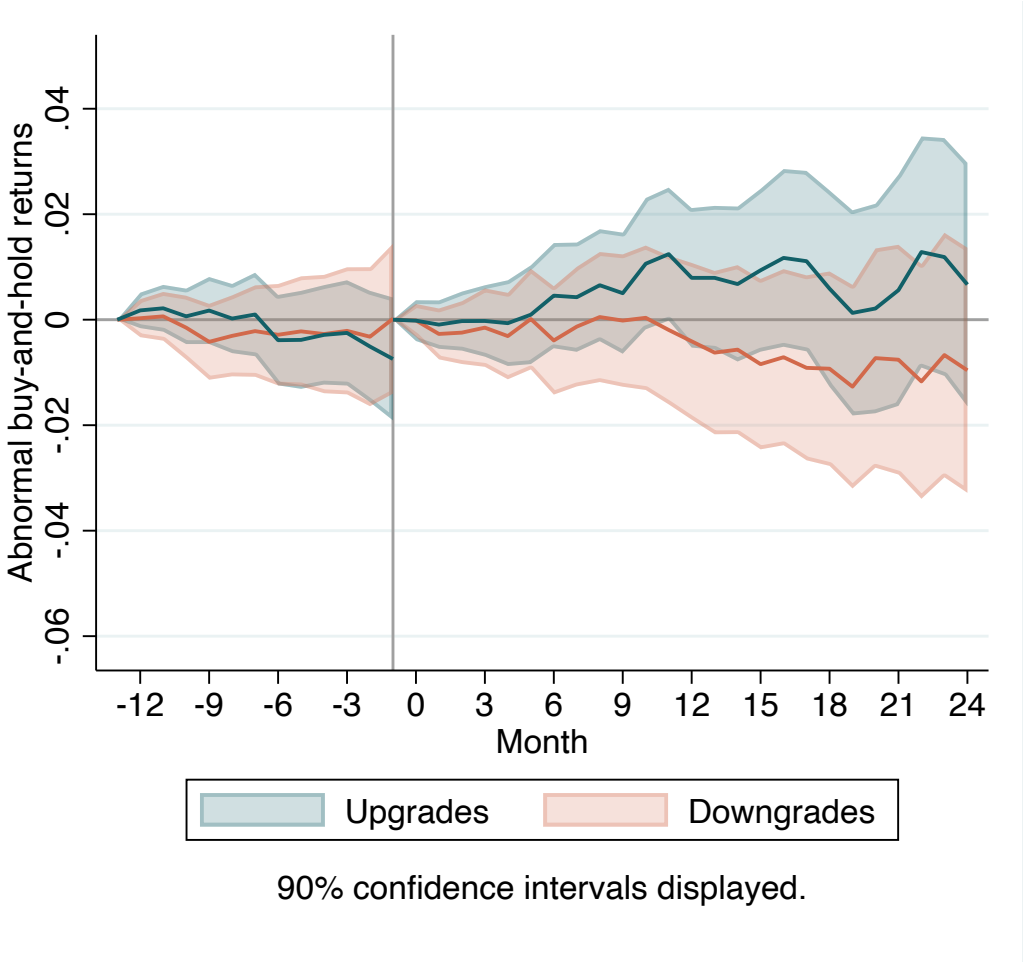


Figure III.A.3. Placebo Panel Event Study, BHRs. This figure shows the results of a series of panel event studies with BHRs for different holding periods as the dependent variables and roughly 2,500 simulated placebo upgrades and 2,500 placebo downgrades between February 2013 and September 2020 as events. For the months greater or equal to zero, the graph displays the abnormal BHRs for holding periods of up to 24 months (corresponding to $\beta_{j\tau}$ and $\gamma_{j\tau}$ in Equation III.6). For the months smaller than zero, the graph displays abnormal BHRs following placebo events 12 months prior to the simulated events, for holding periods reaching up to one month prior to the real events. All panel event studies include treatment leads and lags covering a time period corresponding to the holding period before and after ESG rating upgrades and downgrades, as well as firm and month fixed effects and lagged time-variant firm-level controls. Confidence intervals are based on standard errors clustered at the firm and the month level.

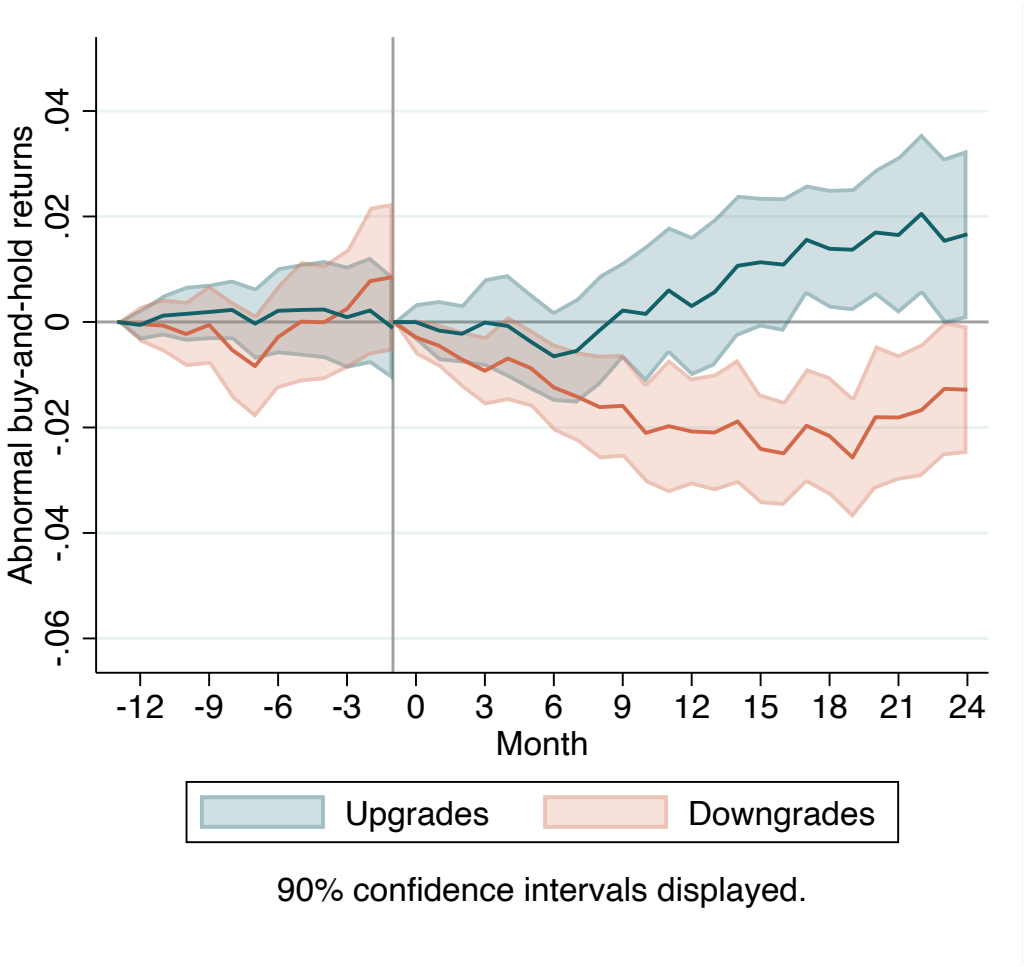


Figure III.A.4. Panel Event Study, BHRs Without Pre- and Post-event Leads and Lags. This figure shows the results of a series of panel event studies with BHRs for different holding periods as the dependent variables and up- and downgrades in the MSCI ESG rating between February 2013 and September 2020 as events. For the months greater or equal to zero, the graph displays the abnormal BHRs for holding periods of up to 24 months (corresponding to $\beta_{j\tau}$ and $\gamma_{j\tau}$ in Equation III.6). For the months smaller than zero, the graph displays abnormal BHRs following placebo events 12 months prior to real events, for holding periods of up to one month prior to the real events. The panel event studies include firm and month fixed effects and lagged time-variant firm-level controls, but they do not include treatment leads and lags. Confidence intervals are based on standard errors clustered at the firm and the month level.

IV How Do Investors Consider Impact in Their Decision-Making?

Do Investors Care About Impact?*

Florian Heeb, Julian F. Kölbel, Falko Paetzold, and Stefan Zeisberger

A version of this paper is forthcoming in *The Review of Financial Studies*.

Abstract

We assess how investors' willingness to pay (WTP) for sustainable investments responds to the impact of those investments, using a framed field experiment. While investors have a substantial WTP for sustainable investments, they do not pay significantly more for more impact. This also holds for dedicated impact investors. When investors compare several sustainable investments, their WTP responds to relative but not to absolute levels of impact. Regardless of investments' impact, investors experience positive emotions when choosing sustainable investments. Our findings suggest that WTP for sustainable investments is primarily driven by an emotional rather than a calculative valuation of impact.

*We thank Tarun Ramadorai and two anonymous reviewers for their thoughtful comments. Further, we thank Erol Bilecen, Timo Busch, Marc Chesney, Alain Cohn, Kornelia Fabisik, Jonathan Harris, Sam Hartzmark, Thorsten Hens, Coralie Jaunin, Pascal Kieren, Christoph Merkle, Jan Schmitz, Julien Senn, Zacharias Sautner, Paul Smeets, Hersh Shefrin, Janneke Toussaint, Michael Ungeheuer, Alexander Wagner, Utz Weitzel, Ziwei Zhao, and participants of seminars and conferences at the AFA Annual Meeting, ARCS, the Dutch Ministry of Finance, the GRASFI, the JEF Seminar, the Liechtenstein Sustainable Finance Workshop, the MPI Bonn, Virtual Experimental Finance Workshop, the SEF Annual Conference, the DGF, Bonn Household Finance Workshop, Radboud University, the SFI Research Days, the SGF, the University of Exeter, and the University of Zurich for helpful comments and suggestions. We gratefully acknowledge financial and participant pool support from the VEB. Florian Heeb was supported by a PRI FIR research award grant. We further thank the VEB and Armand Kersten, Toniic and Kristin Siegel, NEXUS and Tharald Nustad, Credit Suisse and James Gifford, Bank Vontobel and Daniel Grüneisen, Bank Julius Baer and Alessandro Anastassio, Tiedemann Advisors and Brad Harrison, and VALUEworks and Peter Wüthrich for providing support regarding data collection.

IV.1 Introduction

Sustainable investing is seen as a mechanism for curbing negative externalities (Benabou and Tirole, 2010; Broccardo, Hart, and Zingales, 2020; Hong, Wang, and Yang, 2021; Landier and Lovo, 2020; Oehmke and Opp, 2019; Pástor, Stambaugh, and Taylor, 2021). This mechanism is based on the assumption that a substantial pool of investors hold pro-social preferences. Indeed, recent research has demonstrated that pro-social preferences affect investment decisions (Riedl and Smeets, 2017), and that these preferences result in sizable, market-wide fund flows toward sustainable investments (Hartzmark and Sussman, 2019). Some investors show an explicit willingness to pay (WTP) for investments with "impact" (Barber, Morse, and Yasuda, 2021), which we define as a positive externality of an investment.

Standard decision theory would predict that investors are consequentialists—so, that the utility that investors derive from a sustainable investment is proportional to the impact of that investment.[1] This consequentialist view is also adopted widely in current models of sustainable investing. Yet research on charitable giving and public good valuation shows that individuals often display scope insensitivity—that is, they are insensitive to the magnitude of their impact.[2] Scope insensitivity may be prominent in sustainable investing decisions. If it is, this could undermine the effectiveness of sustainable finance as a whole, as the financial industry may not have an incentive to supply products with substantial impact. To shed light on this issue, we ask the question: How does investors' WTP for sustainable investments respond to the impact of these investments?

We investigate this research question in a pre-registered, framed field experiment[3] with experienced investors. They choose between a sustainable investment with a quantified impact and a financially equivalent investment with zero impact. The investment choices are incentivized and consequential—that is, investors can make real money, and their choices have real impact. Based on the investors' choices, we elicit their WTP for the sustainable investment in terms of additional fees they are willing to pay. We operationalize impact in the form of carbon dioxide (CO_2) emissions reductions and vary the

[1] Traditional models of altruistic behavior often assume that individuals contribute to public goods because they derive utility from the well-being of others (Andreoni and Miller, 2002; Becker, 1974; Eckel and Grossman, 1996).

[2] See, for example, Desvousges, Johnson, Dunford, Boyle, Hudson, and Wilson (1992); Kahneman and Knetsch (1992); Karlan and Wood (2017); Metzger and Günther (2019).

[3] According to the classification of Harrison and List (2004).

impact of the sustainable investment by a factor of 10 between two treatment groups. We run the experiment in different variations with a panel of 527 experienced private investors as well as with a unique panel of 125 dedicated high-net-worth impact investors whom we recruited via impact investor networks and specialized wealth managers.

Our main experiment provides evidence that investors' WTP does not respond significantly to the level of impact that a sustainable investment offers. As a starting point, we confirm that investors are willing to pay a substantial amount for a sustainable investment with some impact. However, investors' WTP does not significantly differ between an investment that saves 0.5 tons of CO_2 (tCO_2) emissions and one that saves 5 tons. The experiment does not rule out that some investors are sensitive to impact, but in aggregate that seems to be a second-order effect. I sum, we find that although investors care whether an investment has an impact or not, they barely care about the magnitude of that impact.

We take several measures to ensure the robustness of this finding. First, we make sure that investors intuitively understand what a ton of CO_2 means and that the investment's impact is salient when investors make their choices. Most investors (95%) can exactly recall the impact of the investment after the choice experiment. Second, we run a series of variations of our main experiment. Using the same subject pool, we vary the investments' past financial performance instead of their impact, assuming that investors care about the level of financial performance. We can demonstrate that in this setting investors' WTP is highly sensitive to differences in financial performance. Using an additional sample of 2,800 participants recruited via Amazon Mechanical Turk (MTurk), we first replicate our findings with a larger number of participants (n=1,000). Subsequently, we rely on MTurk participants to run eight variations of the main experiment to probe for the potential influence of experimenter demand and our elicitation method. The result of no significant difference between impact treatments holds across all MTurk experiments. Finally, relying on a sample of 554 university students, we show that our results hold both before and after the COVID-19 crisis.

Extending the main finding, we explore the reasons for the observed insensitivity to impact. First, our results may be due to the fact that investors lack experience in dealing with impact. To test this hypothesis, we repeat the experiment, but this time with a unique sample of high-net-worth impact investors who have substantial experience with impact investing. We find that investors in this sample are just as insensitive to impact as are the private investors in the main experiment. This suggests that the observed

insensitivity to impact is not driven by a lack of experience, and that it is unlikely to disappear as more investors gain experience with sustainable investing.

Second, we investigate whether the ability to directly compare impact information increases investors' sensitivity to impact. To this end, we run our experiment in a joint evaluation setup, where investors receive information on the impact of each of two sustainable investments juxtaposed. Again, the impact of the investments differs by a factor of 10. Yet investors' WTP is only 28 percent higher for the high-impact investment in this setting, even though it has a 900 percent higher impact. Thus, while investors show some reaction to impact information in direct comparison, their WTP per unit of impact remains inconsistent. Analyzing sub-groups, we find that, even in the joint evaluation design, one-third of investors are entirely insensitive to the level of impact. In contrast, the 30 percent of investors with the highest sensitivity to impact pay on average six times more for ten times more impact. Furthermore, we show that in the joint evaluation setup investors' WTP does not depend on the absolute level of impact. Reducing the impact of each investment by a factor of 10 barely changes investors' WTP. This indicates that investors' WTP for sustainable investments depends strongly on the available (arbitrary) choice set and not on the absolute level of impact.

Third, we explore whether WTP for sustainable investments is driven by the emotional experience of choosing a sustainable option rather than by a calculative appraisal of impact. Relying on a post-experiment survey, we find that investors' WTP is correlated with the level of positive emotions they experience when choosing the sustainable investment. The impact of the sustainable investment, however, does not influence these positive emotions. In a regression, we show that investors' WTP per ton of CO_2 is strongly correlated with this level of positive emotions, but not with investors' individual estimates of what it costs to save one ton of CO_2. This suggests that investors' valuation of impact is mainly driven by feelings rather than by calculation. We also explain how an emotional valuation may reconcile the results of the main experiment and of the joint evaluation extension.

Taking all our findings together, we suggest viewing the average pro-social investor as a "warm glow" optimizer,[4] rather than as a consequentialist who optimizes the impact of her or his investments. Our results leave open the possibility that some investors are sensitive

[4]Models of "warm glow" or "impure" altruism focus on emotional valuation. They assume that individuals do not derive utility from the well-being of others, but from an emotional response to the act of behaving pro-socially (Andreoni, 1990).

to impact under specific circumstances, and suggest that a more calculative decision mode can be encouraged by tweaking the choice environment. However, our combined results imply that a calculative appraisal of impact is second-order compared to the emotional "warm glow" investors derive from choosing a sustainable option. Ultimately this suggests that the positive emotions derived from choosing sustainable investments are an important driver of the trend for sustainable investing.

Our paper contributes to the literature that investigates the influence of pro-social preferences on investment decisions (Barber, Morse, and Yasuda, 2021; Bauer, Ruof, and Smeets, 2021; Hartzmark and Sussman, 2019; Riedl and Smeets, 2017). While we confirm previous findings of a substantial WTP for sustainable investments, our results show that this WTP does not scale with the level of impact that those investments offer. Our results suggest that investors' valuation of sustainable investments is more akin to charitable giving than to financial optimization. While scope insensitivity has been shown in other contexts (Desvousges, Johnson, Dunford, Boyle, Hudson, and Wilson, 1992; Hsee and Rottenstreich, 2004; Kahneman and Knetsch, 1992; Karlan and Wood, 2017; Metzger and Günther, 2019; Null, 2011), it has not been demonstrated in the context of financial decision-making, where it may well be the case that investors approach contributing to the public good in a more calculating, consequentialist manner. Our results demonstrate, however, that scope insensitivity is also a relevant issue when individuals express their pro-social preference in an investment context.

Our work is also related to the literature that explores the role of emotions in financial decision-making (Finucane, Alhakami, Slovic, and Johnson, 2000; Kuhnen and Knutson, 2011; Slovic, Finucane, Peters, and MacGregor, 2007). Affective decision-making has been put forward as an explanation for several puzzles in financial markets, including the home bias (Coval and Moskowitz, 1999; Huberman, 2001; Strong and Xu, 2003), or the IPOs of glamorous companies (MacGregor, Slovic, Dreman, and Berry, 2000). Hartzmark and Sussman (2019) suggest that emotions may also drive investors' valuation of sustainable investments. We confirm that investors' WTP for sustainable investments is positively correlated with the positive emotions they experience when choosing a sustainable investment option. This highlights that emotions play an important role in the behavior of pro-social investors.

Finally, we challenge a key assumption in the literature that explores the effects of pro-social preferences on asset pricing. A growing number of theoretical papers model how pro-social investors influence asset prices either because they have a taste for "green"

assets (Heinkel, Kraus, and Zechner, 2001; Pedersen, Fitzgibbons, and Pomorski, 2021; Pástor, Stambaugh, and Taylor, 2021) or because they explicitly care about aggregate externalities (Broccardo, Hart, and Zingales, 2020; Oehmke and Opp, 2019). These models suggest that pro-social investors, by expressing their preferences in the financial market, incentivize companies to reduce externalities. In essence these models assume that pro-social investors' utility increases with the impact of their investments, and that these investors make trade-offs between the financial performance and the impact of their investments. Our results challenge this assumption and suggest that pro-social investors are more likely to maximize financial performance while optimizing the warm glow that they derive from their choices. Modeling investor behavior in such a way would likely emphasize the importance of the structure of the sustainable investment industry, information asymmetry, and the way products are marketed to investors. Without measures in place that align the experience of warm glow with a product's underlying impact, sustainable investing may turn out to be a much less effective mechanism for curbing externalities than previously thought. Thus, future studies may yield important insights by modeling the behavior of pro-social investors in a way that explicitly reflects the role of warm glow.

Our paper is also related to three contemporary working papers. Humphrey, Kogan, Sagi, and Starks (2020) run an investment game where investment returns are positively or negatively linked to charitable benefits. They show that investors allocate less to investment options that entail negative effects on charities, but not more to those that entail positive effects. Although their paper addresses the positive/negative dichotomy and not different levels of impact, its results are consistent with ours in the sense that investors' valuation of externalities is not linear. Bonnefon, Landier, Sastry, and Thesmar (2022) implement an auction of claims on hypothetical corporations that donate parts of their profits to charity. The authors find that respondents' WTP scales with monetary contributions to charities in a linear relationship, irrespective of whether their investment is pivotal to a contribution. Brodback, Günster, and Pouget (2021) employ initial public offerings of assets that have identical financial payoffs but differ in the intensity and timing of their social responsibility aspects. The authors find that participants' WTP increases with positive externalities, implemented as donations to charity. Our findings may seem contradictory to those of Bonnefon, Landier, Sastry, and Thesmar (2022) and Brodback, Günster, and Pouget (2021), as in both works participants' WTP increases with the positive externalities of investment options. A key difference in the experimental setup, however, is that in both of these studies externalities are expressed in monetary units

and implemented as donations. This enables participants to compare their impact one-to-one to the costs of sustainable investments. Expressing impact in monetary terms may be conducive to a calculative valuation of externalities (Hsee and Rottenstreich, 2004), and it relieves investors of the difficult task of valuating impact. Our paper, meanwhile, examines the case where externalities are not monetized, a feature that is widespread in the market for sustainable investment products. In this case, we find that WTP does not scale with the level of impact in a linear fashion, and we suggest that investors' WTP is driven by an emotional valuation. Reading the findings of these different papers together with our results suggests that monetizing impacts could be an effective measure for avoiding scope insensitivity in sustainable investing.

In terms of practical implications, our results highlight that there is a severe risk of greenwashing in sustainable finance. The market for sustainable investing is expanding quickly, in 2020 surging past a total volume of USD 35 trillion (GSIA, 2021). This growth is raising hopes that sustainable investing might help tackle major societal challenges, such as curbing carbon emissions. However, the inconsistent WTP for impact that we demonstrate in this paper creates an incentive problem. If investors' WTP for sustainable investments scales with emotional warm glow rather than with impact, financial institutions have an incentive to create products that offer warm glow rather than impact. This is especially the case when offering impact comes at a cost.[5] Furthermore, our results suggest that financial institutions have an incentive to structure their offerings in such a way that sustainable products with little impact stand out as the most impactful option available. This may result in a market for sustainable investment products that benefits investors in terms of warm glow and financial institutions in terms of profits, but fails to fulfill its potential for solving societal problems.

IV.2 Study Design

We address the question of investor sensitivity to impact in framed field experiments and following a pre-registered experimental procedure.[6] All experiments were conducted between May and September 2020, and our robustness checks in April and May 2022.

[5] Although there are claims to the contrary, most theoretical models imply that pro-social investors need to accept lower financial performance ex ante in order to have impact (Heinkel, Kraus, and Zechner, 2001; Oehmke and Opp, 2019; Pástor, Stambaugh, and Taylor, 2021). In addition, there are—for sustainable investments—additional requirements for data and expertise that are likely to add to management fees.

[6] For pre-registration details, see https://aspredicted.org/w5f8i.pdf.

Based on a series of investment decisions, we assess investors' WTP for a sustainable investment compared to that for a "conventional" investment. Both investment options are presented as equity funds. The sustainable investment has some level of impact while the conventional investment has zero impact, but is otherwise equivalent. We vary the level of impact between subjects, allowing us to investigate how WTP responds to the level of impact. We measure investors' WTP in terms of the front-end fee that investors are willing to pay for the sustainable investment. The investment decisions are consequential, in terms of both payout and impact.

We operationalize the investments' impact in terms of carbon dioxide (CO_2) emissions savings brought about by investing in the product. We choose this measure for several reasons. First, CO_2 emissions and climate change are widely discussed in the financial press. Thus, most investors understand the topic and are familiar with the metric. Second, CO_2 emissions are clearly defined and measurable. Many metrics that are used in the context of sustainable investing are opaque in the sense that it is not clear what they measure and how they measure. In such cases investors might not respond to impact, because it is uncertain what the numbers mean. The measurement of CO_2 emissions, in contrast, is well established. Third, there are markets for CO_2 emissions savings, such as the European Emission Trading Scheme, which means that prices for CO_2 emissions are easily observable in the public domain. Finally, we are able to actually realize the impact of the sustainable investment through verified carbon emission reduction projects (see Section IV.A.1 for details). Thus, quantifying impact in terms of CO_2 emissions savings should allow investors to interpret impact in a quantitative sense.

Figure IV.1 illustrates the flow of the experimental procedure we use to elicit investors' WTP. It proceeds in four steps: instructions, information on investments, investment decisions, and a post-experiment survey.

IV.2.1 Instructions and Incentives

In the first step, investors receive detailed instructions on the investment decisions and on investor incentivization. We make sure that participants understand how the investment decisions work and that it is in their best interest to state their true preferences (see Figure IV.A.1).

The investment decisions are incentivized with relatively high stakes. For 10 randomly selected investors, we make a real €1,000 investment on their behalf, based on their

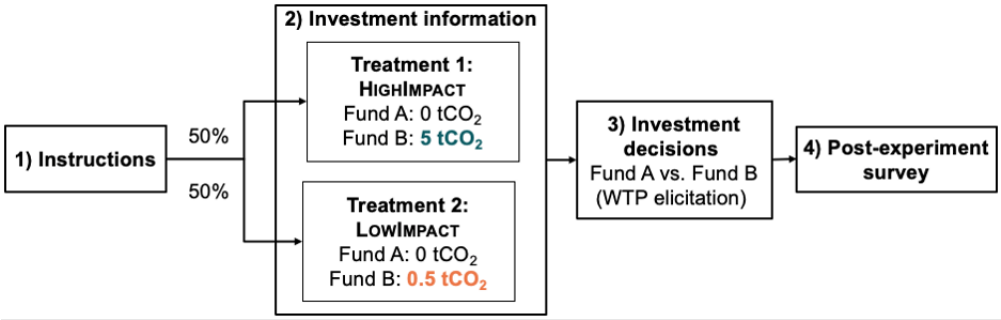


Figure IV.1. Experimental Procedure of the Main Experiment. This figure illustrates the four steps of the experimental procedure we use in our main experiment. In the second step, participants are randomly assigned either to the HIGHIMPACT treatment or the LOWIMPACT treatment.

investment decisions. To guarantee that participants reveal their true WTP, we apply the Becker–DeGroot–Marschak (BDM) mechanism (Becker, Degroot, and Marschak, 1964), a standard procedure in the judgment and decision-making literature. Using the BDM mechanism, we determine in which option to invest and a front-end fee, which we deduct from the investment amount (see Section IV.A.1 for the detailed procedure). After one year, we pay out the full value of this investment to the selected investors.

To familiarize investors with the decision procedure, we guide them through an example (see Figure IV.A.2). We also require them to complete a brief quiz in order to check whether they understand the potential consequences of their decisions. Investors who "fail" the quiz twice receive the correct answers and a short explanation.

IV.2.2 Investment Information

In the second step, investors receive information about the financial performance and the impact of the two investment options. For each of the two investments, we provide information on the asset class, the market segment, the annualized return over the last three years, and the risk level according to Morningstar's risk rating, all of which are identical for both investments. For the sustainable investment option, we additionally provide information on how much CO_2 emissions a €1,000 investment saves (see Figure IV.2 for an example).

To make sure that respondents understand the impact information, we translate the CO_2 savings into more intuitively comprehensible figures. We present the information in terms of trees planted, kilometers of air travel, and daily emissions of an average EU citizen to facilitate comprehension of the indicated amount of CO_2 emissions savings. These figures are in units that most respondents know from personal experience and can directly relate to. Also, to prevent experimenter demand effects—that is, "changes in behavior by experimental subjects due to cues about what constitutes appropriate behavior" (Zizzo, 2010)—in addition to using financial incentives we use an ambiguous framing of sustainable investing: in the information column on the right of Figure IV.2, we provide investors with arguments both for and against investing sustainably being socially desirable.

We randomly assign investors to one of two different treatments, HighImpact and LowImpact. In HighImpact, the sustainable investment saves 5 tons of CO_2, whereas in LowImpact it saves 0.5 tons of CO_2; so, 10 times less. To avoid ordering effects, we randomize whether the sustainable investment option is displayed on the screen's left or its right side. We do not, however, find any significant ordering effects in our results (see Table IV.A.1).

To guarantee that all relevant information is salient when the valuation decision is made, investors again need to participate in a brief quiz on the past performance, the risk level, and the impact of the two investments (see Figure IV.A.3 for a screenshot). Investors who twice fail to answer the quiz questions correctly receive the correct answers and a short explanation.[7]

IV.2.3 Investment Decisions and WTP Elicitation

In the third step, we elicit investors' WTP for the sustainable investment option. As the direct statement of a precise WTP is cognitively demanding for respondents and subject to noisy answers and outliers, we ask investors to make binary choices instead, which is the method most frequently used in the judgment and decision-making literature to measure (risk) preferences (Holt and Laury, 2014). Our respondents repeatedly choose between the sustainable investment and the conventional investment, which are neutrally labeled as investments A and B. For each investment we indicate a onetime, upfront fee, which we

[7]Excluding investors who fail twice in at least one of these quizzes does not substantially affect our results, as shown in Table IV.A.2.

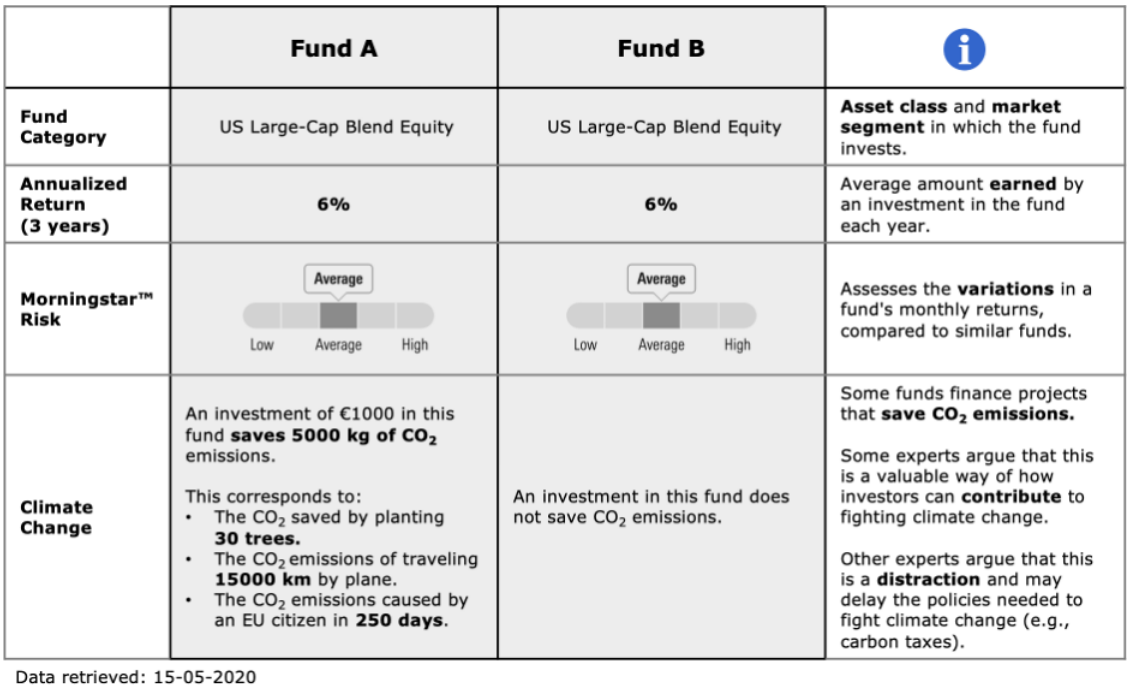


Figure IV.2. Investment Information in the Main Experiment. This figure provides a screenshot of the information the investors participating in our main experiment receive on the two investments if they are assigned to the HighImpact treatment. The investment information investors in the LowImpact treatment receive is shown in Figure IV.A.3.

vary between consecutive choices depending on respondents' answers. Both investment options start with the same fee, of €10. If a participant prefers investment A, we increase the fee for investment A by €40 and ask again. Using the bisection method, also called the midweight method, we iteratively adjust the fee to elicit an investor's WTP (see, e.g., Abdellaoui, 2000; van de Kuilen and Wakker, 2011). We provide an illustrative example in Figure IV.A.4, and additional details on the implementation of the bisection method in Section IV.A.1. Using this method we can determine the investors' WTP with a precision of €1.25 for the €1,000 investment through a series of seven choices.[8] To guarantee that our method yields each investor's true WTP, after the investment decisions have been made we ask respondents to confirm whether the elicited WTP actually reflects their true

[8]Some investors show a censored WTP for the sustainable investment: they do not deviate from the initially preferred investment in any of their seven choices. In this case, we ask these investors to directly state their WTP for the sustainable investment. Table IV.A.3 shows our main results excluding all investors with censored WTP; the exclusion of these investors does not have a material effect on our results.

preferences. Respondents who do not agree with the elicited WTP are asked to repeat the procedure once, if they wish. We exclude investors who disagree with the elicited WTP and are unwilling to repeat the investment decisions.

IV.2.4 Post-experiment Survey

After the respondents have taken their investment decisions they are asked to fill out a survey, which serves two purposes. First, we run a manipulation test to check whether investors understood and remember the investment information provided. We ask investors 1) to recall which investment had a higher impact and 2) how much impact the sustainable investment had exactly. Second, we ask questions about investors' financial expectations with regard to the investments, the feelings they associate with their choices, their perception of the impact of the sustainable investment, and their individual preferences, as well as their demographic characteristics. Table IV.A.4 summarizes all variables elicited in the post-experiment survey.

IV.2.5 Participants

We conduct our main experiment with a sample of experienced private investors. We recruit them from the members of a Dutch investor protection interest group with some 40,000 members.[9] The main activities of this group are the provision of independent information for investors and the coordination of lawsuits that aim to obtain compensation for groups of aggrieved shareholders. Its members hence have substantial experience of, and interest in, making investment decisions, which was our intent with this participant pool. For our experiments, we were able to recruit 527 participants via a general, regular newsletter sent to the members who subscribe to it. Of these participants, 219 take part in our main experiment; the remaining participants take part in two extensions of the main experiment. Table IV.1 shows the demographic characteristics and individual preferences of our sample of private investors in the main experiment. On average, our respondents are older and wealthier, and as a group have a higher share of males than that of the Dutch population. Both treatment groups, HighImpact and LowImpact, are well balanced in terms of demographic variables and individual preferences. We refer to these participants as private investors in the following.

[9]Vereniging van Effectenbezitters (Association of Stockholders).

Table IV.1
Preferences and demographics for the private investors

This table presents the preferences and demographic variables of the sample of private investors in our main experiment, by impact treatment. Time preferences, risk preferences, and altruism are measured on a 10-point scale using an experimentally validated survey module introduced by Falk, Becker, Dohmen, Huffman, and Sunde (2016). In order to improve readability, we transform other variables to a scale from 0 to 10. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments.

	Mean Values		**Mann–Whitney U Test**
	LowImpact ($n = 97$)	HighImpact ($n = 99$)	(HighImpact = LowImpact)
Risk preferences [0,10]	6.938	7.000	$p = 0.307$
Time preferences [0,10]	7.361	7.485	$p = 0.389$
Altruism [0,10]	6.588	6.455	$p = 0.389$
Climate awareness [0,10]	7.423	7.677	$p = 0.694$
Female [0,10]	0.103	0.131	$p = 0.540$
Age	61.660	61.495	$p = 0.700$
Income	€60,000–€74,999	€60,000–€74,999	$p = 0.842$
Net worth	€200,000–€499,999	€200,000–€499,999	$p = 0.887$
Highest education	Bachelor's degree	Bachelor's degree	$p = 0.765$
Investment knowledge [0,10]	6.318	6.234	$p = 0.661$

IV.2.6 Data Processing

In our main experiment, we exclude from our analysis six investors as they do not agree with the statement "Climate change is a serious problem that needs to be solved"—that is to say, investors that state an agreement level of 3 or less on a scale of 1 to 7. CO_2 savings are an inappropriate measure of impact for these investors, and we cannot detect how their WTP for sustainable investments relates to impact. We exclude a further 17 investors as they explicitly disagree with the elicited WTP and are unwilling to repeat the investment decisions, as previously described.[10] This results in a final sample of 196 investors. In accordance with our pre-registered procedure, we winsorize all WTP values at the 5 percent and 95 percent levels, to reduce the influence of extreme values.

[10] We include 19 investors who disagree with the elicited WTP but are willing to repeat the investment decisions. We use the WTP calculated based on these repeated decisions for these investors.

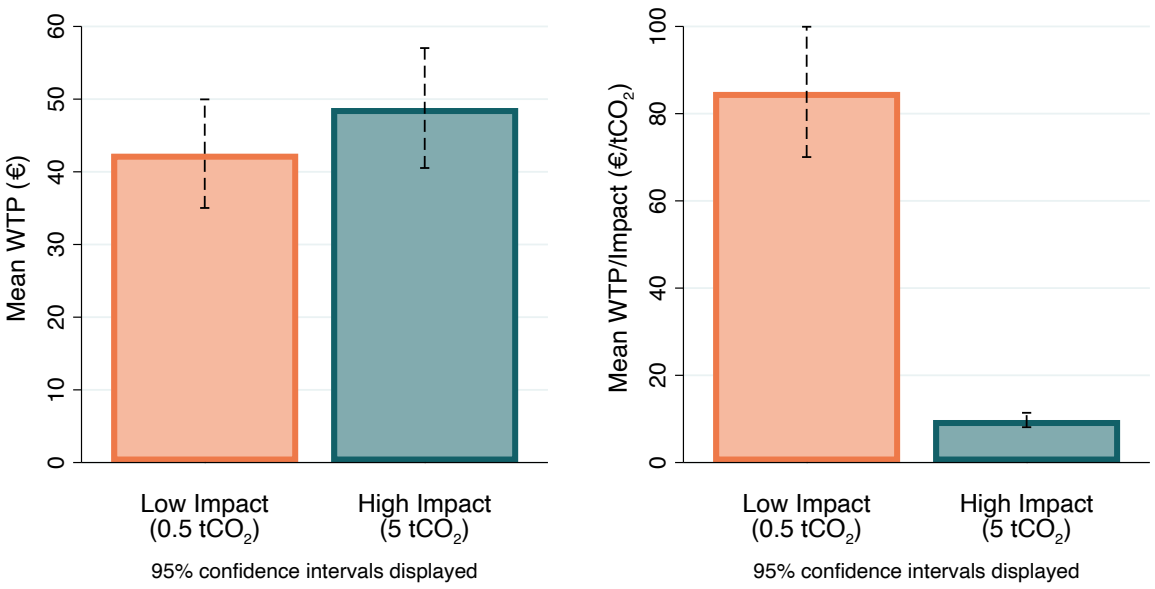
IV.3 Does WTP for Sustainable Investments Scale with Impact?

We observe that investors are willing to pay for investments with impact. Of all investors, 93 percent prefer the sustainable option when fees are equal between the two funds. Pooling investors in the LowImpact and the HighImpact treatment, the average WTP for the sustainable investment is €45.67 for a €1,000 investment. This substantial WTP is an important baseline finding for our investigation.

On this basis, we turn to our main question: Do investors have a higher WTP for investments with a higher impact? When contrasting the treatments, we find that the level of impact of sustainable investments does not significantly affect investors' WTP. Panel A of Figure IV.3 shows that investors' average WTP for the sustainable investment is €42.49 in the LowImpact and €48.78 in the HighImpact treatment. While investors do pay more in the HighImpact treatment, this difference is not significant ($p = 0.363$, Mann–Whitney U test). There is also no significant difference between the two treatments regarding the share of investors who prefer the sustainable investment ($p = 0.798$, Mann–Whitney U test). Also, a visual inspection of the distributions of private investors' WTP reveals no systematic difference between the treatments, as shown in Figure IV.A.6, Panel A.

Panel B of Figure IV.3 shows that the observed insensitivity to impact leads to a substantial inconsistency in investors' WTP per unit of impact. Investors are willing to pay significantly more per tCO_2 saved in the LowImpact treatment than in the HighImpact treatment ($p < 0.001$, Mann–Whitney U test). We observe a difference in the average WTP per ton of CO_2 saved of a factor of 9.7.

The post-experiment survey provides evidence that our results are not driven by differences in risk and return expectations for the sustainable investment between the HighImpact treatment and the LowImpact treatment. Such differences might conceal the influence of the investments' impact. As shown in Table IV.2, neither risk expectations nor return expectations differ significantly between the HighImpact treatment and the LowImpact treatment. Further, as shown in Table IV.A.5, neither investors' risk expectations nor their return expectations correlate significantly with their WTP for the sustainable investment, and the effect of investments' impact on investors' WTP remains insignificant when we control for risk and return expectations.



Panel A Panel B

Figure IV.3. Response of Private Investors' WTP to the Impact of Sustainable Investments. This figure shows the results of our main experiment, which investigates how private investors' WTP for sustainable investments responds to the impact of these investments. Panel A: mean absolute WTP for the sustainable investment, by impact treatment. Panel B: mean relative WTP for the sustainable investment, per ton of CO_2 saved, by impact treatment.

Summing up, we find that while investors have a substantial WTP for an investment with impact, they do not pay significantly more for more impact, even when the impact is increased by a factor of 10. This suggests that investors are quite willing to pay for sustainable investments, yet they have little regard for the amount of impact that such investments offer.

IV.3.1 Robustness Checks

We corroborate this main finding in several robustness checks. First, we confirm that our results are not driven by a lack either of the salience or of the comprehensibility of the impact information provided. Second, we replicate the experiment, with a focus on past financial performance, ensuring that our elicitation method can detect sensitivity when it is present. Third, we replicate the main experiment with a set of MTurk samples and test eight different variations of our elicitation method to address potential concerns that our

Table IV.2
Results of the main experiment

This table presents the results of the main experiment. First, it shows private investor's absolute and relative WTP for the sustainable investment, as well as the share of investors that prefer the sustainable investment when fees are equal. The WTP is elicited using the experimental procedure described in Section IV.2. Second, it shows the results of the post-experiment survey. In order to improve readability, we transform variables from the post-experiment survey to a scale from -10 to 10. For risk expectations, return expectations, and positive emotions, positive values indicate that investors have a more favorable view of the sustainable investment; negative ones that they have a more favorable view of the conventional investment. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments. Detailed descriptions of the variables can be found in Table IV.A.4.

	Mean Values		**Mann–Whitney U Test**
	LowImpact ($n = 97$)	HighImpact ($n = 99$)	(HighImpact = LowImpact)
Experimental Results			
WTP (€)	42.49	48.78	$p = 0.363$
WTP/Impact (€/tCO$_2$)	81.25	8.38	$p < 0.001$
Sustainable investment preference (%)	93.81	92.93	$p > 0.999$
Post-experiment Survey Results			
Risk expectations [-10,10]	-0.526	-0.051	$p = 0.382$
Return expectations [-10,10]	-0.312	-0.707	$p = 0.348$
Positive emotions [-10,10]	6.134	6.465	$p = 0.121$
Perceived investment impact [-10,10]	4.089	5.488	$p = 0.003$
General relevance impact [-10,10]	3.643	4.276	$p = 0.142$
General relevance impact level [-10,10]	2.474	2.896	$p = 0.457$
Estimated cost of saving CO$_2$ (€/tCO$_2$)	94.55	102.43	$p = 0.658$

results are an artefact of this method. Finally, we provide evidence that the COVID-19 crisis is unlikely to have affected our main results.

Is the Impact Information Salient and Comprehensible?

Relying on the post-experiment survey, we examine whether the impact information is salient and comprehensible. If investors were not sufficiently attentive to the information provided, or if investors were unable to evaluate the impact information, this may explain the observed insensitivity.

First, we find that the impact information provided to the investors was salient during the investment decisions. Once the investment decisions had been made, we asked all participants if they could remember the impact information. We find that 99 percent

could correctly identify the sustainable investment, and that 95 percent could, in a free text field, accurately reproduce its exact level of impact in tCO_2 saved.

Second, we observe that on average investors made a realistic estimate of the value of saving a ton of CO_2 emissions. In the post-experiment survey, we ask investors for an estimate of the price of saving a ton of CO_2 emissions. On average, the investors' estimate of CO_2-saving costs is €98.55 per ton, with no significant difference between the HighImpact treatment and the LowImpact treatment ($p = 0.658$, Mann–Whitney U test, 95% confidence interval: €77.08–€120.02). This is higher than the CO_2 prices in the European Union Emissions Trading System during our data collection period, which fluctuated roughly between €25 and €30. The values stated by investors do, however, correspond relatively well to estimates of the cost society incurs from carbon emissions. Based on a survey of a broad panel of climate scientists and economists, Pindyck (2019) estimates that the social cost of emitting a ton of CO_2 lies between \$80 and \$200. Besides the fact that we translate the CO_2 savings of the investments into more intuitively comprehensive units, this finding indicates that the information provided enables investors to evaluate the level of impact of the investments.

Finally, our results suggest that investors expect slightly different levels of impact conditional on the treatment. We asked investors whether they thought the sustainable investment makes a meaningful contribution to mitigating climate change ("Perceived investment impact" in Table IV.2). This variable is significantly higher in the HighImpact treatment than in the LowImpact treatment ($p = 0.003$, Mann–Whitney U test). This suggests that investors seem to understand, at least in an ordinal sense, that the HighImpact treatment offers a higher impact.

Are Our Findings Aligned with What Investors Say Is Important to Them?

We compare investors' WTP for impact with investors' statements about the importance of impact. In the post-experiment survey, we ask two questions to this end. First, we ask investors how important it is to them that their investments contribute to halting climate change ("General relevance impact" in Table IV.2). Second, we ask investors how important it is to them *how much* their investments contribute to halting climate change ("General relevance impact level" in Table IV.2).

Investors assign importance to an investment having an impact and to how much impact an investment has ($p < 0.001$, Wilcoxon signed-rank test). At the same time,

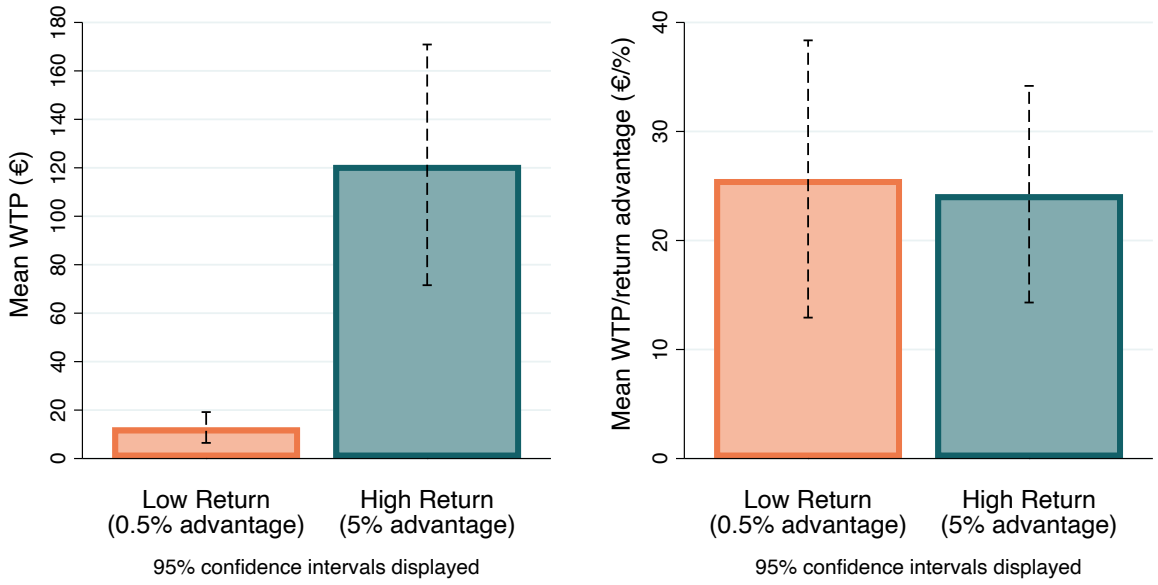
investors assign higher importance to the question of *whether* their investments contribute to climate change mitigation than to the question of *how much* their investments contribute to climate change mitigation ($p < 0.001$, Wilcoxon signed-rank test). These answers align with the fact that we find a WTP for investments with some impact but do not detect any significant differences in WTP between different levels of impact.

Can Our Elicitation Method Detect Sensitivity When It Is Present?

While so far only little is known about investors' sensitivity to impact, there is clear empirical evidence that investors are sensitive to mutual funds' past performance (e.g, Ivković and Weisbenner (2009)). Therefore, we apply our experimental procedure to measure investors' WTP for past financial performance. Concretely, we vary the investments' past performance, rather than their impact, between the two treatments. In each treatment the baseline option has a past performance of 5% per year. In the HighReturn treatment, the second investment outperforms the baseline by 5% (i.e., a total performance of 10% per year). In the LowReturn treatment, the second investment outperforms the baseline by only 0.5% (i.e., a total performance of 5.5% per year). Hence, in accordance with the main experiment, outperformance differs by a factor of 10 between the two treatments. We do not provide information on impact in this setup. Using the same recruitment campaign as for the main experiment, we have a sample of 89 private investors who we randomly assign to this robustness check.[11] Using the same method as in our main experiment, we measure investors' relative WTP for the outperforming investment in terms of additional fees they are willing to pay.

We observe that investors' WTP responds strongly to the past performance of investment options. Figure IV.4 shows that investors are willing to pay significantly more in the HighReturn treatment than in the LowReturn treatment (Panel A), and have a consistent WTP per unit of improved past performance across the treatments (Panel B). The average WTP for the outperforming investment is a factor of 9.5 higher in HighReturn compared to investing in LowReturn: €121.22 vs. only €12.82 ($p < 0.001$, Mann–Whitney U test). This difference is almost exactly proportional to the outperformance, which differs by a factor of 10 between the treatments. Based on these results, we conclude that our experimental design can detect investors' sensitivity when it is present.

[11] We exclude seven investors who explicitly disagree with the elicited WTP and are unwilling to repeat the investment decisions.



Panel A Panel B

Figure IV.4. Response of Private Investors' WTP to the Level of Outperformance of Investments. This figure shows the results of an extension of our main experiment, investigating how private investors' WTP responds to the level of outperformance of investments. Panel A: mean absolute WTP for the investment with a higher past return, by treatment. Panel B: mean relative WTP for the investment with a higher past return, by treatment, expressed relative to outperformance in percentage points.

Do Our Results Replicate with Different Elicitation Methods?

Potentially, our findings are affected by how we elicit investors' WTP. To explore the effect of our elicitation method on our results, we replicate our main experiment with a sample of 2,800 participants recruited through the crowd working platform Amazon Mechanical Turk (MTurk), implementing an exact replication as well as eight variations of our original setup. An overview of the results is shown in Table IV.3 and Figure IV.5. In short, the replications yield two main findings: First, the variations of the elicitation method change the average level of investors' WTP for the sustainable investment by up to ± 30%. Second, none of the variations produces a significant difference in WTP between the HighImpact and the LowImpact treatments.

Table IV.3
Replications of the main experiment

This table shows the key results of nine replications of the main experiment with MTurk samples. The table reports the number of observations, the pooled mean WTP across treatments, and the mean WTP per treatment. Further, it reports p-values of a Mann–Whitney U test, testing for differences in WTP between the treatments. The replication "R1: MTurk Baseline" replicates our main experiment using the original procedure. "R2: Randomized Risk & Return" introduces additional variation regarding the risk and return of the sustainable investment. In "R3: Direct Ask (No Anchor)" we directly ask for investors' WTP for the sustainable investment instead of using a discrete choice design. "R4: Direct Ask (Low Anchor)" and "R5: Direct Ask (High Anchor)" provide different anchors regarding investment fees before the direct ask. "R6: No Neutral Choice" presents investors with an initial choice where the sustainable investment already has a higher fee. "R7: Upward Elicitation" varies the bisection method: we start with a small fee increase and double the difference in fees in subsequent choices. "R8: Lower Scale" and "R9: Higher Scale" keep the original bisection method but decrease/increase the scale of fee differences by 50%. The detailed experimental procedures and pre-registrations for these replications can be found in the appendix under Section IV.A.2.

Replication	**N**	**Mean WTP ($)**			**Mann–Whitney U Test**
		Pooled	LowImpact	HighImpact	(HighImpact = LowImpact)
R1: MTurk Baseline	1000	43.91	42.97	44.93	$p = 0.412$
R2: Randomized Risk & Return	400	32.25	29.26	35.27	$p = 0.263$
R3: Direct Ask (No Anchor)	200	53.46	57.03	49.66	$p = 0.651$
R4: Direct Ask (Low Anchor)	200	33.89	34.60	33.19	$p = 0.093$
R5: Direct Ask (High Anchor)	200	37.06	35.45	38.71	$p = 0.321$
R6: No Neutral Choice	200	47.76	49.11	46.58	$p = 0.506$
R7: Upward Elicitation	200	33.02	35.35	30.87	$p = 0.611$
R8: Lower Scale	200	38.84	39.75	37.92	$p = 0.768$
R9: Higher Scale	200	55.22	48.30	63.18	$p = 0.084$

We pre-registered all replications. To ensure a high quality of observations from MTurk participants,[12] we used a vetted panel of reliable MTurkers provided by CloudResearch, we only accepted participants with outstanding reputations, we added an additional attention check, and we applied more stringent screening criteria with regard to our comprehension questions and the speed at which the experiment was conducted. We kept the variable incentive system constant for all replications, implementing a $1,000 investment for 10 randomly selected MTurk participants, based on their choices. [13] The detailed

[12] While data quality can be a concern with participants sourced on MTurk, experimental results have been shown to be reliable when appropriate quality control measures are applied (Aguinis, Villamor, and Ramani, 2021; Arechar, Gächter, and Molleman, 2018)

[13] As the MTurk participants are based in the U.S., we indicate all monetary amounts in U.S. dollars instead of euros in the replications with these participants.

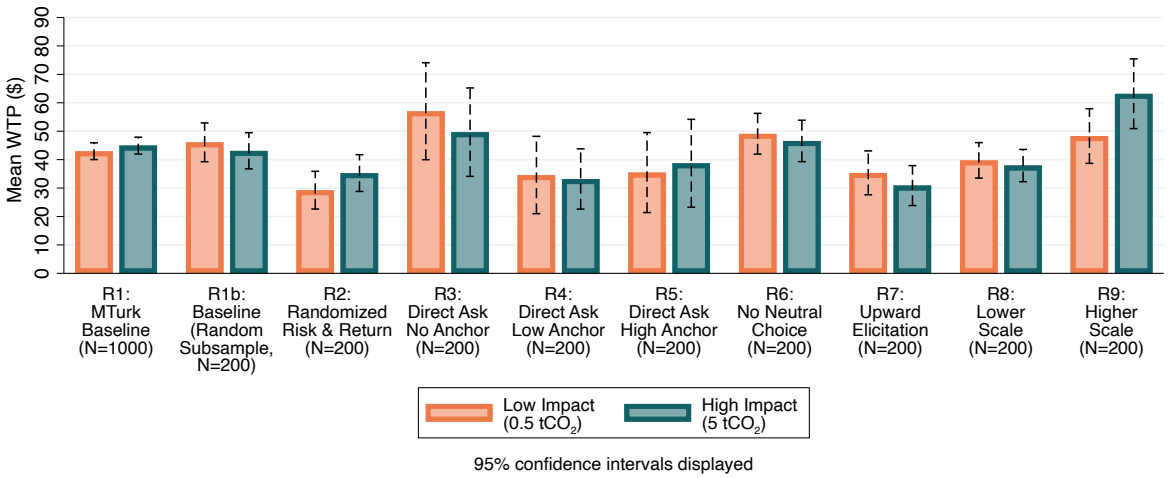


Figure IV.5. Summary of MTurk Results. This figure shows the key results of nine replications of the main experiment with MTurk samples. The bars show the mean WTP per treatment and the error bars denote 5% and 95% confidence intervals. The replication "R1: MTurk Baseline" replicates our main experiment using the original procedure. "R2: Randomized Risk & Return" introduces additional variation regarding the risk and return of the sustainable investment. In "R3: Direct Ask (No Anchor)" we directly ask for investors' WTP for the sustainable investment instead of using a discrete choice design. "R4: Direct Ask (Low Anchor)" and "R5: Direct Ask (High Anchor)" provide different anchors regarding investment fees before the direct ask. "R6: No Neutral Choice" presents investors with an initial choice where the sustainable investment already has a higher fee. "R7: Upward Elicitation" varies the bisection method: we start with a small fee increase and double the difference in fees in subsequent choices. "R8: Lower Scale" and "R9: Higher Scale" keep the original bisection method but decrease/increase the scale of fee differences by 50%. The detailed experimental procedures and pre-registrations for these replications can be found in the appendix under Section IV.A.2.

experimental procedure of the replications and pre-registration links can be found in the appendix under Section IV.A.2.

First, we show that our main results replicate well with MTurks and that our findings are unlikely to be driven by a lack of statistical power. In the experiment "R1: MTurk Baseline," we replicate our main experiment with a sample of 1,000 MTurk participants—so, a more than fivefold increase compared to the original sample size. The results of this replication are very similar to those of the main experiment. The pooled average WTP for the sustainable fund, of $43.91, we obtain for our MTurk sample is close to the original figure of €45.67 for the private investor sample. MTurk participants pay on average $1.96 more in the HighImpact treatment, but the difference between treatments

is insignificant ($p = 0.412$, Mann–Whitney U test). This indicates that our initial finding of a nonsignificant difference is not driven by a lack of statistical power. We also find that the quality of MTurk responses is high: 99 percent correctly remembered which investment had more impact, and 97 percent accurately reproduced the impact figures in the post-experiment survey. Taken together, these results indicate that our MTurk samples are a suitable reference via which to explore variations of the elicitation method in the following, MTurk experiments.

Second, in the variation "R2: Randomized Risk & Return" we explore potential experimenter demand effects that may be induced by our design choices, concluding that such effects are unlikely to affect our findings substantially. In our main experiment, the two investment options differ only in terms of their impact, while the information on risk and return are identical. While we take several measures to reduce experimenter demand, participants may infer from this design aspect what we may consider "appropriate" choices, which may lead to increased WTP estimates. To address this concern, we introduce additional variation in the investments' risk and return characteristics. Keeping the risk and return characteristics of the conventional investment constant, we interact the two impact treatments with two treatment conditions regarding the risk and return of the sustainable investments: in the first condition, the sustainable investment features a lower past return and a better risk profile; in the second condition it features a higher return and an inferior risk profile. As a result, there are four treatment conditions, and in each condition the sustainable investment differs from the baseline fund in all three dimensions (return, risk, and impact). This setup should make it harder for participants to infer what "appropriate" answers could be.

The average WTP for the sustainable investment across all treatments in this replication is \$32.25, which is significantly lower than in the baseline MTurk replication ($p = 0.001$, Mann–Whitney U test). Also, we detect a difference in WTP between the LowImpact treatment (\$29.26) and the HighImpact treatment (\$35.27), similar to the one in our main experiment. However, in this specification too the difference is not significant ($p = 0.263$, Mann–Whitney U test). Interestingly, participants report a lower level of positive emotions experienced from selecting this investment ($p < 0.001$, Mann–Whitney U test) compared to the baseline MTurk replication, which serves as a possible explanation as to why the pooled WTP is lower in this specification. We conclude that while we cannot rule out that the focus on impact differences in our original procedure

leads to somewhat inflated WTP estimates, our results hold in a setting where this focus is substantially reduced.

Third, we explore how the specifications of the method we use to elicit investors' WTP affect our results, concluding that our findings essentially replicate with a broad set of alternative specifications. While in three variations participants' average WTP for the sustainable investment is slightly higher in the HighImpact treatment than in the LowImpact treatment, it is somewhat lower in four variations, so that there is no clear tendency. None of the differences between the two impact treatments are significant.

In a first set of variations, we omit the bisection method that confronts investors with consecutive binary choices and instead directly ask for investors' WTP for the sustainable investment. We implement three versions of this: in "R3: Direct Ask (No Anchor)" we do not provide any anchoring information on costs associated with investments, in "R4: Direct Ask (Low Anchor)" we provide the information that a passively managed fund may charge an annual fee of 0.1 percent per year, and in "R5: Direct Ask (High Anchor)" we provide the information that an actively managed fund may charge a 1 percent fee per year. The direct ask elicitation replication without anchor produces a significantly higher pooled WTP estimate than the baseline MTurk replication (\$53.46, $p = 0.013$, Mann–Whitney U test). Conversely, the average pooled WTP is significantly lower for the replications with lower (\$33.89) and higher (\$37.65) anchors than in the MTurk baseline replication (in both cases $p < 0.001$, Mann–Whitney U test). In none of the direct ask variations do we find a significant difference in investors' WTP between the LowImpact and the HighImpact treatment.

In a second set of four replications, we vary several specifications of our discrete choice method. In the replication "R6: No Neutral Choice" we omit the first discrete choice of our original procedure, where investors decide between the two investments with fees being equal. In the main experiment, most investors prefer the sustainable investment under these conditions. Thus, they may feel obliged to select the same investment in subsequent choices where it is more costly, which would inflate our WTP estimates. This, however, seems unlikely as removing the first choice leads to a slightly (but not significantly) higher pooled average WTP compared to the baseline replication (\$47.76, $p = 0.181$, Mann–Whitney U test).

Next, we vary the bisection approach in the replication "R7: Upward Elicitation." In our original procedure, investors that have selected the sustainable investment in the first choice have to decide whether they are willing to pay \$40 more in fees for this investment

in the next choice. Again, investors may feel urged to accept this relatively high level of fees to remain consistent with their initial choice. Thus, in this variation, we start with a much smaller fee increase of \$1.25. In subsequent choices, we double the difference in fees between the two investments until the participants deviate from their initial choice. Once a participant has switched, we iteratively reduce the fee difference as in the original method to obtain a more precise estimate. With this method, investors' WTP for the sustainable investment is significantly lower than in the baseline replication, but remains substantial (\$33.02, $p < 0.001$, Mann–Whitney U test).

Finally, in the replications "R8: Lower Scale" and "R9: Higher Scale" we keep the original bisection method but decrease or, respectively, increase the initial fee difference by 50 percent. Thus, in their second choice participants have to decide whether they remain with their preferred investment if we increase its fee by \$20 ("R8: Lower Scale") or by \$60 ("R9: Higher Scale"). Decreasing our measurement scale by 50 percent leads to a slightly but significantly lower pooled average WTP for the sustainable investment than in the baseline replication (\$38.84, $p = 0.035$, Mann–Whitney U test). However, this difference is not proportional to the reduction in our measurement scale. When increasing our measurement scale by 50 percent, participants' pooled average WTP is substantially higher than in our baseline replication (\$55.22), but the difference is not significant ($p = 0.112$, Mann–Whitney U test). Again, we do not find a significant difference between the LowImpact and the HighImpact treatments in any of the variations concerning the specifications of our discrete choice method.

The MTurk replications highlight that it is important to interpret the level of our WTP estimates in the context of our experimental design and data processing procedures. The average WTP for the sustainable investment across both impact treatments of our main experiment is €45.67. This corresponds to a fee of roughly 4.5% of the investment, which might seem relatively high. More conservative elicitation methods yield pooled WTP estimates that correspond to fees of 3–3.5% of the investment ("R2: Randomized Risk & Return," "R4: Direct Ask (Low Anchor)," and "R7: Upward Elicitation"). Also, the data processing affects the level of our WTP estimates. In the pre-registration, we commit to retaining observations from participants with censored WTP,[14] and to mitigate the effect of outliers by winsorizing observations at the 5% and 95% level. If we deviate from the pre-registered procedure and exclude investors with censored WTP, we obtain a pooled

[14]Investors' WTP is censored if they do not deviate from the initially preferred investment in all seven investment choices. We cannot elicit WTP directly from the choices of these investors. We additionally ask these investors to state their WTP for the sustainable investment.

average WTP corresponding to a fee of roughly 3% for the main experiment (see IV.A.3). In addition, it is relevant to consider that we measure WTP as an upfront fee for a one-year €1,000 investment. Retail investors might be used to relatively high entry and exit fees.[15] We conclude that our WTP estimates are somewhat sensitive to the elicitation method and data processing protocol, and that investors' WTP in the field might be lower than our estimates suggest. Importantly, however, our main result of insignificant treatment effects remains robust across different elicitation methods.

Are Our Results Affected by the COVID-19 Crisis?

The COVID-19 crisis has been found to have affected investors' behavior in various ways (see, e.g., Ramelli and Wagner (2020)). As our data collection took place mid-2020, the pandemic may also have affected our results. To investigate the effect of the crisis, we make use of a preliminary version of our experiment, which we ran with 311 students at a large Dutch university in September 2019, well before the emergence of SARS-CoV-2. We repeated this experiment in September 2020, after the emergence of SARS-CoV-2, under the same conditions and with a corresponding sample of 243 students at the same university. We find no substantial differences between the results before and after the appearance of the virus. The detailed results can be found in Table IV.A.6. In light of these results, it seems unlikely that COVID-19 had a relevant effect on our results.

IV.3.2 Discussion

We find that investors are willing to pay to obtain a sustainable investment with some impact, but do not pay significantly more for more impact. We provide evidence that this finding is unlikely to result from methodological issues or from the fact that investors did not understand the impact information provided to them. We do not rule out completely that some investors have some degree of sensitivity to impact, but this seems to be a second- or third-order effect. What drives investors' WTP is their preference for a sustainable investment with some impact over a conventional investment with no impact.

[15] Entry and exit costs and brokerage charges can be substantial in the retail investing market. Khorana, Servaes, and Tufano (2009) estimate that the average shareholder cost for a five-year investment period is 2.09% per year, of which 0.66% are entry and exit fees. Adjusted for a one-year investment period, total costs would amount to 4.76%. According to more recent (2018) data the average entry fee for equity funds in Europe was 3.65% (https://ec.europa.eu/info/publications/180425-retail-investment-products-distribution-systems_en).

This finding extends existing experimental studies in important ways. Prior studies have demonstrated that there is a WTP for sustainable investments (Barber, Morse, and Yasuda, 2021; Bauer, Ruof, and Smeets, 2021; Riedl and Smeets, 2017), and that there are substantial financial flows toward sustainable investment funds (Hartzmark and Sussman, 2019). These studies do not, however, explore sensitivity to the level of impact. Our results suggest that investors have a positive WTP for sustainable investments as a category, but do not adjust their WTP according to the impact of these investments. This dovetails with an earlier finding of Riedl and Smeets (2017), who show that pro-social preferences explain whether investors invest in sustainable funds, but do not explain how much of their wealth investors allocate to sustainable funds. While Riedl and Smeets (2017) treat "sustainability" as a binary attribute of investments, we complement their evidence by differentiating between different levels of "how" sustainable investments are and by showing how investors' WTP responds to this attribute of investments.

Our findings are in contrast with traditional models of altruistic behavior. These models assume that individuals are consequentialists, in the sense that they contribute to public goods because they derive utility from the level of the public good, beyond the direct benefit they experience themselves from the good (Andreoni and Miller, 2002; Becker, 1974; Eckel and Grossman, 1996). Such models, which are often labeled models of "pure" altruism, imply that the benefit an individual receives from performing a pro-social act depends on the act's impact on a public good. Thus, if pro-social investors were driven by pure altruism, we would expect their WTP for sustainable investments to increase with the impact of such investments. Yet, as our results show, even an increase in impact by a factor of 10 does not lead to a significant increase in investors' WTP. This observed behavior is not in line with a pure altruism decision model.[16]

Insensitivity to quantity has been demonstrated in other contexts, such as public good valuation and philanthropic donations. However, it is important to reconsider this phe-

[16] Pure altruism could possibly explain the insensitivity we observe if the marginal societal utility of CO_2 emissions savings is strongly decreasing. This, however, seems unlikely in our setting. The "pure" altruist's marginal utility is proportional to the societal utility of an additional unit of impact. Thus, to explain the observed behavior one would have to assume that marginal societal utility is high up until 0.5 tons of CO_2 emissions savings and strongly decreases between 0.5 and 5 tons of CO_2 emissions savings. Such a case is hard to make from a societal perspective. A recent report by the United Nations Environment Programme (UNEP) concludes that annual global greenhouse gas emissions need to be reduced by 32 billion tCO_2 by 2030 if the internationally agreed target of limiting global warming to 1.5 degrees Celsius above preindustrial levels is to be reached (UNEP, 2020). In light of this emission reduction gap, it seems highly unlikely that the marginal societal benefit decreases substantially for impact levels below 5 tCO_2, which is the highest impact level we use in our experiment.

nomenon in the context of sustainable investing. First, the classic contingent valuation result by Desvousges, Johnson, Dunford, Boyle, Hudson, and Wilson (1992) that individuals are willing to pay roughly the same amount to save 2,000, 20,000, or 200,000 birds relies on stated preferences. In a revealed preference setting where choices are consequential, individuals might evaluate their options more critically. Second, investment decisions tend to trigger a calculative decision mode, and thus public goods may be considered differently when they are part of an investment choice. Third, theorists currently model the behavior of sustainable investors as if they were consequentialists. Fourth, two contemporaneous working papers find that, within their experimental setups, respondents' WTP for sustainable investments scales with positive externalities (Bonnefon, Landier, Sastry, and Thesmar, 2022; Brodback, Günster, and Pouget, 2021). Against this backdrop, our rejection of the consequentialist view of sustainable investors is an important piece of evidence. In the following, we explore reasons why investors' WTP does not significantly respond to the level of impact.

IV.4 Exploring Reasons Why WTP for Sustainable Investments Does Not Scale with Impact

From the existing literature, we identify three potential explanations for the observed behavior. First, investors may lack the knowledge or experience necessary to evaluate impact information in an investment context. Second, investors may only be able to discriminate between products in terms of impact when they can directly compare several options. Third, investors' WTP may be driven by positive emotions, or, the "warm glow" that is associated with choosing a sustainable option, rather than by a calculative appraisal of that option's impact.

IV.4.1 Do Investors Lack the Necessary Experience to Evaluate the Impact of Investments?

Investors' lack of experience in evaluating impact information in an investment context could explain why their WTP for sustainable investments does not respond to the impact of those investments. We have already put several measures in place to ensure that

investors understand the impact of the available investment options. Nevertheless, investors may still be unable to valuate impact in the context of an investment decision. Even if they are able to differentiate between the impact of planting three trees and the impact of planting 30 trees, they may still be unable to evaluate whether planting 30 trees has an impact that is meaningful for a €1,000 investment. Research shows that the ability to perform such evaluations can increase with experience (Hsee and Zhang, 2010).

To investigate the effect of experience on sensitivity to impact, we repeat our experiment with a unique sample of dedicated impact investors. We recruit this sample through different channels: First, from the alumni network of the University of Zurich's training programs for high-net-worth impact investors. Second, via two associations of high-net-worth impact investors,[17] whose members come together to share knowledge and participate in events on impact investing. Third, from among the customers of five wealth advisors and private banks; those recruited were identified as experienced impact investors by their advisers.[18] In total, we recruited 125 impact investors through these channels.[19] The final sample comprises 118 investors as we excluded one investor who does not think that climate change is a serious problem and six who explicitly disagreed with the elicited WTP and were unwilling to repeat the investment decisions.

All of the investors in this sample have indicated their intention of exerting a positive impact with their investments, by taking courses, joining a network, or by instructing their advisors. Most have considerable experience in dealing with impact investments. While their individual levels of experience vary, we are confident that on average they have a much higher level of experience than the sample of private investors we recruited for our main experiment. If lacking the ability to evaluate impact information drives insensitivity to impact, we would expect these impact investors to be more sensitive to impact than the private investors are.

Table IV.4 presents the characteristics of the impact investor sample. Compared to the private investors in the main experiment, the impact investors have greater wealth, state a higher level of investment experience, and are younger, and the sample has a larger fraction of female investors. The median household net worth lies between €1

[17] Toniic and the NEXUS Working Group on Impact Investments.

[18] Credit Suisse, Bank Vontobel, Bank Julius Baer, Tiedemann Advisors, and VALUEworks.

[19] Five impact investors participated in the experiment shortly after the end of the sampling period specified in the pre-registration. If we exactly follow the pre-registration procedure, the sample is slightly smaller ($n = 120$). However, the results do not substantially differ from those obtained with the full sample (see Figure IV.A.5). As the sampling period does not seem to have an effect on our results, we analyze the entire sample.

Table IV.4
Preferences and demographics for the impact investors

This table presents the preferences and demographic variables of our sample of impact investors, by impact treatment. Time preferences, risk preferences, and altruism are measured on a 10-point scale using an experimentally validated survey module introduced by Falk, Becker, Dohmen, Huffman, and Sunde (2016). In order to improve readability, we transform other variables to a scale from 0 to 10. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments.

	Mean Values		**Mann–Whitney U Test**
	LowImpact ($n = 59$)	HighImpact ($n = 59$)	(HighImpact = LowImpact)
Risk preferences [0,10]	7.169	6.898	$p = 0.521$
Time preferences [0,10]	8.508	8.068	$p = 0.119$
Altruism [0,10]	7.763	7.169	$p = 0.131$
Climate awareness [0,10]	9.096	8.983	$p = 0.814$
Female [0,10]	0.356	0.407	$p = 0.705$
Age	41.424	38.966	$p = 0.456$
Income	€125,000–€149,999	€150,000–€174,999	$p = 0.543$
Net worth	€1M–€9.9M	€1M–€9.9M	$p = 0.931$
Highest education	Master's degree	Master's degree	$p = 0.828$
Investment knowledge [0,10]	6.877	6.707	$p = 0.650$

million and €10 million.[20] Regarding their preferences, the impact investors are more altruistic and have a stronger long-term orientation than our private investors. Again, both treatment groups, HighImpact and LowImpact, are well balanced in terms of demographic variables and individual preferences.

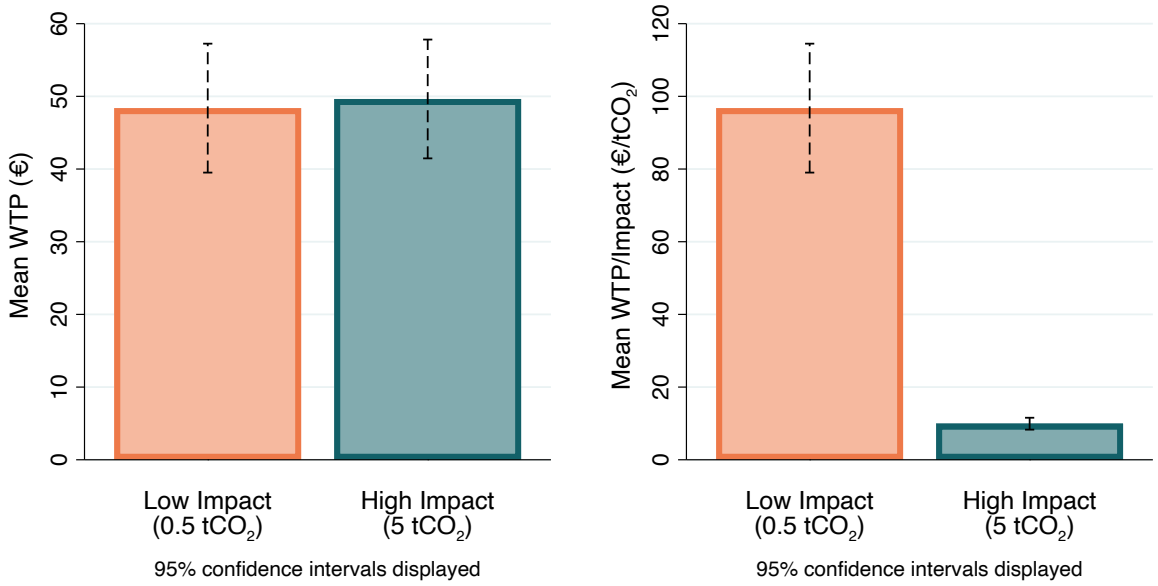
The results of our experiment with dedicated impact investors are overall very similar to those of the main experiment with private investors. We find that impact investors too have a positive WTP for sustainable investments. Of all our impact investors, 97 percent prefer the sustainable investment when fees are equal. Pooling investors in the LowImpact and the HighImpact treatments, the average WTP for the sustainable investment is €49.01 for a €1,000 investment. This is slightly more than the figure for the private investors; the difference, however, is not significant ($p = 0.096$, Mann–Whitney U test).[21]

[20] This figure may be an understatement in numerous cases as many of the impact investors are embedded in family structures that collectively own much more, often several billion euros.

[21] A similar reasoning as above applies to the magnitude of the measured WTP.

Further, we find that for impact investors too the level of impact of sustainable investments does not significantly affect their WTP. Figure IV.6, Panel A contrasts impact investors' WTP for sustainable investments between the LowImpact and HighImpact treatments. There is no significant difference in the WTP for the sustainable investment between the treatments ($p = 0.767$, Mann–Whitney U test, Table IV.5). In the LowImpact treatment, impact investors have an average WTP of €48.38 for an investment that saves 0.5 tCO_2, while in the HighImpact treatment the average WTP for an investment that saves 5 tCO_2 is €49.64. The distribution of impact investors' WTP for the sustainable investment can be found in Figure IV.A.6, Panel B.

Figure IV.6, Panel B shows that for the impact investors too the WTP per unit of impact is inconsistent between the treatments. Impact investors are willing to pay significantly more per tCO_2 saved in the LowImpact treatment than in the HighImpact treatment ($p < 0.001$, Mann–Whitney U test). The difference in the average WTP per ton of CO_2 is of a factor of 9.2.



Panel A Panel B

Figure IV.6. Response of Impact Investors' WTP to the Impact of Sustainable Investments. This figure shows the results of our experiment investigating how impact investors' WTP for sustainable investments responds to the impact of these investments. Panel A: mean absolute WTP for the sustainable investment, by impact treatment. Panel B: mean relative WTP for the sustainable investment, per ton of CO_2 saved, by impact treatment.

As with the private investors, neither the risk expectations nor the return expectations of impact investors differ significantly between the HighImpact treatment and the LowImpact treatment (Table IV.5). In comparison to the private investors, the impact investors have more positive expectations with regard to the financial performance of the sustainable investment. The impact investors expect the sustainable investment to have slightly better returns ($p = 0.047$, Mann–Whitney U test) and lower risk ($p = 0.006$, Mann–Whitney U test) than do the private investors.

Table IV.5
Results of the experiment with impact investors

This table presents the results of the experiment with impact investors. First it shows impact investors' absolute and relative WTP for the sustainable investment, as well as the share of investors that prefer the sustainable investment when fees are equal. The WTP is elicited using the experimental procedure described in Section IV.2. Second it shows the results of the post-experiment survey for the sample of impact investors. In order to improve readability, we transform variables from the post-experiment survey to a scale from -10 to 10. For risk expectations, return expectations, and positive emotions, positive values indicate that investors have a more favorable view of the sustainable investment, negative ones that they have a more favorable view of the conventional investment. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments. Detailed descriptions of the variables can be found in Table IV.A.4.

	Mean Values		**Mann–Whitney U Test**
	LowImpact	HighImpact	(HighImpact = LowImpact)
	($n = 59$)	($n = 59$)	
Experimental Results			
WTP (€)	48.38	49.64	$p = 0.767$
WTP/Impact (€/tCO$_2$)	96.76	9.93	$p < 0.001$
Sustainable investment preference (%)	96.61	98.31	$p > 0.999$
Post-experiment Survey Results			
Risk expectations [-10,10]	0.678	0.593	$p = 0.991$
Return expectations [-10,10]	0.169	0.254	$p = 0.952$
Positive emotions [-10,10]	7.797	6.864	$p = 0.209$
Perceived investment impact [-10,10]	3.898	5.085	$p = 0.314$
General relevance impact [-10,10]	6.158	6.158	$p = 0.820$
General relevance impact level [-10,10]	5.763	4.746	$p = 0.182$
Estimated cost of saving CO$_2$ (€/tCO$_2$)	404.57	291.47	$p = 0.258$

Taken together, the results of our experiment with impact investors demonstrate that a lack of experience in evaluating the impact of investments is an unlikely explanation for the observed insensitivity to impact. We find that even dedicated and experienced impact investors do not respond, via their WTP, to different levels of investment impact.

We therefore conclude that it is not a mere lack of experience that drives insensitivity to impact. This finding prefigures one important implication—namely, that investor training and the building up of experience may not be sufficient to address the problems that come with the observed insensitivity to impact.

IV.4.2 Does Comparability Increase Investors' Sensitivity to Impact?

The choice investors face in our main experiment corresponds to one that many retail investors face when their bank advisor offers them a binary choice between a conventional and a sustainable investment product. However, this may not necessarily correspond to the choice faced by more experienced investors, who can evaluate a broader set of investment options. As demonstrated by Hartzmark and Sussman (2019), these investors may consider information sources like the Morningstar "Globe" Rating, which allow them to compare a range of different investment options.

To investigate the importance of relative comparison, we run our experiment in a joint evaluation setup. Investors receive information on three investments: Fund A has zero impact, Fund B has a comparatively low level of impact, and Fund C has a considerably higher level of impact. Figure IV.7 illustrates the experimental procedure of our joint evaluation extension. We divide investors into two treatments, which we denote as the HighImpactRange treatment and the LowImpactRange treatment. In the HighImpactRange treatment, Fund B saves 0.5 tCO_2 and Fund C saves 5 tCO_2. This corresponds to the impact values of the two treatments in our main experiment. In the LowImpactRange treatment, Fund B only saves 0.05 tCO_2 and Fund C only saves 0.5 tCO_2. Again, for each treatment we translate this impact into more intuitively comprehensible units, as shown in Figure IV.8 for the treatment HighImpactRange.

For each investor, we assess the WTP for Fund B and for Fund C one after the other relative to Fund A, using the same procedure as in the main experiment. Hence, once one WTP is determined, the participant goes through the same mechanism again with the other sustainable investment in comparison to Fund A. During both WTP elicitations, investors always see the information on all three investments, as illustrated in Figure IV.8. The order in which WTP is elicited, first for Fund B or first for Fund C, is randomized.

We run the experiment with 219 additional private investors from the same recruitment campaign that we used for the main experiment. We exclude 11 investors who do

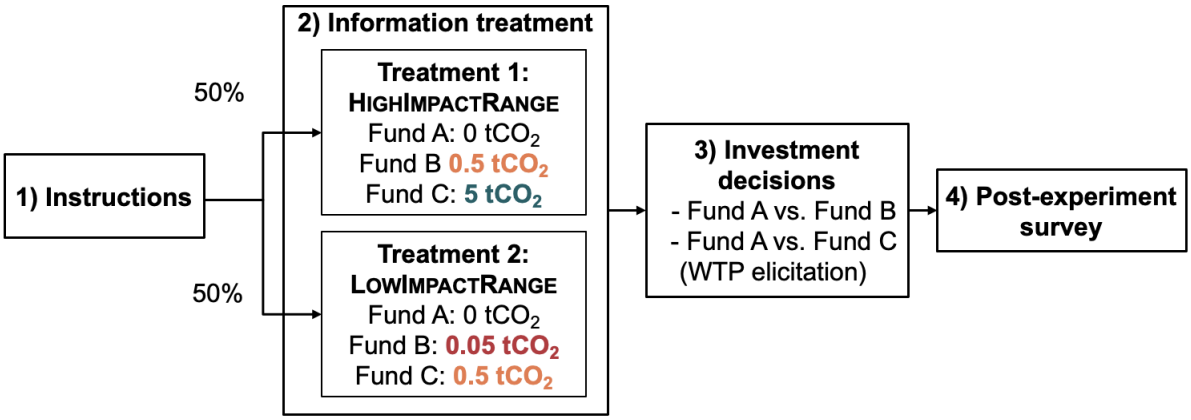
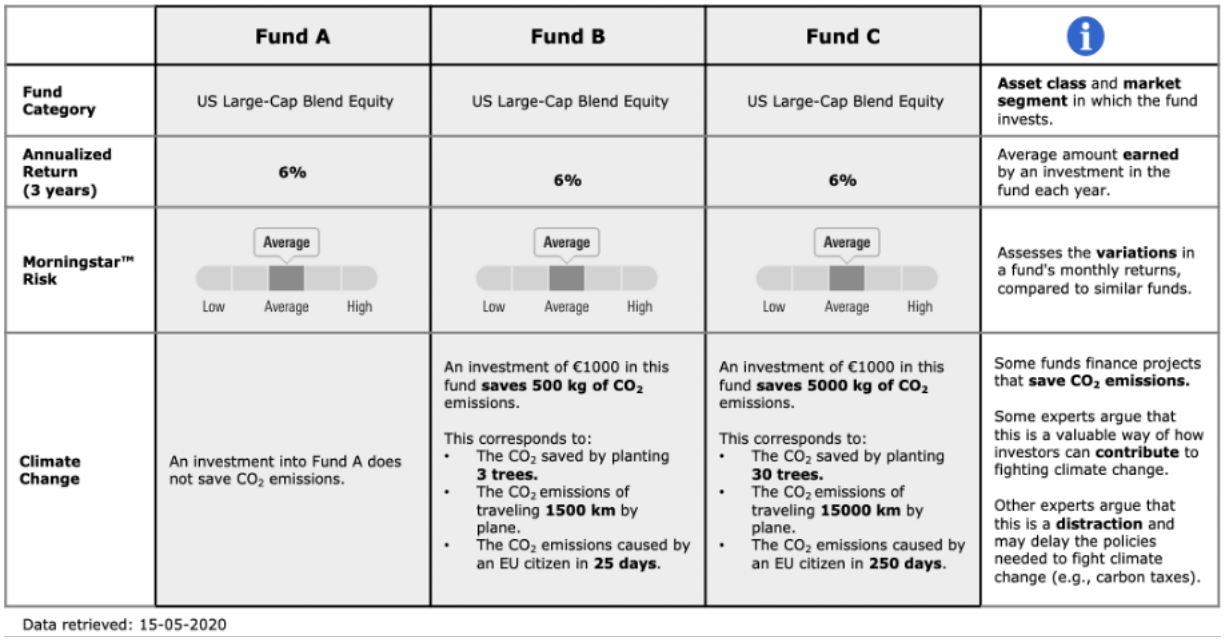


Figure IV.7. Experimental Procedure of the Joint Evaluation Extension. This figure provides an illustration of the experimental procedure we use in the joint evaluation extension of our main experiment.

not agree that climate change is a serious problem and 11 who explicitly disagree with the detected WTP and are unwilling to repeat the investment decisions. This results in a final sample of 197 investors.

The findings of the joint evaluation extension show that when comparing different sustainable investments investors' WTP reacts to some extent to differences in impact (Figure IV.9). The results of the HIGHIMPACTRANGE treatment, in which investors see the same sustainable investments we use in our main experiment, provide evidence that increased comparability leads to a significant difference in WTP between the two sustainable investments ($p < 0.001$, Wilcoxon signed-rank test). The mean WTP is €31.09 for the sustainable investment that saves 0.5 tCO_2 and €40.07 for the sustainable investment that saves 5 tCO_2. While this indicates that investors respond to differences in impact in the case of directly and easily comparable options, their sensitivity remains limited: an increase in impact by a factor of 10 increases investors' WTP by only 28 percent. Thus, the value investors assign to a unit of impact remains inconsistent; there is a significant difference in the WTP per unit of impact between the two available sustainable investments ($p < 0.001$, Wilcoxon signed-rank test).

Further, the results show that even in the joint evaluation setup, investors' WTP does not respond to the absolute level of impact. The results in the LOWIMPACTRANGE treatment are very similar to those observed in the HIGHIMPACTRANGE treatment: investors' WTP for the sustainable investment that saves 0.05 tCO_2 is €28.01; the figure is €36.89 for the investment that saves 0.5 tCO_2. Despite the tenfold difference in impact,



	Fund A	Fund B	Fund C	
Fund Category	US Large-Cap Blend Equity	US Large-Cap Blend Equity	US Large-Cap Blend Equity	**Asset class** and **market segment** in which the fund invests.
Annualized Return (3 years)	**6%**	**6%**	**6%**	Average amount **earned** by an investment in the fund each year.
Morningstar™ Risk	Average Low Average High	Average Low Average High	Average Low Average High	Assesses the **variations** in a fund's monthly returns, compared to similar funds.
Climate Change	An investment into Fund A does not save CO_2 emissions.	An investment of €1000 in this fund **saves 500 kg of CO_2** emissions. This corresponds to: • The CO_2 saved by planting **3 trees.** • The CO_2 emissions of traveling **1500 km** by plane. • The CO_2 emissions caused by an EU citizen in **25 days**.	An investment of €1000 in this fund **saves 5000 kg of CO_2** emissions. This corresponds to: • The CO_2 saved by planting **30 trees.** • The CO_2 emissions of traveling **15000 km** by plane. • The CO_2 emissions caused by an EU citizen in **250 days**.	Some funds finance projects that **save CO_2 emissions.** Some experts argue that this is a valuable way of how investors can **contribute** to fighting climate change. Other experts argue that this is a **distraction** and may delay the policies needed to fight climate change (e.g., carbon taxes).

Data retrieved: 15-05-2020

Figure IV.8. Screenshot of the Investment Information in the Joint Evaluation Extension. This figure provides an example of the information the investors receive with regard to the three investments in the joint evaluation extension of our experiment. The screenshot corresponds to the investment information investors in the HighImpactRange treatment receive. The investment information investors in the LowImpactRange treatment receive can be found in Figure IV.A.5.

we find no significant difference if we compare the WTP for the more impactful sustainable investment (Fund C) between the HighImpactRange and the LowImpactRange treatment ($p = 0.394$, Mann–Whitney U test). There is also no significant difference if we compare the WTP for the sustainable investment with a lower impact (Fund B) between the HighImpactRange and the LowImpactRange treatments ($p = 0.273$, Mann–Whitney U test).

On average, investors' WTP is lower in the joint evaluation extension than in our main experiment. The pooled average WTP over both sustainable investments in the HighImpactRange treatment of the joint evaluation experiment is €34.12. This is significantly lower than the pooled average over both impact treatments in the main experiment, where the sustainable funds feature the same levels of impact ($p < 0.001$, Mann–Whitney U test).

The joint evaluation allows us to analyze within-subject sensitivity to the impact of investments, and we observe substantial heterogeneity in individual investors' sensitivity.

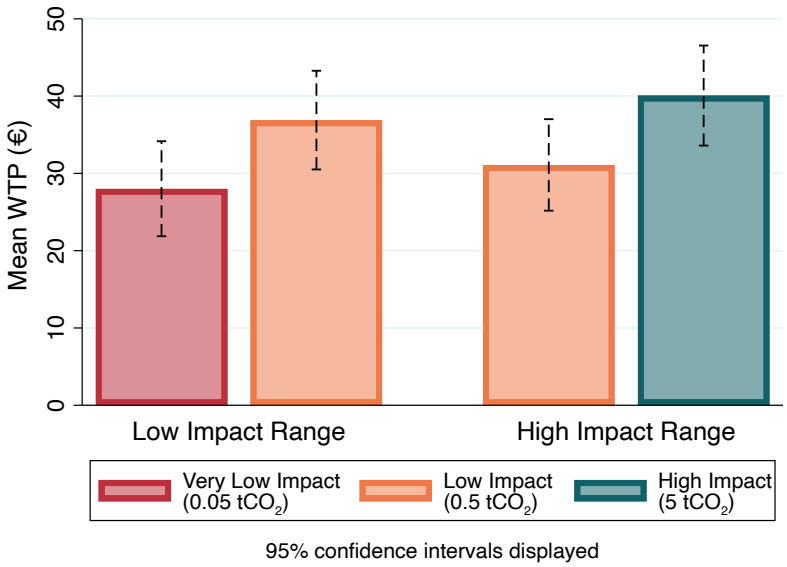


Figure IV.9. The Effect of Joint Evaluation on Private Investors' Response to Impact. This figure shows the results of our experiment investigating how increased comparability affects the response of private investors' WTP to the impact of investments. The graph illustrates the mean WTP for the sustainable investments, by the investment's impact and treatment group.

Table IV.6
Investors' level of sensitivity to the impact of investments

This table provides an overview of investors' level of sensitivity to impact in the joint evaluation experiment. We define the variable *sensitivity* as each investor's WTP for Fund C (higher impact) divided by the WTP for Fund B (lower impact). The values reported in the table pool observations from the HighImpactRange and the LowImpactRange treatments. We divide investors into four groups with respect to their level of sensitivity. Investors with "inverse sensitivity" pay less for the investment with a higher impact. "Insensitive" investors have the same WTP for both investments. The remaining investors pay more for the investment with higher impact and are split into two equally large groups ("low sensitivity" and "high sensitivity"). For each group, we report the number of observations, the mean WTP for Funds B and C individually and pooled, and the mean sensitivity.

	N	**Frequency**	**Mean WTP (€)**			**Mean Sensitivity**
			Fund B	Fund C	Average	
Inverse sensitivity	9	5%	40.3	22.7	31.5	0.5
Insensitive	67	34%	34.5	34.5	34.5	1.0
Low sensitivity	61	31%	41.6	51.3	46.5	1.3
High sensitivity	60	30%	10.1	32.2	21.2	6.0

We define the variable *sensitivity* as investors' WTP for Fund C (higher impact) divided by the WTP for Fund B (lower impact). Table IV.6 shows that about one-third of investors are entirely insensitive to investments' impact (labeled "insensitive"). Even with a side-by-side comparison, these investors' WTP is equal for both sustainable investments, at on average €34.5. We split those investors that are willing to pay more for more impact into two groups at their median sensitivity (labeled "low sensitivity" and "high sensitivity"). Investors in the "low sensitivity" group are, on average, willing to pay 1.3 times more for the more impactful investment. Their average WTP for the sustainable investments is higher than that of the other investors (€46.5, $p < 0.001$, Mann–Whitney U test). In the "high sensitivity" group, investors' WTP is, on average, 6 times higher for an investment with 10 times more impact. Yet their average WTP for both sustainable investments is significantly lower than that of the other sensitivity groups (€21.2, $p = 0.005$, Mann–Whitney U test). Investors in the "high sensitivity" group have a higher level of education and earn more than the other investors ($p = 0.033$ and $p = 0.004$, Mann–Whitney U test). These findings suggest that the modest overall increase in sensitivity in the joint evaluation setup is driven by a sub-group of investors who are really sensitive to impact.

Taken together, the joint evaluation demonstrates that comparability creates some sensitivity to impact. However, even with options to compare, the average investor's WTP for sustainable investments is still far from proportional to the impact of these

investments. First, within the two treatments a clearly visible tenfold increase in impact leads to a WTP that is approximately 30 percent higher. Second, the choice set strongly influences investors' WTP per unit of impact. This indicates that investors evaluate the impact of investments relative to other available options. Further, the results show that while comparability increases investor's sensitivity to impact, it also tends to diminish their baseline WTP for sustainable investments.

IV.4.3 Is Investors' Valuation of Impact Driven by Emotion Rather than by Calculation?

A third potential explanation for investors' insensitivity to impact is that their valuation is driven by the emotional experience of choosing the sustainable option rather than by a calculative appraisal of the impact of this choice. This idea is in accordance with models of "warm glow," where individuals' utility is unrelated to the level of the public good, being instead related to an emotional response that comes from the pro-social act itself (Andreoni, 1989, 1990). Further, Hsee and Rottenstreich (2004) argue that when individuals value a good's characteristic based on emotional perception rather than calculative appraisal, their WTP tends to be a step function of the characteristic. They show, for example, that the willingness to donate money to save pandas depends on the emotional importance of pandas in general, not on the number of pandas that will be saved. Applying this to our context, the emotional response to choosing a sustainable investment could explain the WTP for this sustainable investment.

Our post-experiment survey shows that choosing a sustainable investment feels good to investors (Table IV.7). We ask investors how good it feels to invest in the sustainable investment compared to in the non-sustainable one. Both private investors and impact investors report that it feels better to invest in the sustainable investment ($p < 0.001$, Mann–Whitney U test). The impact investors report a higher level of positive emotions as a consequence of choosing the sustainable investment than do the private investors ($p = 0.005$, Mann–Whitney U test). However, for both private investors and impact investors the investment's impact does not affect these positive emotions. There is no significant difference in the reported positive emotions between HighImpact and LowImpact for both samples. This result also holds for the replication of our main experiment with a larger MTurk sample. If positive emotions drive investors' WTP for sustainable

investments, it would explain why there is no difference in the WTP between the treatments.

Table IV.7
Positive emotions associated with investing sustainably

This table presents the self-stated emotions investors experience when choosing the sustainable investment, per treatment, and sample. Positive emotions are denoted on a scale of -10 to 10, where positive values indicate that it feels better for investors to choose the sustainable investment, and negative values that it feels better to choose the conventional investment. The first two columns report mean values of the self-stated positive emotions, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments. *Private investors* is the sample from the main experiment, *impact investors* is the sample of impact investors, and *MTurks* is the sample from the "R1: MTurk Baseline" replication.

	N	**Mean Positive Emotions [-10,10]**		**Mann–Whitney U Test**
		LowImpact	HighImpact	(HighImpact = LowImpact)
Private investors	196	6.1	6.5	$p = 0.121$
Impact investors	118	7.8	6.9	$p = 0.209$
MTurks	1000	6.5	6.5	$p = 0.710$

Further, we show that investors' valuation of a unit of impact increases with positive emotions, but not with their cost estimate for a unit of impact. In Table IV.8, we regress investors' WTP per ton of CO_2 saved on the level of positive emotions investors experience and their estimate of the cost of saving a ton of CO_2. We find that for both private investors and impact investors, the WTP per ton of CO_2 saved correlates significantly with the level of reported positive emotions, but not with investors' estimate of the cost of saving CO_2. For our larger sample of MTurks in the replication of our main experiment, there is a significant correlation between participants' WTP per ton of CO_2 saved and their estimate of the cost of saving CO_2. However, the effect is relatively weak: a €1 increase in the CO_2 cost estimate corresponds to less than a cent increase in the WTP per ton of CO_2 saved. The significance does not persist if we control for demographics. Consistent with our previous findings, for all samples the impact treatment has a highly significant effect on participants' valuation of a unit of impact. The regression results suggest that investors' valuation of impact is strongly influenced by emotions. A calculative appraisal, where investors determine their WTP by estimating what might be an appropriate price for the impact that is offered to them, does not seem to play an important role.

An emotional valuation mode seems a reasonable explanation not only for the result of our main experiment but also for the results of the extensions. Regarding the exten-

Table IV.8
Emotions, cost estimates, and investors' valuation of impact

This table presents the results of an ordinary least squares (OLS) regression with WTP for the sustainable investment per unit of impact as the dependent variable. In all specifications, investors' level of self-stated positive emotions experienced when choosing the sustainable investment as well as their estimate of the cost of saving a ton of CO_2 are included as independent variables, together with an indicator variable taking the value of 0 for the LowImpact treatment and 1 for the HighImpact treatment. In addition, specifications (2) and (4) also include controls for investors' demographics, as described in detail in Table IV.A.4. Specifications (1) and (2) report the results for our sample of private investors in the main experiment; specifications (3) and (4) report the results for our sample of impact investors; specifications (5) and (6) report the results for our sample of MTurks in the "R1: MTurk Baseline" replication of the main experiment. Standard errors are shown in parentheses.

	Private Investors		**Impact Investors**		**MTurks**	
	(1)	(2)	(3)	(4)	(5)	(6)
	WTP/tCO_2	WTP/tCO_2	WTP/tCO_2	WTP/tCO_2	WTP/tCO_2	WTP/tCO_2
Positive emotions	4.341***	4.314***	3.408**	2.668*	4.826***	4.680***
	(0.968)	(0.957)	(1.229)	(1.326)	(0.384)	(0.388)
Estimated cost of saving 1 ton of CO_2	0.0318	0.0165	-0.0117	-0.00976	0.00232*	0.00197
	(0.0235)	(0.0237)	(0.00781)	(0.00831)	(0.00113)	(0.00114)
Impact treatment	-77.76***	-78.21***	-86.57***	-86.07***	-76.17***	-76.14***
	(7.132)	(6.998)	(8.919)	(9.295)	(2.919)	(2.916)
Demographics	No	Yes	No	Yes	No	Yes
Constant	56.20***	-63.67	74.91***	15.94	52.51***	83.37***
	(8.124)	(44.96)	(11.80)	(46.45)	(3.349)	(17.38)
Observations	195	195	117	117	1000	1000
R^2	0.416	0.486	0.492	0.524	0.464	0.474
F	45.40	10.53	36.54	9.534	287.0	74.12

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

sion with dedicated impact investors, experience can be expected to lead to an improved calculative valuation, as subjects have a better frame of reference via which to price impact. However, when investors' WTP depends mainly on their level of positive emotions, greater precision in estimating costs may be irrelevant. Regarding the role of comparability, "relative" emotional valuation could explain why in the joint evaluation extension investors respond to differences in impact but not to the absolute level of impact. In a direct comparison, it is obvious which option offers greater impact; thus, in a calculative valuation mode WTP for this better option is likely to be greater. This, however, may also be true in an emotional valuation mode. As argued by Hsee and Rottenstreich (2004), given a choice between saving two or saving three pandas it is easy to see that the emotional return will be lower when saving only two. Along the same lines, Ferguson

and Flynn (2016) propose a model of relative warm glow in which the warm glow individuals derive from choosing an option depends on how "good" this option is relative to other options in a given choice set. Such a relative emotional valuation is consistent with both findings of our comparative settings. Investors' WTP responds to impact within the choice set, but it remains constant when the choice set as a whole shifts to a different level of impact. When WTP is determined by emotions, these two observations are easily reconciled.

The level of positive emotions investors experience when selecting sustainable investments may depend on the choice environment. It appears that those choice settings that emphasize a more calculative valuation of impact reduce the positive emotions associated with choosing the sustainable option. Both our joint evaluation extension and the MTurk replication in which the investments differ in their risk and return ("R2: Randomized Risk & Return") likely increase the focus on calculative valuation, and we observe lower average levels of WTP in these experiments. While we do not measure the positive emotions associated with selecting a sustainable choice in the joint evaluation extension, participants in the replication "R2: Randomized Risk & Return" report a significantly lower level of such emotions. Along similar lines, Karlan and Wood (2017) find that small-scale donors are less likely to give again if provided with quantitative information about a charity's effectiveness, presumably because it reduces the warm glow they derive from donating. Overall, the best explanation for our results throughout the present paper remains that investors' WTP for sustainable investments is primarily driven by emotional experience or warm glow.

IV.5 Implications

Based on our findings, we argue that the average sustainable investor does not behave like a consequentialist, but rather like a warm-glow optimizer. Investors explicitly give up financial wealth to invest sustainably, but this WTP is driven by positive emotions regarding making an impact rather than by a calculative valuation of impact.

A key implication of this conclusion is that there may be incentives for greenwashing, or impact washing,[22] in the market for sustainable investment products. With an estimated size of USD 35 trillion assets under management (GSIA, 2021), the market for sustainable investment products is already substantial, and it is set to grow further. If we accept the premise that sustainable investments may help to reduce negative externalities in the real economy, this rapid growth is promising. Our results, however, illustrate that investors pay for positive emotions regarding impact rather than for impact itself. Assuming impact is costly,[23] this creates an incentive for financial institutions to create products that optimize investors' emotions rather than impact.

The current market for sustainable investment products already reflects this focus on emotions. Most sustainable investment products cater to investors' desire to make an impact with their investing, at least in an implicit way. But it is often a vague promise, and few products underpin it with evidence. A review of 230 European investment funds with a focus on sustainability found that 52 percent made explicit environmental impact claims in their marketing material, but that none of the claims were substantiated.[24] Our study suggests that one key reason for this might be that investors are not demanding such evidence, because their preference is to make a choice that is emotionally rewarding. In turn, fund managers are not competing on impact and have little incentive to provide quantitative evidence of it. This is in line with the observation made by Hartzmark and Sussman (2019), who find that market-wide flows strongly increase when funds are clearly labeled as sustainable, while a fund's performance in more quantitative sustainability indicators has a negligible effect on fund flows. A further indication comes from Gibson, Glossner, Krueger, Matos, and Steffen (2021), who find that the holdings of institutional investors that are committed to sustainability are only marginally more sustainable, especially in terms of environmental sustainability. These findings are consistent with a market for sustainable investing that is driven by a demand for positive feelings rather than a calculative valuation of impact.

[22] Busch, Bruce-Clark, Derwall, Eccles, Hebb, Hoepner, Klein, Krueger, Paetzold, Scholtens, and Weber (2021) define impact washing as "the dilution of the term impact investing using the term impact as a marketing tool to attract capital or boost reputations without actually focusing on material solutions to environmental and societal challenges."

[23] Several economic models imply that pro-social investors need to accept lower financial performance in order to have impact (Heinkel, Kraus, and Zechner, 2001; Oehmke and Opp, 2019; Pástor, Stambaugh, and Taylor, 2021). Also, sustainable investments have additional data and expertise requirements that are likely to add to product costs and fees.

[24] https://2degrees-investing.org/wp-content/uploads/2020/03/Report-EU-Retail-Funds-Environmental-Impact-Claims-Do-Not-Comply-with-Regulatory-Guidance.pdf

Certainly, catering to investors' emotions is a business opportunity. Laudi, Smeets, and Weitzel (2021) provide experimental evidence that professional financial advisors actively exploit the sustainability preferences of their clients. In real settings, products may be marketed strategically in such a way that a sustainable product with relatively modest impact is presented as the sustainable option. This would enable financial institutions to collect investors' WTP for sustainable investments, while avoiding the costs of developing investment products with substantial positive impact.[25] This creates the risk of a market for sustainable investment products that benefits investors in terms of warm glow, and financial institutions in terms of fees, but falls well short of fulfilling its potential for solving important societal problems such as curbing carbon emissions to combat climate change.

Our findings also challenge an important assumption in the literature on asset pricing and sustainable finance. Several theoretical papers present models that explore pro-social investors' impact on asset prices (Broccardo, Hart, and Zingales, 2020; Fama and French, 2007; Oehmke and Opp, 2019). These models suggest that pro-social investors shift asset prices, and reward firms that reduce negative externalities with a lower cost of capital. An important assumption in these models, embedded in the utility function of investors, is that pro-social investors optimize between the cost and the impact of sustainable investment opportunities.

We find that such an assumption may be problematic. We suggest that rather than optimizing for impact, pro-social investors optimize for warm glow—and that, in turn, warm glow depends on the way in which investment options are presented and marketed to them. As a consequence, the predictions of these models may be overly optimistic regarding the effect of pro-social preferences on asset prices and externalities. Rather than rewarding firms that effectively reduce negative externalities, investors with pro-social preferences may be rewarding financial intermediaries that design and market products that offer warm glow. Exploring the consequences of warm-glow optimization for models that link pro-social preferences to asset prices is therefore an important avenue for future research.

The notion of financial intermediaries designing products primarily to trigger a warm glow might seem in contradiction with the findings of Berg, Heeb, and Kölbel (2022),

[25] This finding has parallels to evidence from the market for retail structured products, showing that financial institutions' margins increase with the complexity of products (Célérier and Vallée, 2017; Chesney, Fattinger, Krakow, and Straumann, 2022).

who find that funds with a dedicated ESG strategy actually adjust their holdings based on changes in ESG ratings. If investors' demand for ESG funds is primarily driven by warm glow, why should financial intermediaries make an effort to adjust funds' holdings after such changes? We think that, for two reasons, the findings of Berg, Heeb, and Kölbel (2022) are compatible with the results of the present study. First, as shown by Berg, Heeb, and Kölbel (2022), the impact of funds that consider ESG performance in their allocation decisions seems to be limited, and—as demonstrated by Friede, Busch, and Bassen (2015)—hardly, or at least so far, seems to come at the cost of reduced financial performance. Second, funds that market themselves as being sustainable are increasingly under regulatory pressure to act upon their ESG claims.[26] Thus, as such, adjusting holdings based on ESG ratings may constitute the minimal effort required by regulators if financial intermediaries want to market products in ways that trigger a positive emotional response in investors.

Finally, our results have implications for the metrics that are used in sustainable investing. Problems of measurement are pervasive in sustainable finance. Claims that an investment is "sustainable" or the fact that it has one additional Morningstar "globe" are difficult to interpret, even for experts. Confronted with this challenge, investors may appear insensitive to quantity simply because it is unclear to them what quantity means. Our study shows that scope insensitivity persists when impact is presented as a relatively well-defined quantity. Tons of CO_2 emissions are an established metric that investors can understand. And yet investors remain insensitive to quantity when confronted with CO_2 emissions. This suggests that improved sustainability accounting, while desirable in many ways, will not help to realign WTP with impact, because valuation is being driven by emotions.

One potential solution to this problem is revealed by a comparison between our results and those of Bonnefon, Landier, Sastry, and Thesmar (2022) and Brodback, Günster, and Pouget (2021). In their experiments, it appears that investors consistently value positive externalities if these are expressed in monetary terms. Of course, monetizing environmental and social impacts is difficult and the typical ESG product is a long way from providing such a figure. Efforts to provide monetized impact measures are, however,

[26] This is reflected both in new regulatory frameworks that aim to avoid greenwashing in the financial sector, including the EU's Sustainable Finance Disclosure Regulation, and in recent litigation based on greenwashing allegations against large financial intermediaries, including Goldman Sachs, DWS, and BNY Mellon, as documented by *The Economist* in its "Special report ESG investing" (July 23, 2022), available at https://www.economist.com/special-report/2022-07-23.

underway.[27] Equipped with monetized impact measurements, investors may indeed one day behave as consequentialists who adjust their WTP to the level of impact that a product offers them.

Another way to realign emotional preferences with quantitative product characteristics is labels for sustainability funds. In Europe, a number of labeling initiatives to this end have already been established, and our results underscore the importance of such efforts.[28] In the US, meanwhile, there are currently no established labels for sustainable investment products. There are, of course, also risks inherent in labeling. Our findings demonstrate that investors have a substantial WTP even for investments with a rather modest impact if these are perceived as more impactful than the alternative. It is thus important to ensure that labels actually align with socially desirable outcomes. So, for example, a product receiving a high score in a climate rating should respect limiting global warming to 1.5° Celsius above preindustrial levels—the international goal agreed at the Paris climate summit in 2015.

IV.6 Conclusion

We present evidence that investors' willingness to pay for sustainable investments is largely independent of the impact of such investments. We arrive at this result for both experienced private investors and dedicated high-net-worth impact investors. The results replicate with a larger MTurk sample and for variations of the elicitation method. Being able to exclude a series of alternative explanations, we suggest that pro-social investors are best understood as warm-glow optimizers who prefer investments that feel good rather than as consequentialists who derive utility from optimizing their impact.

Our findings have important implications for modeling investors' pro-social preferences in asset pricing and for policy makers who want to harness the growing demand for sustainable investments in order to support efforts to achieve societal goals. Current theoretical models routinely assume pro-social investors to be consequentialists. Incorporating the importance of warm glow for decision-making might affect these models' conclusions. For policy makers, our findings indicate a risk of greenwashing by providers

[27] See, e.g., https://www.hbs.edu/impact-weighted-accounts/Pages/default.aspx.

[28] There are eight different sustainability labels provided by NGOs for funds in Europe; there is the Morningstar Sustainability Rating, described in Hartzmark and Sussman (2019); and there is an ongoing effort by the European Union to develop an "ecolabel" for sustainable investment products.

of financial products and a potential equilibrium of "light green" products. This is particularly relevant against the backdrop of achieving internationally agreed sustainability goals, and given the fact that we observe our results even for experienced and dedicated impact investors.

IV.A Appendix

IV.A.1 Detailed Procedure: WTP Elicitation and Incentivization

WTP Elicitation

We elicit investors' WTP for the sustainable investment through a series of seven binary choices between Fund A and Fund B. In the first choice, both funds feature a onetime, upfront fee of €10, which will be deducted from the €1,000 investment. This first choice reveals which investment option investors prefer if fees are equal. By way of explanation, assume an investor initially chooses Fund A, and that this is the sustainable option.[29] In the next step, we add €40 to the fee for Fund A, so that the investor now has the choice between Fund A with a fee of €50, and Fund B with a fee of €10. If an investor switches to Fund B under these conditions, we deduct €20 from the fee for Fund A for the following choice. In such a case, the investor now faces a choice between Fund A with a fee of €30, and Fund B with a fee of €10. If, rather, an investor still prefers Fund A, we add another €20 to the fee for Fund A for the following choice, so that the investor has the choice between Fund A with a fee of €70, and Fund B with a fee of €10. We repeat this procedure over four additional choices, dividing the amount we deduct from or add to the fee for Fund A by a factor of two in each consecutive round. During each choice, all information on the two investments is visible to the investors.

After these seven choices, we calculate the midpoint between the highest fee the investor is willing to accept for Fund A and the lowest fee for Fund A at which the investor decided to switch to Fund B. We calculate the investor's WTP for Fund A as this midpoint minus the baseline fee of €10. For investors that prefer Fund B in the initial choice, we proceed accordingly by varying the fees for Fund B. If Fund B is the investment without impact, we elicit the WTP for this investment in the same way as described above, but we use the negative of this value as the investor's WTP for the sustainable investment. This procedure allows us to measure WTP values ranging from €-78.125 to €+78.125, with a precision of €1.25.

[29] Note that we randomize whether Fund A or Fund B corresponds to the sustainable investment, to avoid ordering effects. We checked for order effects, but did not find any (see Table IV.A.1).

Incentivization

After we have elicited each participant's stated WTP for his or her preferred investment, we draw a random amount between the highest and lowest WTP that we can detect with our design—that is, between €-78.125 and €+78.125. If this random amount is smaller than the elicited WTP, we invest €1,000 in the investor's chosen investment, deducting the randomly drawn amount. If the random amount is larger than or equal to the elicited WTP, we invest €1,000 in the other investment, as is standard in the BDM mechanism. This is to ensure that investors have no incentive to deviate from their true preferences (Becker, Degroot, and Marschak, 1964).

Impact Realization

We realize the impact component of the selected investment by purchasing the corresponding amount of carbon credits from a verified greenhouse gas emissions reduction project. Such projects save greenhouse gas emissions, which are measured in tons of CO_2 equivalents, according to standardized methodologies (e.g., as defined by the Verified Carbon Standard or the Gold Standard). All emissions savings are verified by an independent third party. As it has been questioned in various cases whether investments in emissions reduction projects lead to the claimed emissions savings (Alexeew, Bergset, Meyer, Petersen, Schneider, and Unger, 2010), we implement our emissions savings with a project that reduces methane emissions from organic waste treatment in Vietnam.[30] In a report commissioned by the European Commission, Cames, Harthan, Füssler, Lazarus, Lee, Erickson, and Spalding-Fecher (2016) conclude that, as opposed to other project types, methane reduction projects are highly likely to lead to the claimed emissions reductions.

IV.A.2 Detailed Procedure: MTurk Replications

The MTurk replications largely follow the original experimental procedure we have used in our main experiment. All deviations from the original procedure are reported for each replication in the corresponding subsection below. To ensure comparability with our main results, all replications are incentivized in the same way: we randomly select ten

[30] For more detailed information on the project, see https://market.southpole.com/home/offset-emissions/project-details/71.

participants from among all MTurk participants, and make a real $1,000 investment for each of them, following the procedure described under IV.A.1.

R1: MTurk Baseline

In this experiment, we replicate our main experiment with a sample of MTurks. We follow the original experimental design, adjusting the currency from euros to U.S. dollars. We recruit MTurk participants over the Approved Participants Panel offered by Cloud Research. MTurks in this panel have delivered evidence of a high level of engagement and attention. In addition, we restrict our study to MTurks living in the U.S. with a substantial track record of completing tasks well (at least 1,000 tasks completed on the platform and an acceptance rate for these tasks of at least 96%). As a further measure to ensure data quality, we add an attention check early in the experiment. As in the initial experiment, we exclude participants who do not agree that climate change is a serious problem that needs to be solved (stating an agreement level of 3 or less on a scale of 1–7). We also exclude participants who explicitly disagree with the elicited WTP and are unwilling to repeat the elicitation. In addition, we apply the following exclusion criteria: we exclude participants who fail the attention check; we exclude participants who "fail" more than once in at least one of the two quizzes; we exclude participants who take less than four minutes to complete the experiment. The same recruitment and quality control procedures are applied for all MTurk replications described below. For this experiment, we collect observations until we have exactly 1,000 observations that pass the abovementioned screening criteria. The pre-registration of this replication can be found at https://aspredicted.org/M34_P2R.

R2: Randomized Risk & Return

In this replication, we introduce two randomized treatment conditions regarding the past risk and return of the sustainable investment to the experimental design used in our baseline MTurk experiment (R1). In both risk/return treatment groups, the conventional investment features an "average" Morningstar Risk rating and an annualized past return of 6%, as in the original experimental design. In the first treatment condition regarding risk and return, the sustainable investment features a "below-average" Morningstar Risk rating and an annualized past return of 5%. The second fund features an "above-average" Morningstar Risk rating and an annualized past return of 7%. In addition to the original

design, we provide the information that a lower Morningstar Risk rating corresponds to a more favorable risk profile to avoid confusion about the directionality of this measure. Also, to find real mutual funds matching these characteristics, we broaden the fund category from "US Large-Cap Blend Equity" to "US Large-Cap Equity." We collect observations until we have precisely 400 observations that pass the screening criteria described above under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/B2P_PTF.

R3: Direct Ask (No Anchor)

In this replication, instead of applying the discrete choice method of the original design we directly ask investors how much more they are willing to pay for the sustainable investment. Specifically, we ask the following question: "Please consider a $1,000 investment in one of the two funds presented above. How much more would you be willing to pay for Fund A (B) compared to Fund B (A), in terms of upfront fees? Please enter an amount in $," where Fund A (B) corresponds to the sustainable investment. Participants receive the identical fund information as in the baseline MTurk experiment (R1). We collect observations until we have precisely 200 observations that pass the above screening criteria under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/JXQ_ZGD.

R4: Direct Ask (Low Anchor)

In this replication, we follow the experimental setup described under "R3: Direct Ask (No Anchor)," but provide information on common fees for equity funds. Specifically, we display the following statement directly before the question on investors' WTP: "A standard passive US Large Cap Equity fund such as Fund A (B) charges a fee of about 0.1% per year. This is $1 for a $1,000 investment", where Fund A (B) corresponds to the conventional investment. We collect observations until we have exactly 200 observations that pass the screening criteria described above under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/JXQ_ZGD.

R5: Direct Ask (High Anchor)

In this replication, we follow the experimental setup described under "R3: Direct Ask (No Anchor)," but provide information on common fees for equity funds. Specifically, we display the following statement directly before the question on investors' WTP: "A standard active US Large Cap Equity fund such as Fund A (B) charges a fee of about 1% per year. This is $10 for a $1,000 investment", where Fund A (B) corresponds to the conventional investment. We collect observations until we have exactly 200 observations that pass the screening criteria described above under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/JXQ_ZGD.

R6: No Neutral Choice

In this replication, we follow the experimental setup described under "R1: MTurk Baseline." We omit, however, the first investment choice, where participants decide between the two funds with upfront fees being equal for both funds. In the first choice, the sustainable fund already costs $40 more in terms of upfront fees than the conventional one. We collect observations until we have exactly 200 observations that pass the screening criteria described above under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/GZL_LCN.

R7: Upward Elicitation

In this replication, instead of applying the midpoint method of our main experiment, we use a discrete choice method that gradually increases the fee differences between the sustainable investment and the conventional one. As in the main experiment, participants first face a choice between the two investments where for each investment an upfront fee of $10 is deducted. In the next choice, we increase the upfront fee for the investment preferred in the first choice by $1.25. If a participant still prefers the same investment with this level of fees, we increase the upfront fee of this investment so that the difference in fees between the two investments doubles. So, in the third choice, the fee for the preferred investment would be $2.50 higher than that for the other investment. As long as a participant keeps selecting the initially preferred investment, we keep doubling the fee differences. Once a participant switches to the other investment in any of their choices, we start applying the bisection method used in the original setup until we reach

a precision of $1.25 for the WTP elicitation. If the difference in fees exceeds $80, we directly ask participants for the maximum difference in fees they would be willing to pay for the preferred fund, analogous to the procedure of the main experiment. We collect observations until we have exactly 200 observations that pass the screening criteria described above under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/4KB_V7F.

R8: Lower Scale

In this replication, we apply the midpoint method used in our main experiment, but use a lower measurement scale. After the first choice, where participants face a choice between the two investments with an upfront fee of $10 for each, we increase the fee of the preferred investment by $20. This is in contrast to the main experiment, where we increase the fee for the preferred investment by $40 in the second choice. As in the original procedure, investors make seven investment choices. For participants that do not deviate from the initially preferred fund in all subsequent choices, we directly ask for the maximum difference in fees they would be willing to pay for the preferred fund, as we do in the main experiment. This setup increases our measurement precision for investors' WTP by a factor of two and reduces the maximal range of our discrete choices by a factor of two. We collect observations until we have exactly 200 observations that pass the screening criteria described above under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/GZL_LCN.

R9: Higher Scale

In this replication, we apply the midpoint method used in our main experiment, but use an increased measurement scale. After the first choice, where participants face a choice between the two investments with an upfront fee of $10 for each, we increase the fee of the preferred investment by $60. This is in contrast to the main experiment, where we increase the fee for the preferred investment by $40 in the second choice. As in the original procedure, investors make seven investment choices. For investors that do not deviate from the initially preferred fund in all subsequent choices, we directly ask for the maximum difference in fees they would be willing to pay for the preferred fund, as we do in the main experiment. This setup decreases our measurement precision for investors' WTP by a factor of two but also increases the maximal range of our discrete choices by

a factor of two. We collect observations until we have exactly 200 observations that pass the screening criteria described above under "R1: MTurk Baseline." The pre-registration of this replication can be found at https://aspredicted.org/PNQ_L29.

Table IV.A.1
Order effects

This table presents the results of an ordinary least squares (OLS) regression with WTP for the sustainable investment as the dependent variable. In both specifications, an indicator variable taking the value of 0 for participants seeing the sustainable investment on the left-hand side and 1 for participants seeing the sustainable investment on the right-hand side is included, together with an indicator variable taking the value of 0 for the LowImpact treatment and 1 for the HighImpact treatment. Specification (1) reports the results for our sample of private investors in the main experiment; specification (2) reports the results for the baseline experiment with MTurks. Standard errors are shown in parentheses ("R1: MTurk Baseline").

	Private Investors		**MTurks**	
	(1)	(2)	(3)	(4)
	WTP	WTP	WTP	WTP
Right-hand design	-3.892	-9.362	0.0539	-0.146
	(5.680)	(8.090)	(2.135)	(2.969)
Impact treatment	6.838	1.514	1.952	1.748
	(5.680)	(7.981)	(2.136)	(2.996)
Right-hand design Impact treatment		10.79		0.415
		(11.36)		(4.275)
Constant	44.14***	46.45***	42.95***	43.04***
	(4.661)	(5.260)	(1.779)	(2.019)
Observations	196	196	1000	1000
R^2	0.009	0.013	0.001	0.001
F	0.859	0.873	0.420	0.283

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

Table IV.A.2
Main results, excluding investors who "fail" the quizzes

This table presents private investors' absolute and relative WTP for the sustainable fund, as elicited in the main experiment, but excluding investors who twice "fail" in at least one of the two quizzes that test their comprehension of the information that is provided to them. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments.

	Mean Values		**Mann–Whitney U Test**
	LowImpact ($n = 91$)	HighImpact ($n = 94$)	(HighImpact = LowImpact)
WTP (€)	42.66	48.77	$p = 0.383$
WTP/Impact (€/tCO_2)	85.32	9.75	$p = 0.001$

Table IV.A.3
Main results excluding investors with censored WTP

This table presents private investors' absolute and relative WTP for the sustainable investment, as elicited in the main experiment, excluding investors with censored WTP. Investors' WTP is censored if they do not deviate from their initially preferred investment in all seven investment choices. We cannot elicit WTP directly from the choices of these investors. We additionally ask these investors to state their WTP for the sustainable investment. These stated values are excluded from the results presented in this table. For each treatment group, the mean WTP is lower when excluding investors with censored WTP. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments.

	Mean Values		**Mann–Whitney U Test**
	LowImpact ($n = 81$)	HighImpact ($n = 74$)	(HighImpact = LowImpact)
WTP (€)	30.24	28.22	$p = 0.903$
WTP/Impact (€/tCO_2)	60.48	5.64	$p = 0.001$

Table IV.A.4
Definition of variables elicited in the post-experiment survey

This table shows the definitions and measurement of the variables we elicit in our post-experiment survey.

Variable	Measure
Risk expectations	Answer to the question "How do you expect that Fund A and Fund B compare in terms of risk?" on a 5-point scale from "Investing in Fund A is much riskier" to "Investing in Fund B is much riskier." Values are normalized to a scale from -10 to 10, where positive values indicate that investors expect the sustainable investment to be less risky, and negative ones that they expect the conventional investment to be less risky.
Return expectations	Answer to the question "How do you expect that Fund A and Fund B compare in terms of return?" on a 5-point scale from "Fund A will deliver a much higher return" to "Fund B will deliver a much higher return." Values are normalized to a scale from -10 to 10, where positive values indicate that investors expect the sustainable investment to deliver higher returns, and negative ones that they expect the conventional investment to deliver higher returns.
Positive emotions	Answer to the question "How do Fund A and Fund B compare in terms of how it feels to invest in the fund?" on a 5-point scale from "It feels much better to invest in Fund A" to "It feels much better to invest in Fund B." Values are normalized to a scale from -10 to 10, where positive values indicate that it feels better for investors to choose the sustainable investment, and negative values that it feels better to choose the conventional investment.
Perceived investment impact	Agreement with the statement "Investing in Fund [A,B] makes a relevant contribution to fighting climate change," on a 7-point Likert scale. Values are normalized to a scale from -10 to 10.
General relevance impact	Agreement with the statement "When investing, it is important to me whether I contribute to fighting climate change," on a 7-point Likert scale. Values are normalized to a scale from -10 to 10.
General relevance impact level	Agreement with the statement "When investing, it is important to me *how much* I contribute to fighting climate change," on a 7-point Likert scale, where [A,B] corresponds to the sustainable investment. Values are normalized to a scale from -10 to 10.
Estimated cost of saving 1 ton of CO_2	Answer to the question "What do you think: What are the average costs of saving 1,000 kg of CO_2 emissions (in €)?" To reduce the influence of extreme values, values are winsorized at the 5% and 95% levels.
Risk preferences	Answer to the question "In general, how willing or unwilling are you to take risks?" on a 10-point scale (1 = "Completely unwilling to take risks"; 10 = "Very willing to take risks"), according to the experimentally validated survey module of Falk, Becker, Dohmen, Huffman, and Sunde (2016).
Time preferences	Answer to the question "How willing are you to give up something that is beneficial for you today in order to benefit more from that thing in the future?" on a 10-point scale (1 = "Completely unwilling"; 10 = "Very willing to do so"), following Falk, Becker, Dohmen, Huffman, and Sunde (2016).
Altruism	Answer to the question "How do you assess your willingness to share with others without expecting anything in return?" on a 10-point scale (1 = "Completely unwilling to share"; 10 = "Very willing to share"), following Falk, Becker, Dohmen, Huffman, and Sunde (2016).
Climate awareness	Agreement with the statement "Climate change is a serious problem that needs to be solved," on a 7-point Likert scale. Values are normalized to a scale from -10 to 10.
Female	The dummy variable *Female* takes the value of 1 if the investor chooses *Female* from among the options *Female*, *Male*, and *Other*, and zero if not.
Age	Investor's self-stated age.
Income	Self-reported annual household income, with options ranging from "less than €10,000" to "€200,000 or more," in steps of €5,000.
Net worth	Self-reported household net worth, with seven options ranging from "less than €50,000" to "more than €10 million."
Highest education	Self-reported highest degree or level of schooling the investor has completed.
Investment knowledge	Agreement with the statement "Compared to the average of the population, my investment knowledge is good," on a 7-point Likert scale (1 = "Strongly disagree"; 7 = "Strongly agree"), following Dorn and Huberman (2005) and Riedl and Smeets (2017). Values are normalized to a scale from 0 to 10.

Table IV.A.5
Risk and return expectations and investors' WTP for sustainability

This table presents the results of an ordinary least squares (OLS) regression with the WTP for the sustainable investment as the dependent variable. In all specifications, investors' risk and return expectations for the sustainable investment, as elicited in our post-experiment survey, are included as independent variables. Both risk and return expectations are transformed to a scale from -10 to 10, where positive values indicate that investors have a more favorable view of the sustainable investment, negative ones that they have a more favorable view of the conventional investment. All specifications include an indicator variable taking the value of 0 for the LowImpact treatment and 1 for the HighImpact treatment. In addition, specifications (2) and (4) also include controls for investors' demographics, as described in detail in Table IV.A.4. Specifications (1) and (2) report the results for our sample of private investors in the main experiment; specifications (3) and (4) report the results for our sample of impact investors. Standard errors are shown in parentheses.

	Private Investors		**Impact Investors**	
	(1)	(2)	(3)	(4)
	WTP	WTP	WTP	WTP
Risk expectations	0.252	0.0689	-1.386	-2.258
	(0.891)	(0.873)	(1.288)	(1.288)
Return expectations	0.851	0.745	-0.712	-0.126
	(0.710)	(0.695)	(1.116)	(1.188)
Impact treatment	7.450	6.739	1.204	1.970
	(5.658)	(5.567)	(6.020)	(6.107)
Demographics	No	Yes	No	Yes
Constant	41.94***	9.232	49.44***	-34.23
	(4.060)	(31.38)	(4.335)	(31.98)
Observations	194	194	118	118
R^2	0.017	0.120	0.020	0.121
F	1.078	2.058	0.779	1.204

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

Table IV.A.6
WTP before and after the onset of the COVID-19 crisis

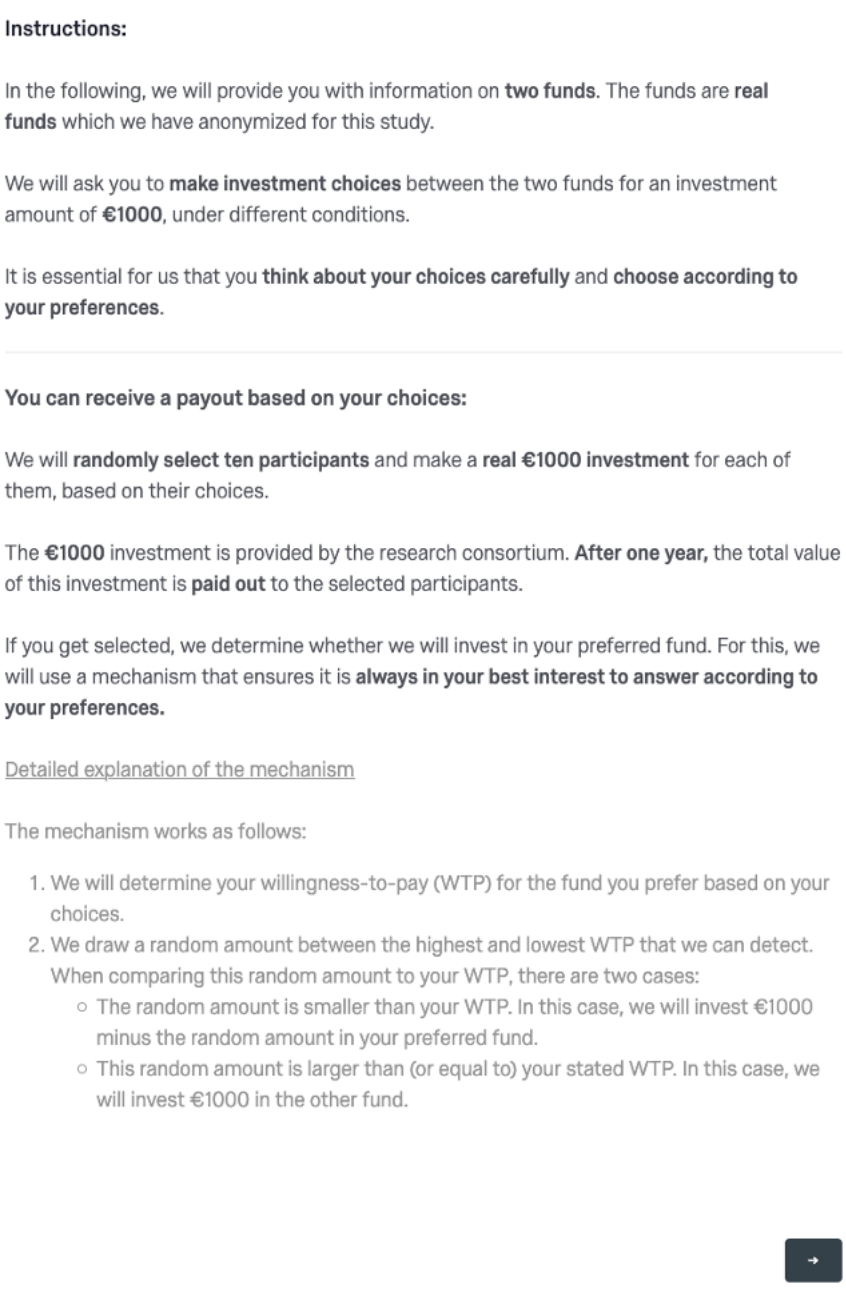
This table reports the results of a preliminary version of our experiment, which we ran with students at Radboud University in September 2019, well before the emergence of SARS-CoV-2, as well as in September 2020. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments.

	Mean Values		**Mann–Whitney U Test**
	LowImpact	HighImpact	(HighImpact = LowImpact)
September 2019			
N	159	152	
WTP (€)	27.64	29.82	$p = 0.533$
WTP/Impact (€/tCO_2)	55.28	5.96	$p < 0.001$
September 2020			
N	119	123	
WTP (€)	32.03	27.85	$p = 0.262$
WTP/Impact (€/tCO_2)	64.04	5.57	$p < 0.001$

Table IV.A.7
Results for the impact investors during the pre-registration period

This table presents investors' absolute and relative WTP for the sustainable investment, as elicited in our experiment with the sample of impact investors, strictly following the pre-registered procedure and thereby excluding five investors who participated in the experiment starting more than three months after its launch. The first two columns report mean values of the variables, by impact treatment; the third column reports p-values of a Mann–Whitney U test, testing for differences between the two treatments.

	Mean Values		**Mann–Whitney U Test**
	LowImpact ($n = 56$)	HighImpact ($n = 57$)	(HighImpact = LowImpact)
WTP (€)	48.79	49.45	$p = 0.850$
WTP/Impact (€/tCO_2)	97.59	9.89	$p < 0.001$



Instructions:

In the following, we will provide you with information on **two funds**. The funds are **real funds** which we have anonymized for this study.

We will ask you to **make investment choices** between the two funds for an investment amount of **€1000**, under different conditions.

It is essential for us that you **think about your choices carefully** and **choose according to your preferences**.

You can receive a payout based on your choices:

We will **randomly select ten participants** and make a **real €1000 investment** for each of them, based on their choices.

The **€1000** investment is provided by the research consortium. **After one year,** the total value of this investment is **paid out** to the selected participants.

If you get selected, we determine whether we will invest in your preferred fund. For this, we will use a mechanism that ensures it is **always in your best interest to answer according to your preferences.**

Detailed explanation of the mechanism

The mechanism works as follows:

1. We will determine your willingness-to-pay (WTP) for the fund you prefer based on your choices.
2. We draw a random amount between the highest and lowest WTP that we can detect. When comparing this random amount to your WTP, there are two cases:
   - The random amount is smaller than your WTP. In this case, we will invest €1000 minus the random amount in your preferred fund.
   - This random amount is larger than (or equal to) your stated WTP. In this case, we will invest €1000 in the other fund.

→

Figure IV.A.1. Instructions. This figure shows a screenshot of the instructions that investors receive.

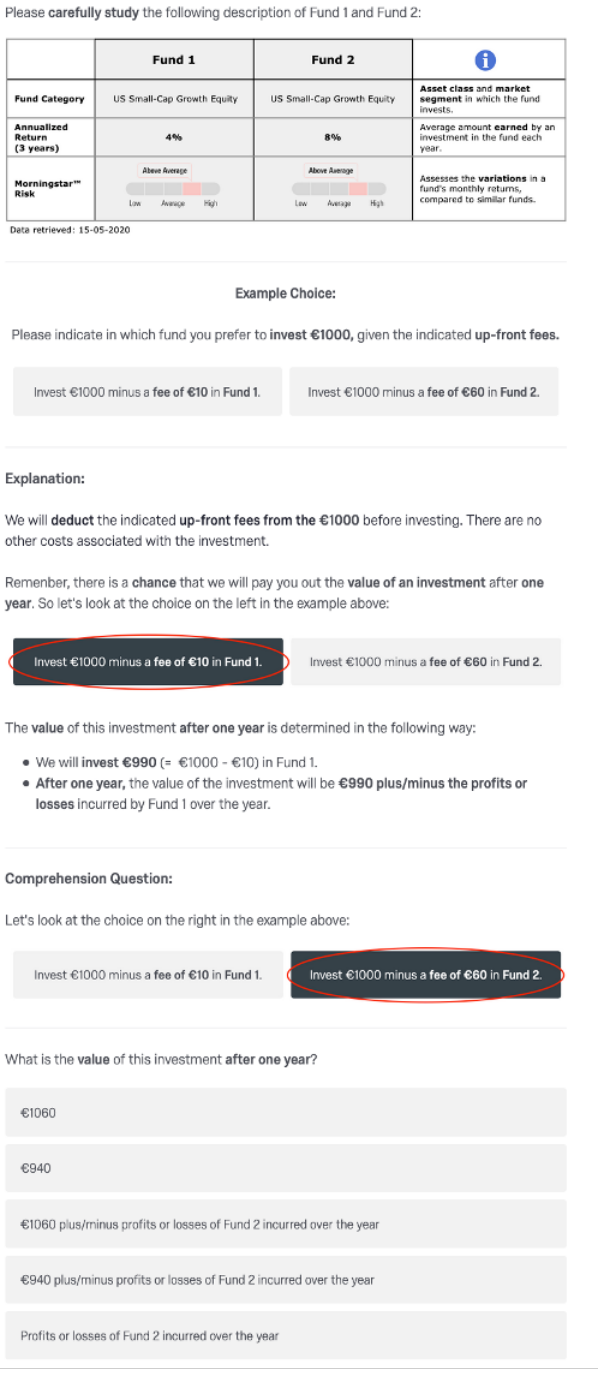
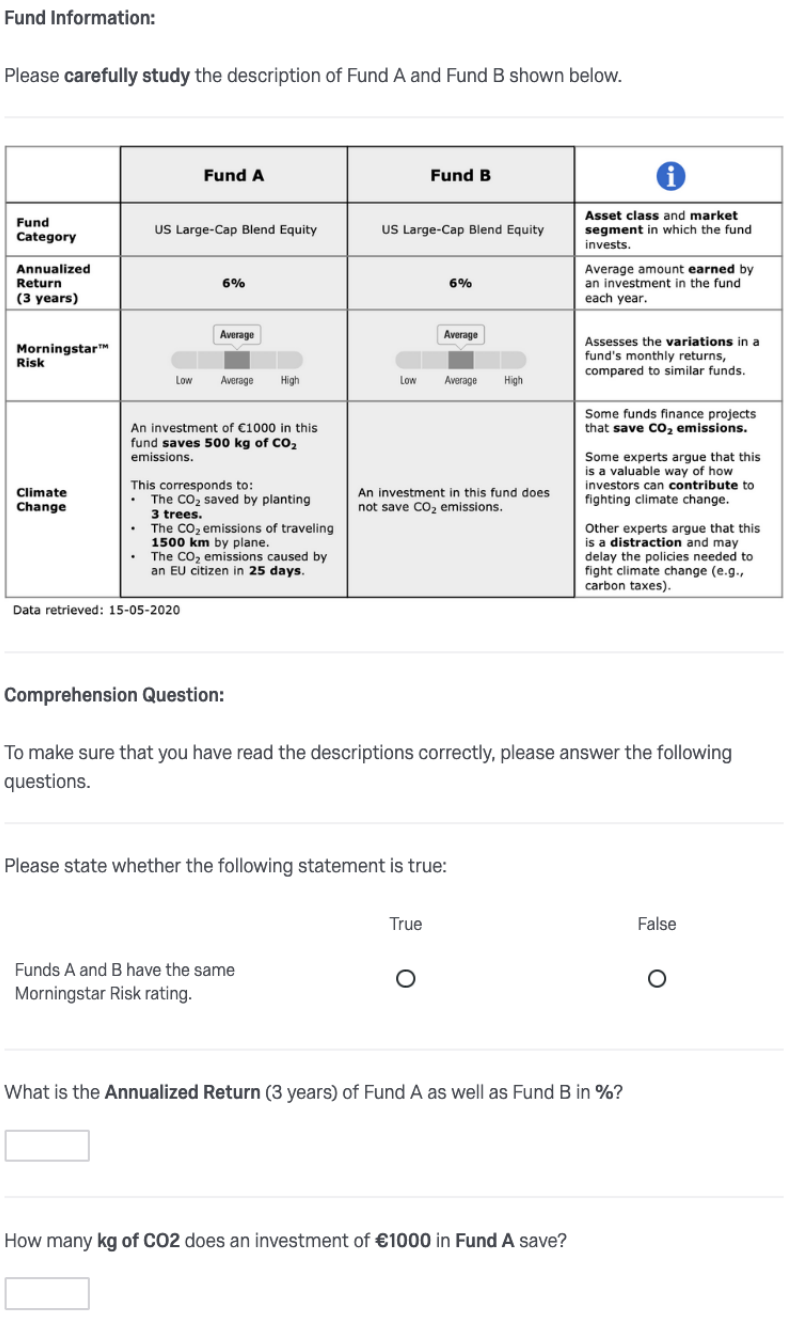


Figure IV.A.2. Example Choice. This figure shows a screenshot of the quiz investors participate in before taking the investment decisions.



Fund Information:

Please **carefully study** the description of Fund A and Fund B shown below.

	Fund A	**Fund B**	ℹ
Fund Category	US Large-Cap Blend Equity	US Large-Cap Blend Equity	**Asset class** and **market segment** in which the fund invests.
Annualized Return (3 years)	**6%**	**6%**	Average amount **earned** by an investment in the fund each year.
Morningstar™ Risk	Average — Low Average High	Average — Low Average High	Assesses the **variations** in a fund's monthly returns, compared to similar funds.
Climate Change	An investment of €1000 in this fund **saves 500 kg of CO_2** emissions. This corresponds to: • The CO_2 saved by planting **3 trees.** • The CO_2 emissions of traveling **1500 km** by plane. • The CO_2 emissions caused by an EU citizen in **25 days**.	An investment in this fund does not save CO_2 emissions.	Some funds finance projects that **save CO_2 emissions.** Some experts argue that this is a valuable way of how investors can **contribute** to fighting climate change. Other experts argue that this is a **distraction** and may delay the policies needed to fight climate change (e.g., carbon taxes).

Data retrieved: 15-05-2020

Comprehension Question:

To make sure that you have read the descriptions correctly, please answer the following questions.

Please state whether the following statement is true:

	True	False
Funds A and B have the same Morningstar Risk rating.	○	○

What is the **Annualized Return** (3 years) of Fund A as well as Fund B in **%**?

How many **kg of CO2** does an investment of **€1000** in **Fund A** save?

Figure IV.A.3. Investment Information and Comprehension Quiz. This figure provides a screenshot of the information the investors participating in our main experiment receive on the two investments if they are assigned to the LowImpact treatment, as well as of the comprehension quiz investors have to "pass" if they are to continue.

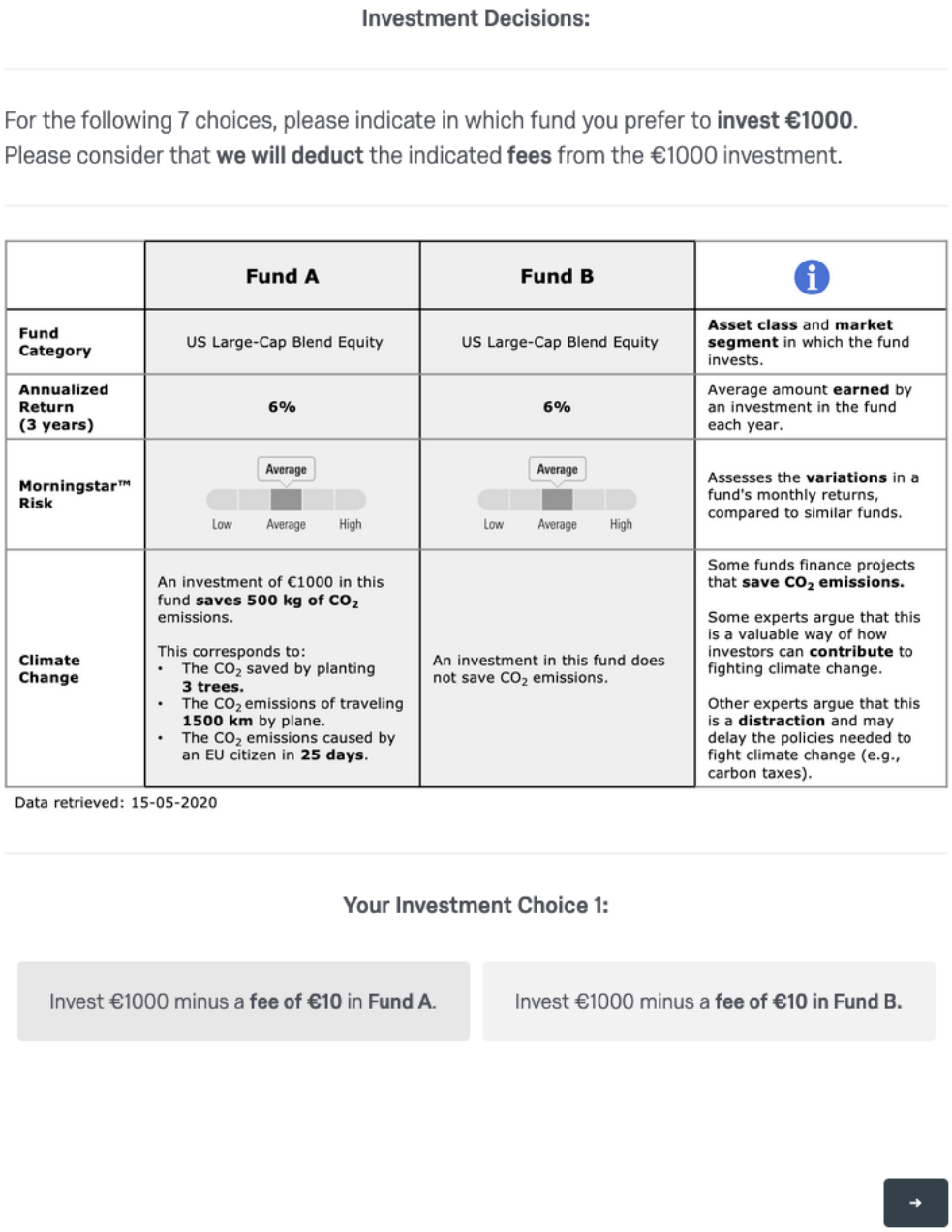
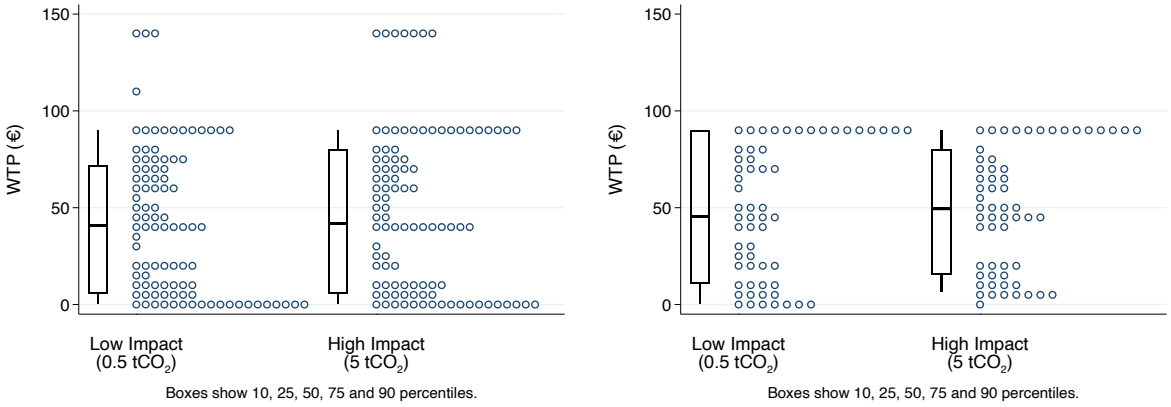


Figure IV.A.4. Investment Choice. The screenshot in this figure shows an example of the investment choices the investors face.

	Fund A	Fund B	Fund C	
Fund Category	US Large-Cap Blend Equity	US Large-Cap Blend Equity	US Large-Cap Blend Equity	**Asset class** and **market segment** in which the fund invests.
Annualized Return (3 years)	**6%**	**6%**	**6%**	Average amount **earned** by an investment in the fund each year.
Morningstar™ Risk	Average Low Average High	Average Low Average High	Average Low Average High	Assesses the **variations** in a fund's monthly returns, compared to similar funds.
Climate Change	An investment into Fund A does not save CO_2 emissions.	An investment of €1000 in this fund **saves 50 kg of CO_2** emissions. This corresponds to: • **30%** of the CO_2 saved by planting a **tree.** • The CO_2 emissions of traveling **150 km** by plane. • The CO_2 emissions caused by an EU citizen in **2.5 days**.	An investment of €1000 in this fund **saves 500 kg of CO_2** emissions. This corresponds to: • The CO_2 saved by planting **3 trees.** • The CO_2 emissions of traveling **1500 km** by plane. • The CO_2 emissions caused by an EU citizen in **25 days**.	Some funds finance projects that **save CO_2 emissions.** Some experts argue that this is a valuable way of how investors can **contribute** to fighting climate change. Other experts argue that this is a **distraction** and may delay the policies needed to fight climate change (e.g., carbon taxes).

Data retrieved: 15-05-2020

Figure IV.A.5. Screenshot of the Investment Information in the Joint Evaluation Extension. This figure provides an example of the information the investors receive in the joint evaluation extension of our experiment on the three investment options. The screenshot corresponds to the investment information investors in the LowImpactRange treatment receive.



Panel A: Private Investors

Panel B: Impact Investors

Figure IV.A.6. Distribution of Investors' WTP for Sustainable Investments. This figure shows vertical histograms of investors' WTP for the sustainable investment. Panel A shows the WTP in the main experiment with private investors, Panel B shows the WTP of impact investors. Each dot reflects one individual WTP, and dots are grouped and stacked to the right in brackets with a width of €5. Note that the data has been winsorized at the 5 percent and 95 percent levels, according to the pre-registration.

Bibliography

Abdellaoui, Mohammed, 2000, Parameter-free elicitation of utility and probability weighting functions, *Management Science* 46, 1497–1512.

Admati, Anat R., and Paul Pfleiderer, 2009, The "wall street walk" and shareholder activism: Exit as a form of voice, *The Review of Financial Studies* 22, 2645–2685.

Aguinis, Herman, Isabel Villamor, and Ravi S. Ramani, 2021, MTurk research: Review and recommendations, *Journal of Management* 47, 823–837.

Ahern, Kenneth R., 2014, Do common stocks have perfect substitutes? Product market competition and the elasticity of demand for stocks, *Review of Economics and Statistics* 96, 756–766.

Alexeew, Johannes, Linda Bergset, Kristin Meyer, Juliane Petersen, Lambert Schneider, and Charlotte Unger, 2010, An analysis of the relationship between the additionality of CDM projects and their contribution to sustainable development, *International Environmental Agreements: Politics, Law and Economics* 10, 233–248.

Amel-Zadeh, Amir, and George Serafeim, 2018, Why and how investors use ESG information: Evidence from a global survey, *Financial Analysts Journal* 74(3), 87–103.

Andreoni, James, 1989, Giving with impure altruism: Applications to charity and ricardian equivalence, *Journal of Political Economy* 97, 1447–1458.

Andreoni, James, 1990, Impure altruism and donations to public goods: A theory of warm-glow giving, *The Economic Journal* 100, 464–477.

Andreoni, James, and John Miller, 2002, Giving according to GARP: An experimental test of the consistency of preferences for altruism, *Econometrica* 70, 737–753.

Ansar, Atif, B. L. Caldecott, and James Tilbury, 2013, Stranded assets and the fossil fuel divestment campaign: What does divestment mean for the valuation of fossil fuel assets?, Smith School of Enterprise and the Environment, University of Oxford.

Arechar, Antonio A., Simon Gächter, and Lucas Molleman, 2018, Conducting interactive experiments online, *Experimental Economics* 21, 99–131.

Bagwell, Laurie S., 1992, Dutch auction repurchases: An analysis of shareholder heterogeneity, *The Journal of Finance* 47, 71–105.

Baker, Malcolm, Daniel Bergstresser, George Serafeim, and Jeffrey Wurgler, 2018, Financing the response to climate change: The pricing and ownership of us green bonds, NBER Working Paper No. 25194.

Baker, Malcolm, Jeremy C. Stein, and Jeffrey Wurgler, 2003, When does the market matter? Stock prices and the investment of equity-dependent firms, *The Quarterly Journal of Economics* 118, 969–1005.

Baker, Malcolm, and Jeffrey Wurgler, 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives* 21, 129–152.

Bamberger, Michael, and Linda Mabry, 2019, *RealWorld Evaluation: Working Under Budget, Time, Data, and Political Constraints* (Sage Publications).

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

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barko, Tamas, Martijn Cremers, and Luc Renneboog, 2017, Activism on corporate social responsibility, *SSRN Electronic Journal*.

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

Beck, Thorsten, Aslı Demirgüç-Kunt, Luc Laeven, and Vojislav Maksimovic, 2006, The determinants of financing obstacles, *Journal of International Money and Finance* 25, 932–952.

Beck, Thorsten, Aslı Demirgüç-Kunt, and Maksimovic Vojislav, 2008, Financing patterns around the world: Are small firms different?, *Journal of Financial Economics* 89, 467–487.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak, 1964, Measuring utility by a single-response sequential method, *Behavioral Science* 9, 226–232.

Becker, Gary S., 1974, A theory of social interactions, *Journal of Political Economy* 82, 1063–1093.

Beltratti, Andrea, 2005, Capital market equilibrium with externalities, production and heterogeneous agents, *Journal of Banking & Finance* 29, 3061–3073.

Benabou, Roland, and Jean Tirole, 2010, Individual and corporate social responsibility, *Economica* 77, 1–19.

Beneish, Messod D., and Robert E. Whaley, 1996, An anatomy of the "S&P game": The effects of changing the rules, *The Journal of Finance* 51, 1909–1930.

Berg, Florian, Florian Heeb, and Julian F. Kölbel, 2022, The economic impact of ESG rating changes, *SSRN Electronic Journal*.

Berg, Florian, Julian F. Kölbel, and Roberto Rigobon, 2022, Aggregate Confusion: The Divergence of ESG Ratings, *Review of Finance* (in press).

Berger, Allen N., and Gregory F. Udell, 1998, The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle, *Journal of Banking & Finance* 22, 613–673.

Berk, Jonathan, and Jules H. van Binsbergen, 2021, The impact of impact investing, *SSRN Electronic Journal*.

Bessembinder, Hendrik, Michael J. Cooper, and Feng Zhang, 2019, Characteristic-based benchmark returns and corporate events, *The Review of Financial Studies* 32, 75–125.

Betti, Gianni, Costanza Consolandi, and Robert G. Eccles, 2018, The relationship between investor materiality and the sustainable development goals: A methodological framework, *Sustainability* 10, 2248.

Blanchard, Olivier, Changyong Rhee, and Lawrence Summers, 1993, The stock market, profit, and investment, *The Quarterly Journal of Economics* 108, 115–136.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts, 2010, Why do firms in developing countries have low productivity?, *American Economic Review* 100, 619–23.

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

Bolton, Patrick, and Marcin Kacperczyk, 2021b, Global pricing of carbon-transition risk, NBER Working Paper No. 28510.

Bonnefon, Jean-Francois, Augustin Landier, Parinitha Sastry, and David Thesmar, 2022, The moral preferences of investors: Experimental evidence, *SSRN Electronic Journal*.

Brest, Paul, and Kelly Born, 2013, Unpacking the impact in impact investing, *Stanford Social Innovation Review* 11, 22–31.

Brest, Paul, Ronald J. Gilson, and Mark A. Wolfson, 2018, Essay: How investors can (and can't) create social value, *The Journal of Corporation Law* 44, 205.

Briere, Marie, and Stefano Ramelli, 2021, Green sentiment, stock returns, and corporate behavior, *SSRN Electronic Journal*.

Broadhurst, Dominic, Janette Watson, and Jane Marshall, 2003, *Ethical and Socially Responsible Investment: A Reference Guide for Researchers* (Walter de Gruyter).

Broccardo, Eleonora, Oliver Hart, and Luigi Zingales, 2020, Exit vs. voice, NBER Working Paper No. 27710.

Brodback, Daniel, Nadja Günster, and Sébastien Pouget, 2021, The valuation of corporate social responsibility: A willingness-to-pay experiment, Working Paper.

Busch, Timo, Peter Bruce-Clark, Jeroen Derwall, Robert Eccles, Tessa Hebb, Andreas Hoepner, Christian Klein, Philipp Krueger, Falko Paetzold, Bert Scholtens, and Olaf Weber, 2021, Impact investments: A call for (re)orientation, *SN Business & Economics* 1, 33.

Busch, Timo, Matthew Johnson, and Thomas Pioch, 2022, Corporate carbon performance data: Quo vadis?, *Journal of Industrial Ecology* 26, 350–363.

Cames, Martin, Ralph O. Harthan, Jürg Füssler, Michael Lazarus, Carrie M. Lee, Pete Erickson, and Randall Spalding-Fecher, 2016, How additional is the clean development mechanism?, available at: https://www.oeko.de/publikationen/p-details/how-additional-is-the-clean-development-mechanism.

Carlos, W. Chad, and Ben W. Lewis, 2018, Strategic silence: Withholding certification status as a hypocrisy avoidance tactic, *Administrative Science Quarterly* 63, 130–169.

Carpenter, Robert E., and Bruce C. Petersen, 2002, Is the growth of small firms constrained by internal finance?, *Review of Economics and statistics* 84, 298–309.

Ceccarelli, Marco, Simon Glossner, Mikael Homanen, and Daniel Schmidt, 2021, Which institutional investors drive corporate sustainability?, *SSRN Electronic Journal*.

Célérier, Claire, and Boris Vallée, 2017, Catering to investors through security design: Headline rate and complexity, *The Quarterly Journal of Economics* 132, 1469–1508.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression discontinuity and the price effects of stock market indexing, *The Review of Financial Studies* 28, 212–246.

Chatterji, Aaron K., Rodolphe Durand, David I. Levine, and Samuel Touboul, 2016, Do ratings of firms converge? Implications for managers, investors and strategy researchers, *Strategic Management Journal* 37, 1597–1614.

Chatterji, Aaron K., and Michael W. Toffel, 2010, How firms respond to being rated, *Strategic Management Journal* 31, 917–945.

Chava, Sudheer, 2014, Environmental externalities and cost of capital, *Management science* 60, 2223–2247.

Chesney, Marc, Felix Fattinger, Nils J. Krakow, and Simon Straumann, 2022, Interest rates, competition, and complexity: Demand and supply of retail financial products, *SSRN Electronic Journal*.

Chesney, Marc, Jonathan Gheyssens, Anca C. Pana, and Luca Taschini, 2016, *Environmental Finance and Investments* (Springer).

Chowdhry, Bhagwan, Shaun William Davies, and Brian Waters, 2019, Investing for impact, *The Review of Financial Studies* 32, 864–904.

Clarke, Damian, and Kathya Tapia-Schythe, 2021, Implementing the panel event study, *The Stata Journal* 21, 853–884.

Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *The Journal of Finance* 54, 2045–2073.

Cravo, Tulio, and Caio Piza, 2016, The impact of business support services for small and medium enterprises on firm performance in low- and middle-income countries: A meta-analysis, World Bank Policy Research Working Paper No. 7664.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier, 2012, Testing for altruism and social pressure in charitable giving, *The Quarterly Journal of Economics* 127, 1–56.

Desvousges, William H., F. Reed Johnson, Richard W. Dunford, Kevin J. Boyle, Sara P. Hudson, and K. Nicole Wilson, 1992, *Measuring Nonuse Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy* (RTI Press).

Diller, Christian, and Christoph Kaserer, 2009, What drives private equity returns? Fund inflows, skilled GPs, and/or risk?, *European Financial Management* 15, 643–675.

Dimson, Elroy, Oğuzhan Karakaş, and Xi Li, 2015, Active ownership, *The Review of Financial Studies* 28, 3225–3268.

Dimson, Elroy, Oğuzhan Karakaş, and Xi Li, 2021, Coordinated engagements, European Corporate Governance Institute Finance Working Paper No. 721.

Dorn, Daniel, and Gur Huberman, 2005, Talk and action: What individual investors say and what they do, *Review of Finance* 9, 437–481.

Dyck, Alexander, Karl V. Lins, Lukas Roth, and Hannes F. Wagner, 2019, Do institutional investors drive corporate social responsibility? International evidence, *Journal of Financial Economics* 131, 693–714.

Eccles, Robert G., and Judith C. Stroehle, 2018, Exploring social origins in the construction of ESG measures, *SSRN Electronic Journal*.

Eckel, Catherine C., and Philip J. Grossman, 1996, Altruism in anonymous dictator games, *Games and Economic Behavior* 16, 181–191.

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2012, The real effects of financial markets: The impact of prices on takeovers, *The Journal of Finance* 67, 933–971.

Egli, Florian, Bjarne Steffen, and Tobias S. Schmidt, 2018, A dynamic analysis of financing conditions for renewable energy technologies, *Nature Energy* 3, 1084–1092.

Evans, David S., and Boyan Jovanovic, 1989, An estimated model of entrepreneurial choice under liquidity constraints, *Journal of Political Economy* 97, 808–827.

Falk, Armin, Anke Becker, Thomas J. Dohmen, David Huffman, and Uwe Sunde, 2016, The preference survey module: A validated instrument for measuring risk, time, and social preferences, *SSRN Electronic Journal*.

Fama, Eugene F., and Kenneth R. French, 2007, Disagreement, tastes, and asset prices, *Journal of Financial Economics* 83, 667–689.

Ferguson, Eamonn, and Niall Flynn, 2016, Moral relativism as a disconnect between behavioural and experienced warm glow, *Journal of Economic Psychology* 56, 163–175.

Finucane, Melissa L., Ali Alhakami, Paul Slovic, and Stephen M. Johnson, 2000, The affect heuristic in judgments of risks and benefits, *Journal of Behavioral Decision Making* 13, 1–17.

Fischer, Stanley, and Robert C. Merton, 1984, Macroeconomics and finance: The role of the stock market, in *Carnegie-Rochester Conference Series on Public Policy* vol. 21 pp. 57–108. Elsevier.

Frey, Bruno S., and Stephan Meier, 2004, Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment, *The American Economic Review* 94, 1717–1722.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, and Jesse M. Shapiro, 2021, Visualization, identification, and estimation in the linear panel event-study design, NBER Working Paper No. 29170.

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

Gantchev, Nickolay, Mariassunta Giannetti, and Rachel Li, 2022, Does money talk? Divestitures and corporate environmental and social policies, *Review of Finance* (forthcoming).

Geddes, Anna, Tobias S. Schmidt, and Bjarne Steffen, 2018, The multiple roles of state investment banks in low-carbon energy finance: An analysis of Australia, the UK and Germany, *Energy Policy* 115, 158–170.

Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch, 2016, *Impact Evaluation in Practice* (World Bank Publications).

Gibson, Rajna, Simon Glossner, Philipp Krueger, Pedro Matos, and Tom Steffen, 2021, Do responsible investors invest responsibly?, *SSRN Electronic Journal*.

Gibson Brandon, Rajna, Philipp Krueger, and Peter S. Schmidt, 2021, ESG rating disagreement and stock returns, *Financial Analysts Journal* 77, 104–127.

Gilbert, Daniel Todd, Susan T. Fiske, and Gardner Lindzey, 1998, *The Handbook of Social Psychology* (Oxford University Press).

Glück, Maximilian, Benjamin Hübel, and Hendrik Scholz, 2021, ESG rating events and stock market reactions, *SSRN Electronic Journal*.

Gollier, Christian, and Sébastien Pouget, 2014, The "washing machine": Investment strategies and corporate behavior with socially responsible investors, Working Paper.

Gompers, Paul, and Josh Lerner, 2000, Money chasing deals? The impact of fund inflows on private equity valuations, *Journal of Financial Economics* 55, 281–325.

Goranova, Maria, and Lori Verstegen Ryan, 2014, Shareholder activism: A multidisciplinary review, *Journal of Management* 40, 1230–1268.

Greenwood, Robin, Toomas Laarits, and Jeffrey Wurgler, 2022, Stock market stimulus, NBER Working Paper No. 29827.

Greiner, Sandra, and Axel Michaelowa, 2003, Defining investment additionality for CDM projects—practical approaches, *Energy Policy* 31, 1007–1015.

GSIA, 2018, Global sustainable investment review 2018, available at: http://www.gsi-alliance.org/trends-report-2018.

GSIA, 2021, Global sustainable investment review 2020, available at: http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf.

Hachenberg, Britta, and Dirk Schiereck, 2018, Are green bonds priced differently from conventional bonds?, *Journal of Asset Management* 19, 371–383.

Hadlock, Charles J., and Joshua R. Pierce, 2010, New evidence on measuring financial constraints: Moving beyond the KZ index, *The Review of Financial Studies* 23, 1909–1940.

Haidt, Jonathan, 2007, The new synthesis in moral psychology, *Science* 316, 998–1002.

Harrison, Glenn W., and John A. List, 2004, Field experiments, *Journal of Economic Literature* 42, 1009–1055.

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

Heeb, Florian, Julian F. Kölbel, Falko Paetzold, and Stefan Zeisberger, 2022, Do investors care about impact?, *The Review of Financial Studies* (forthcoming).

Heinkel, Robert, Alan Kraus, and Josef Zechner, 2001, The effect of green investment on corporate behavior, *Journal of Financial and Quantitative Analysis* 36, 431–449.

Hertin, Julia, Frans Berkhout, Marcus Wagner, and Daniel Tyteca, 2008, Are EMS environmentally effective? The link between environmental management systems and environmental performance in European companies, *Journal of Environmental Planning and Management* 51, 259–283.

Hirschman, Albert O., 1970, *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States* (Harvard University Press).

Hoepner, Andreas G. F., Ioannis Oikonomou, Zacharias Sautner, Laura T Starks, and Xiaoyan Zhou, 2018, ESG shareholder engagement and downside risk, *SSRN Electronic Journal*.

Holt, Charles A., and Susan K. Laury, 2014, Assessment and estimation of risk preferences, in Mark Machina, and Kip Viscusi, ed.: *Handbook of the Economics of Risk and Uncertainty* (Elsevier).

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen, 1994, Sticking it out: Entrepreneurial survival and liquidity constraints, *Journal of Political Economy* 102, 53–75.

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

Hong, Harrison, Neng Wang, and Jinqiang Yang, 2021, Welfare consequences of sustainable finance, NBER Working Paper No. 28595.

Hsee, Christopher K., and Yuval Rottenstreich, 2004, Music, pandas, and muggers: On the affective psychology of value, *Journal of Experimental Psychology: General* 133, 23–30.

Hsee, Christopher K., and Jiao Zhang, 2010, General evaluability theory, *Perspectives on Psychological Science* 5, 343–355.

Huberman, Gur, 2001, Familiarity breeds investment, *The Review of Financial Studies* 14, 659–680.

Humphrey, Jacquelyn, Shimon Kogan, Jacob Sagi, and Laura T. Starks, 2020, The asymmetry in responsible investing preferences, *SSRN Electronic Journal*.

IFC, 2019, Operating principles for impact management, International Finance Corporation.

IPCC, 2018, *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (Cambridge University Press).

Ivković, Zoran, and Scott Weisbenner, 2009, Individual investor mutual fund flows, *Journal of Financial Economics* 92, 223–237.

Kahneman, Daniel, and Jack L. Knetsch, 1992, Valuing public goods: The purchase of moral satisfaction, *Journal of Environmental Economics and Management* 22, 57–70.

Karlan, Dean, and Daniel H. Wood, 2017, The effect of effectiveness: Donor response to aid effectiveness in a direct mail fundraising experiment, *Journal of Behavioral and Experimental Economics* 66, 1–8.

Kaul, Aditya, Vikas Mehrotra, and Randall Morck, 2000, Demand curves for stocks do slope down: New evidence from an index weights adjustment, *The Journal of Finance* 55, 893–912.

Kersten, Renate, Job Harms, Kellie Liket, and Karen Maas, 2017, Small firms, large impact? A systematic review of the SME finance literature, *World Development* 97, 330–348.

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

Kim, Soohun, and Aaron Yoon, 2020, Analyzing active managers' commitment to ESG: Evidence from the United Nations Principles for Responsible Investment, *SSRN Electronic Journal*.

Knight, Richard, 1990, Sanctions, disinvestment, and US corporations in South Africa, in *Sanctioning Apartheid* (Africa World Press Trenton).

Kölbel, Julian F., Florian Heeb, Falko Paetzold, and Timo Busch, 2020, Can sustainable investing save the world? Reviewing the mechanisms of investor impact, *Organization & Environment* 33, 554–574.

Kuhnen, Camelia M., and Brian Knutson, 2011, The influence of affect on beliefs, preferences, and financial decisions, *Journal of Financial and Quantitative Analysis* 46, 605–626.

Landier, Augustin, and Stefano Lovo, 2020, ESG investing: How to optimize impact?, *SSRN Electronic Journal*.

Laudi, Marten, Paul Smeets, and Utz Weitzel, 2021, Do Financial Advisors Exploit Responsible Investment Preferences?, *SSRN Electronic Journal*.

Lins, Karl V., Henri Servaes, and Ane Tamayo, 2017, Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis, *The Journal of Finance* 72, 1785–1824.

Loderer, Claudio, John W. Cooney, and Leonard D. Van Drunen, 1991, The price elasticity of demand for common stock, *The Journal of Finance* 46, 621–651.

Luo, H. Arthur, and Ronald J. Balvers, 2017, Social screens and systematic investor boycott risk, *Journal of Financial and Quantitative Analysis* 52, 365–399.

Lynch, Anthony W., and Richard R. Mendenhall, 1997, New evidence on stock price effects associated with changes in the S&P 500 index, *The Journal of Business* 70, 351–383.

MacGregor, Donald G., Paul Slovic, David Dreman, and Michael Berry, 2000, Imagery, affect, and financial judgment, *Journal of Psychology and Financial Markets* 1, 104–110.

McCahery, Joseph A., Zacharias Sautner, and Laura T. Starks, 2016, Behind the scenes: The corporate governance preferences of institutional investors, *The Journal of Finance* 71, 2905–2932.

McKinsey Global Institute, 2022, The net-zero transition: What it would cost, what it could bring, available at: https://www.mckinsey.com/business-functions/sustainability/our-insights/the-net-zero-transition-what-it-would-cost-what-it-could-bring.

Melnyk, Steven A., Robert P. Sroufe, and Roger Calantone, 2003, Assessing the impact of environmental management systems on corporate and environmental performance, *Journal of Operations Management* 21, 329–351.

Metzger, Laura, and Isabel Günther, 2019, Making an impact? The relevance of information on aid effectiveness for charitable giving. A laboratory experiment, *Journal of Development Economics* 136, 18–33.

Morck, Randall, Andrei Shleifer, Robert W. Vishny, Matthew Shapiro, and James M. Poterba, 1990, The stock market and investment: Is the market a sideshow?, *Brookings Papers on Economic Activity* 1990, 157–215.

Null, Clair, 2011, Warm glow, information, and inefficient charitable giving, *Journal of Public Economics* 95, 455–465.

OECD, 2020, Global outlook on financing for sustainable development 2021, available at: https://www.oecd.org/development/global-outlook-on-financing-for-sustainable-development-2021-e3c30a9a-en.htm.

Oehmke, Martin, and Marcus M. Opp, 2019, A theory of socially responsible investment, *SSRN Electronic Journal*.

Paetzold, Falko, Timo Busch, and Marc Chesney, 2015, More than money: Exploring the role of investment advisors for sustainable investing, *Annals in Social Responsibility*.

Pedersen, Lasse H., Shaun Fitzgibbons, and Lukasz Pomorski, 2021, Responsible investing: The ESG-efficient frontier, *Journal of Financial Economics* 142, 572–597.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *The Review of Financial Studies* 22, 435–480.

Pindyck, Robert S., 2019, The social cost of carbon revisited, *Journal of Environmental Economics and Management* 94, 140–160.

Pástor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2021, Sustainable investing in equilibrium, *Journal of Financial Economics* 142, 550–571.

Rajan, Raghuram, and Luigi Zingales, 1996, Financial dependence and growth, NBER Working Paper No. 5758.

Ramelli, Stefano, and Alexander F. Wagner, 2020, Feverish stock price reactions to COVID-19, *The Review of Corporate Finance Studies* 9, 622–655.

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

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

Rzeznik, Aleksandra, Kathleen Weiss Hanley, and Loriana Pelizzon, 2022, Investor reliance on ESG ratings and stock price performance, *SSRN Electronic Journal*.

Schmidheiny, Kurt, and Sebastian Siegloch, 2019, On event study designs and distributed-lag models: Equivalence, generalization and practical implications, Working Paper.

Schreiner, Mark, and Jacob Yaron, 2001, *Development Finance Institutions: Measuring Their Subsidy* (World Bank Publications).

Searcy, Cory, and Doaa Elkhawas, 2012, Corporate sustainability ratings: An investigation into how corporations use the Dow Jones Sustainability Index, *Journal of Cleaner Production* 35, 79–92.

Seele, Peter, and Marc Chesney, 2017, Toxic sustainable companies: A critique on the shortcomings of current corporate sustainability ratings and a definition of 'financial toxicity', *Journal of Sustainable Finance & Investment* 7, 139–146.

Shanaev, Savva, and Binam Ghimire, 2022, When ESG meets AAA: The effect of ESG rating changes on stock returns, *Finance Research Letters* 46, 102302.

Shang, Jen, and Rachel Croson, 2009, A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods, *The Economic Journal* 119, 1422–1439.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *The Journal of Finance* 41, 579–590.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

Slovic, Paul, Melissa L. Finucane, Ellen Peters, and Donald G. MacGregor, 2007, The affect heuristic, *European Journal of Operational Research* 177, 1333–1352.

Stern, Nicholas, 2007, *The Economics of Climate Change: The Stern Review* (Cambridge University Press).

Strong, Norman, and Xinzhong Xu, 2003, Understanding the equity home bias: Evidence from survey data, *The Review of Economics and Statistics* 85, 307–312.

Tang, Dragon Yongjun, and Yupu Zhang, 2020, Do shareholders benefit from green bonds?, *Journal of Corporate Finance* 61, 101427.

Teoh, Siew Hong, Ivo Welch, and C. Paul Wazzan, 1999, The effect of socially activist investment policies on the financial markets: Evidence from the South African boycott, *The Journal of Business* 72, 35–89.

UNEP, 2020, The emissions gap report 2020, available at: https://www.unep.org/emissions-gap-report-2020.

van de Kuilen, Gijs, and Peter P. Wakker, 2011, The midweight method to measure attitudes toward risk and ambiguity, *Management Science* 57, 582–598.

Vörösmarty, C. J., V. Rodríguez Osuna, D. A. Koehler, P. Klop, J. D. Spengler, J. J. Buonocore, A. D. Cak, Z. D. Tessler, F. Corsi, P. A. Green, et al., 2018, Scientifically assess impacts of sustainable investments, *Science* 359, 523–525.

Wurgler, Jeffrey, and Ekaterina Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks?, *The Journal of Business* 75, 583–608.

Zerbib, Olivier David, 2019, The effect of pro-environmental preferences on bond prices: Evidence from green bonds, *Journal of Banking & Finance* 98, 39–60.

Zizzo, Daniel John, 2010, Experimenter demand effects in economic experiments, *Experimental Economics* 13, 75–98.

Curriculum Vitae

Florian Heeb

Date of birth: February 19, 1987

Education

- 2017–2022: PhD studies in Finance, University of Zurich
- 2014–2016: St.Gallen Management Certificate (CAS), University of St.Gallen (HSG)
- 2008–2011: MSc in Environmental Science, ETH Zurich
- 2005–2010: BSc in Environmental Science, ETH Zurich

Main Professional Experience

- 08/2017–08/2022: Research Associate, University of Zurich
- 02/2017–06/2017: Chief Operating Officer, South Pole (Zurich)
- 01/2015–01/2017: Director Consultancy & Services, South Pole (Zurich)
- 10/2013–12/2014: Head of Swiss Offset Projects, South Pole (Zurich)
- 11/2011–10/2013: Project Manager, Foundation myclimate (Zurich)