Sustainable Investing, Social Preferences and ESG Commitment

Renfang Zhang

Doctor of Philosophy
The University of Edinburgh
2023

Acknowledgements

I would express my sincere gratitude to my supervisors Dr. Angelica Gonzalez and Dr. Maria Boutchkova for their support during my PhD studies. Your enthusiasm in research is inspiring and motivating. You have always been my source of inspiration, support and guidance. The care and patience you had for me and our fruitful discussions helped me as a student to gain confidence, overcome difficulties in this journey and become a qualified researcher.

I am very grateful to my friends and colleagues with whom I spend a happy time in work and life in Edinburgh in the past four years. Most importantly, my study would not have been possible without the love and unconditional support from my parents, my partner Terry and my two dogs Eva and Lucas. This work is also dedicated to you.

Table of contents

Abstract

This thesis examines sustainable investment and social preferences of investors and fund managers. The first chapter examines individuals' social preferences and the potential drivers behind such preferences. Although abundant evidence exists in psychological and sociological studies that show people demonstrate other-regarding behaviors. It is unclear how and whether this is manifested in the economic life of agents. In the first chapter of this thesis, I examine the response of philanthropic flows in the UK and the US to exogenous shocks that potentially motivate or disincentivize giving. I find humanitarian crises, such as an earthquake taking place far from the population being studied, are unlikely to induce strong social preference changes. Whereas there is some evidence that shocks related to environmental protection may result in more charitable donations. In addition, I show that the preference for "fair play" may also manifest in the form of demanding effectiveness and a good reputation from non-profit organizations as individuals reduce donations to scandalous charities. I also show that the extent to which economic agents act altruistically and make donations could depend on the financial benefits of doing so.

In Chapter 3, I illustrate that social preferences do not drive sustainable investing (SI) by examining the response of SI flows to the exogenous shocks in Chapter 2 that trigger changes in donation flow. The research design of this analysis is grounded in the theory that considers a philanthropy-conventional investment bundle as an alternative to SI, given their common social preference and risk-return motives. While environmental shocks trigger altruism, as evidenced by a significant effect on granular US and UK charity flows, there is no response in SI. Adverse shocks to charities' reputation and tax shield benefits also do not affect SI. In addition, the thesis addresses alternative explanations: risk-return effects and institutional and informational differences between SI and philanthropy. Chapter 3 contributes to the increasing body of literature on SI by presenting new evidence on a commonly-assumed argument that, compared to conventional investors, SI investors make decisions based on a different utility function modified by their social preferences.

Chapter 4 of this thesis examines differences in portfolio choices between sustainable investing (SI) mutual funds and conventional mutual funds. In line with existing literature, I show preliminary evidence that SI and conventional funds share similar financial characteristics, although labeled differently. Additionally, anecdotal evidence suggests SI funds do not underweight firms that incurred well-publicized ESG scandals. To empirically examine the ESG commitment of SI funds, I employ a difference-in-differences approach to compare portfolio holding changes of SI and conventional funds after ESG controversies. I control for financial interests in rebalancing by categorizing controversies into positive, negative, and insignificant cumulative abnormal returns (CAR). I find that both SI and conventional funds hold their portfolios unchanged when there are clear market reactions to controversies (positive and negative CAR). Conventional and SI funds act in opposite directions when the market disagrees on the stock price implications of ESG controversies, with SI being buyers of controversial firms. I suggest this is because SI funds are motivated by the fact that controversies lead to an increase in ESG ratings. The novelty of this work lies firstly in using the percentage of shares (stake) in companies as an objective measure of portfolio-level ESG commitment of SI funds, as opposed to most existing literature that uses noisy and subjective fund-level ESG ratings. The research design also addresses endogeneity in portfolio decision making by SI/Conventional mutual funds by using firm level “ESG compliance” measured by an objective measure - the number of recent ESG controversies. The findings have important implications for the theoretical underpinnings for SI investing. While most theoretical models assume SI has a distinct taste for ESG, I do not find such empirical evidence regarding ESG controversies. The findings in this thesis suggest that SI fund investors and policy makers should not rely on ratings when evaluating SI fund commitment to ESG using portfolio level ESG ratings.

Lay Summary

Over the past decades, sustainable investing (SI) has gained increasing investor interest, and the SI mutual fund industry has expanded substantially. Academics have also devoted extant research to this topic.

Many existing studies suggested that investors of SI mutual funds act differently compared to investors of conventional mutual funds when making their investment decisions because SI investors are driven by non-financial motives. In this thesis, I address whether social preferences drive SI investment. Meanwhile, the “sustainable” label on SI funds is being challenged and doubted as purely a marketing tool. In this thesis, I also examine whether SI mutual funds and conventional mutual funds differ in their portfolio compositions and reactions to ESG-related controversies.

To achieve these goals, I employ a quasi-experiment approach to test if SI funds react to shocks to social preferences. I first identify that certain types of shocks affect social preferences through changes in individuals’ donations to charities. However, investment flows to SI funds do not react to these shocks to social preferences, indicating that non-financial social preferences may not be the primary driver for SI investment. In addition, this thesis finds that SI and conventional funds have substantial overlaps in their portfolios. For example, after portfolio firms experience environmental, social, or governance (ESG) related controversies, SI funds increase holdings in these companies relative to conventional funds, likely because SI funds are more concerned about their fund-level ESG ratings than negative externalities exerted by controversial firms.

My thesis contributes to the literature on sustainable investment by providing two critical pieces of evidence: investors' demand for SI is not primarily driven by social preferences; SI mutual funds are more concerned with how they are perceived than the actual ESG performance of their holdings. As such, work presented in this thesis is essential to investors, businesses, researchers, and policy makers because it has important implications on how one should understand investor preferences and tastes, laying ground for theoretical and empirical analysis of SI, as well as critically evaluate ESG commitment of SI mutual funds.

Chapter 1 Introduction

Sustainable investing (SI), also known as socially responsible investing (SRI), refers to investing incorporating environmental, social and governance (ESG) considerations. Over the past two decades, the interest in SI has grown substantially. The US Sustainable Investing Forum (US SIF) estimates a $16.6 trillion of assets at the beginning of 2020 (US SIF, 2020). The United Nations Principal for Responsible Investment (PRI) records over 3,000 PRI signatories at the end of 2020 (PRI, 2020). With the growth of SI, an expanding body of academic literature now examines SI investors and SI mutual funds theoretically and empirically. In this thesis, I document the empirical behavior of both SI investors (Chapter 3) and mutual funds managers (Chapter 4) when faced with events that reveal their social preferences and/or commitment to sustainability.

In line with Fama and French (2007), several theoretical papers model SI investors as individuals with specific preferences for sustainable assets who also gain utility from aligning their investment portfolios with social values, thus leading to a new sustainable priced factor under a new market equilibrium (e.g., Heinkel et al., 2001; Oehmke & Opp 2020; Pastor et al., 2021; Pedersen et al., 2021; De Angelis et al., 2022). In these models, the market is composed of different types of investors. For example, green investors and non-green investors (Heinkel et al., 2020; De Angelis et al., 2022), socially responsible versus financial investors (Oehmke & Opp, 2020), or agents with different preferences for green or ESG holdings (Pastor et al., 2021; Pederson et al., 2021). Some empirical evidence support this assumption where socially investors are willing to accept a returns discount compared to conventional investors (e.g., Bialkowski & Starks, 2016; Geczy et al., 2021; Renneboog et al., 2008, 2011). Other empirical studies point to a positive association between investing in ESG-tilted portfolios and superior returns (e.g., Edmans, 2011; Gibson et al., 2021). The contrasting empirical evidence, combined with evidence on SI fund flow resilience to poor past performance (e.g., Renneboog et al., 2011), lays the ground for doubt regarding whether SI investors hold any fundamentally different preferences compared to conventional investors. Recent investor surveys find that both social motives and financial motives are behind the decisions to invest in ESG (Boffo & Patalano, 2020). Therefore, the question addressed in this thesis is: is the growth in

SI fund flows driven by socially motivated investors? Answering this question would provide insight into the motivation behind the overall market movement of SI investors, leading to more successful policy interventions attuned to actual investor preferences. In addition, establishing empirical evidence for whether social preferences drive SI provides foundations for making realistic assumptions about SI investors' utility functions, which can then be used in theoretical models to derive and characterize market equilibrium.

SI mutual funds are often advertised and assumed to incorporate ESG issues in their portfolio construction and management. Most empirical studies on the difference between SI mutual funds and conventional funds examine their financial performance and find SI funds do not consistently and significantly under- or over-perform conventional funds (e.g., Hong & Kacperczyk, 2009; Bollen, 2007; Renneboog, Ter Horst and Zhang, 2008; Bolton & Kacperczyk, 2021). Other papers attempt to distinguish SI from conventional funds by their ESG performance and tend to conclude that SI (the definition of which varies between studies) funds have better fund-level ESG performance (e.g., Gibson et al., 2021). However, most such studies are limited by the scope, accuracy, and representativeness of ESG ratings, which they use to quantify ESG performance. As illustrated in existing research on the credibility of ESG ratings (e.g., Berg et al., 2022; Chatterji et al., 2016), issues such as the low correlation between data vendors, data rewrite, and over time methodology changes may lead to misleading ESG performance evaluations. Evidence from these studies show that investors cannot safely infer whether underlying assets of SI funds are different compared to conventional ones and if SI funds rebalance stakes in these assets in a manner consistent with their claimed ESG commitment. Theoretical work by Pederson et al. (2021) constructs a market equilibrium model with heterogeneity in how investors incorporate ESG information. This heterogeneity induces different portfolio choices along an ESG-efficient frontier between investors who consider ESG scores as risk-return signals, investors who prefer high ESG ratings, and investors who are purely mean-variance optimizing, leading to a four-fund separation. A related proposition is made by Paster et al. (2021) where there is a three-fund separation: in equilibrium, agents hold the market portfolio, risk-free asset, and a socially responsible portfolio depending on agents' ESG preferences. Both models suggest that SI and conventional funds should hold different assets and, thus, portfolios. My empirical analysis in Chapter 4 compares the portfolio holdings of SI and conventional funds.

To isolate social preferences, start by analyzing philanthropy in Chapter 2 since many aspects of social preferences are drivers of philanthropy: altruism, fairness, reciprocity, inequity aversion, guilt, and spitefulness. This thesis is the first to link philanthropy and SI to test if investment in SI mutual funds is driven by investors' social preferences (Chapter 2). Comparing SI to philanthropy - rather than, for example, conventional investment– has several advantages for identification purposes. Further, by using charity data, I can test the hypotheses on market-wide participants rather than on surveys or experimental data. In the fourth chapter, I move from investors' decisions to those of fund managers and scrutinize if SI mutual fund portfolios differ from conventional funds such that SI funds demonstrate a distinct taste for ESG. Overall, I present evidence that investors of SI funds are unlikely to be motivated by pro-social preferences under stringent empirical design using exogenous shocks. Further, several pieces of evidence suggest that SI funds do not hold substantially different assets than conventional funds. Importantly, SI funds' rebalancing decisions after ESG controversies show that they do not act in line with strong ESG preferences but are more likely to act as mean-variance investors under different scenarios.

I begin by showing in Chapter 2 that individuals make more charitable contributions after environmental shocks and lower contributions after shocks to charities' reputation or a decrease in the tax shield of philanthropy. In Chapter 3, I employ these shocks that triggered reactions in philanthropy to compare SI and conventional mutual fund flows. I find consistently insignificant SI fund flow responses regardless of the type of shock employed. The first two chapters demonstrate that SI flows are unlikely to be driven by social preferences. Thus, the motivation may come from risk-return considerations, which suggest that SI mutual funds might be managed similarly to conventional funds. Therefore, I turn to fund managers in Chapter 4. I document a considerable overlap between SI and conventional mutual fund portfolios and find evidence that SI and conventional funds hold similar assets. In addition, their portfolio choices in response to ESG controversies differ when the market disagrees on events' price implications: SI funds increase holdings of controversial firms. Although this may appear counterintuitive, I reveal the underlying reason by showing the positive response of ESG ratings to firm-level controversies. On the other hand, once ESG controversies are priced, SI funds do not attempt to signal higher ESG ratings by increasing holdings and thus behave as conventional mean-variance optimizing funds.

The first two chapters contribute to the existing literature on the social preferences of SI investors. Under the stringent empirical design and the use of granular data, I show that investment flows toward SI mutual funds are unlikely to be driven by social preferences. Prior literature identifying social preferences in SI uses investor survey data, OLS, or logit setting (such as in Riedl & Smeets, 2019; Hartzmark & Sussman, 2019; Barber et al., 2019). In contrast, I use archival data with a quasi-experimental approach to explore fund flow reactions to well-publicized and exogenous events that affect different potential motives for investing in SI. My research design has several benefits, the first of which is that I directly isolate social preferences. The typical strategy used in existing research explains differences in investment outcomes between SI versus conventional funds with observables, which control for their common monetary motive. Any conclusions about alternative explanations (including social preferences) are based on the part of the variation that is unexplained and remains in the residual (Geczy et al., 2021; Renneboog et al., 2008; 2011; Barber et al., 2021). An improvement on the typical design is to exploit shocks to the preferences for sustainable investments, especially climate-related (Bialkowski & Starks, 2016; Ramelli, Ossola et al., 2021). My approach is similarly shock-based, exploiting a variety of quasi-natural experiments (both environmental catastrophes and policy changes), but I first establish their relevance to social preferences. I only interpret SI flow results as driven by shocks to social preferences if they are also shown to have impacted philanthropy. Importantly, by explicitly including shocks to social preferences in the regressions, I do not rely on the residual variation but are able to directly measure their effect on SI.

Another benefit of the approach in Chapters 2 and 3 is that I consider a second group of shocks to the perceived effectiveness/attractiveness of the entire charity sector to deliver pro-social goals: a shock to the reputation of charitable institutions and a shock to the tax shield of philanthropy.[1] Tests based on these shocks allow me to eliminate alternative explanations, such as shock irrelevance to SI, and strengthen the

[1] For evidence that donation decisions are influenced by the efficiency of charities see Chen, 2009; Berman et al., 2018; Michel and Rieunier, 2012.

conclusions of whether social preferences are behind the demand for SI instead of risk/return motives. To illustrate this, consider an environmental disaster that triggers an increase in philanthropy. On the one hand, finding no effect on SI flows could mean that either social preferences are not a driver of SI or that the shock activates an aspect of social preferences better satisfied by philanthropy rather than SI. However, a shock that does not activate social preferences, such as removing the tax shield of philanthropy, makes it more expensive relative to SI to satisfy non-monetary motives. Thus, finding no effect on SI cannot be attributed to shock irrelevance to a certain aspect of social preferences. On the other hand, finding an effect of an environmental disaster on SI flows could mean that either investors are driven by social preferences or that they are motivated by strategic forward-looking risk/return implications on the firms in their SI portfolios. However, finding an effect of the tax shield shock on SI flows is independent of a particular risk exposure because it affects only the philanthropy sector and not the constituent firms of SI portfolios.

This second set of shocks also allows us to circumvent complications regarding investor heterogeneity. Such heterogeneity may concern the perceived substitutability or complementarity between SI and philanthropy under different shocks. When the perceived attractiveness of charities decreases due to a shock to their tax shield or reputation, I expect pro-social investors to increase resource allocation to SI regardless of their prior preference over SI or philanthropy.

In Chapter 3, I deepen the analysis by analyzing the behavior of SI fund managers. My research design benefits from using an objective measurement of mutual fund ESG performance. Most existing studies use ESG ratings to measure commitment/preference for ESG. These ratings, however, suffer from subjectivity and noise. In this thesis, I use objective ESG controversies as my identification strategy and measure how portfolio compositions change after these events while controlling for the market reactions to these controversies. My controversies measure is defined as the objective count of the number of ESG incidents reported in the media, thus likely exogenous to mutual fund demand of the affected company shares, given no advance insider knowledge spillovers. Since both SI and conventional funds are motivated to earn financial returns, and it is essential to separate risk-return motivations from ESG preferences, I examine how their reactions differ, controlling for the price implications of controversies. I achieve this by computing cumulative abnormal returns around controversies and running difference-in-difference models with interactions between the type of CAR (zero, positive or negative) and the SI fund indicator. This approach allows us to understand whether ESG preferences of SI funds are enhanced or depressed by pecuniary interests after controversies take place. I also argue that the reason SI acts similarly to conventional funds is that they prefer high portfolio ESG scores and not the avoidance of firms that produce negative externalities.

Unlike the assumptions made in existing theoretical and empirical papers, I find mutual funds currently marketed as SI do not react to negative social externalities caused by firms' actual ESG behavior (i.e., controversies). Instead, SI-labelled mutual fund managers are more concerned with ESG ratings and the price implications of controversies. I observe that when market participants disagree on the stock price effect of controversies (i.e., zero CAR, meaning there is scope for a financial gain by taking risks), SI funds increase their holdings of controversial firms while conventional funds decrease their holdings. Additional tests show that zero-CAR controversies lead to ESG rating increases contemporaneously and in a forward-looking manner. These results suggest that SI funds are willing to risk increasing their holdings on a controversial asset if better ESG ratings can justify a potential monetary loss.

In contrast, since conventional funds do not pursue better ratings, the risk of holding these controversial assets cannot be mitigated, and therefore, under these circumstances, they decrease their holdings. The actions of SI managers are in line with the socially responsible investors in Ohmeke & Opp (2020). They direct their scarce capital to firms with larger room for obtaining ESG improvement (Ohmeke & Opp refer to this as "avoided externalities"). However, I argue that because ESG ratings are subject to various methodological discrepancies across rating providers that lead to divergence across ratings and back-propagation issues, ratings poorly measure actual social impact. Thus, the incentive for SI mutual funds to make portfolio rebalancing decisions lies in the interaction of monetary factors and higher "perceived ESG performance" rather than "actual ESG performance". This finding also corresponds to my discussion in Chapters 1 and 2. Although various papers suggest strong and significant inflows to "high ESG profile" funds (Hartzmark & Sussman, 2019), I do not find individual SI investors to exhibit strong social preferences after impactful shocks to donation.

In addition, I indicate that the investment decisions of SI fund managers conform with those of other conventional market participants once clear price implications of ESG controversy news are established. When ESG controversy information is realized as positive or negative cumulative abnormal returns (CAR), there are no further opportunities for any market participants to outperform in the short term. Thus, both SI and conventional funds continue to hold their current stakes in controversial firms. My findings suggest that when facing positive CAR ESG controversies, both SI and conventional funds are akin to those in Pederson, Fitzgibbons, and Pomorski (2020), who are mean-variance optimizers and use ESG scores only to update their views on risk-return. Both types of funds find that the controversy-induced higher ESG ratings no longer hold any additional informational value and thus end up with identical portfolios. This is similar to a scenario where if there is no dispersion in ESG tastes across investors, asset prices fully adjust to reflect such “average” ESG taste, and all agents hold the optimal market portfolio (Pastor et al., 2021).

My results also have implications for individual investors. Hartzmark and Sussman (2019) find net inflow towards funds that received high Morningstar sustainability ratings. The positive association between ESG controversies and ESG ratings suggests that investors with pro-social preferences should not rely on using overall ESG ratings to examine firms’ social externalities. In addition, neither the Morningstar labelled SI funds nor the UN PRI funds can guarantee avoidance of or divestment from controversial firms.

Chapter 2 Donation flows and social preferences

2.1 Introduction

Prosocial behaviors are widely observed in the socioeconomic life of individuals. Economic agents spend leisure time volunteering, donate money to charities that materially benefit other people (e.g., deprived children, disabled people), and support causes that may not directly benefit themselves (e.g., protecting the environment, funding cancer research). The existence of social preferences beyond purely material self-interest is acknowledged by economists and psychologists (Andreoni, 1990; Becker, 1974; Preston & de Waal, 2002; Singer et al., 2004a; Berman et al., 2018; Harbaugh, 1998). The different motivations behind behaving prosocially modify the implications of how individuals maximize utility (Andreoni & Miller, 2002), the provision of public goods from private giving (Andreoni et al., 2017; Jennie & Loewenstein, 1997; Duncan, 1999; Feldman, 2010), and government subsidy and tax policy decisions (e.g., Reece & Zieschang, 1985; Hickey et al., 2019).

In this chapter, I examine motivations behind other-regarding (social) preferences and pro-social behavior: pure and impure altruism (Andreoni, 1990; 1990), empathy (Singer et al., 2004b; Burgoyne et al., 2005), reciprocity (e.g., Levine, 1998; Charness & Rabin, 2002), wealth and social image signaling (e.g., Bénabou & Tirole, 2006; Glazer & Konrad, 1996). Many of these papers present well-developed theoretical models for social preferences and use lab experiments or incentivized games to confirm the existence of social preferences. In this chapter, I instead perform tests using impactful exogenous shocks and observe if they lead to changes in charitable contributions and thus indicate changes in individuals' social preferences.

My first set of shocks is oriented toward humanitarian and natural disasters that affect other people or the environment with as little confounding as possible by other adjacent events. Humanitarian crises may stimulate compassion and emotional connection with humanitarian causes, so any altruism surge is purely selfless. I also examine the response of charitable flows to environmental or climate-related shocks similar to those studied by Bialkowski and Starks (2016) and Ramelli, Ossola et al. (2021). While environmental shocks contain a self-interest element as far as climate change affects the present quality of life of investors and the expected future quality of life of their descendants, I see them as also eliciting the desire to help others and/or the environment and ensure the welfare of future generations.

In attributing the charity flow effects of my shocks to social preferences, I exclude several alternative possibilities. First, philanthropy can be driven by motives beyond altruism: monetary (tax shield for donations - Feldstein & Taylor, 1976; Clotfelter, 1980; Auten et al., 2002; List, 2011; Hickey et al., 2019), wealth and social image signaling (Glazer & Konrad, 1996; Bénabou & Tirole, 2006; Harbaugh, 1998). I show the design and result from the second set of tests which exploits shock to the attractiveness of charitable contributions: shock to tax incentives of giving and shock to the reputation of charities. In these tests, I am able to clearly distinguish between social preferences and two alternative motives for donation: monetary incentive of tax reduction and the perceived effectiveness of charities in making social impact.

In my first set of tests, I exploit the following shocks: humanitarian crises caused by the Haiti earthquake in 2010 and the Syrian refugee crisis in 2015; the environment-related shocks caused by the BP oil spill in 2011, the US presidential election in 2016, and the US announcing its withdrawal from the Paris Agreement in 2017. In my second set of tests, when faced with a shock to the reputation of charities or a reduction in the tax shield of charitable giving, I expect individuals to reduce their charitable donations due to changes in the monetary incentives of donating. My second set of tests uses the US Tax Cut and Jobs Act 2017 and three charity scandals: the Oxfam misconduct scandal in 2011, the Kids Company financial scandal in 2013, and the Cancer Fund of America scandal in 2013. Results from my first set of tests consistently demonstrate that social preferences drive donations after environmental shocks (except for the BP oil spill shock). In my second set of tests, the shocks used already entail a decrease in the attractiveness of philanthropy and the monetary incentives of donating. I establish that charitable donations decreased after each of the second set of shocks. To exclude the social image and reputation signaling motives, I employ a separate test on the change in the volume of named versus anonymous donations around one of my shocks.

An additional benefit of my empirical design stems from the granularity of my philanthropy data, which allows us to use a strict definition of treated and control charities, match them based on multiple pre-shock covariates, and design as stringent tests as possible. Consider the different existing preference orderings of investors among charitable causes: climate change, social justice, animal welfare, arts, etc. An exogenous shock, such as an environmental crisis, would trigger compassion and enhanced altruistic behavior in donors who prioritize the environment relative to the ones who prioritize the arts. Thus, I would expect stronger demand for donating to charities with an environmental focus and a possible corresponding increase in flows towards SI funds supporting the same cause. I employ granular data of the population of UK and US charities (ranked first in the world and Europe, respectively, by the percentage of GDP donated to non-profit organizations by individuals (CAF, 2016) to perform shock strength tests on carefully designed treated and control groups by charity primary purpose classification.[2]

This Chapter is structured as follows. Section 2.2 reviews the literature on charitable giving, highlighting potential motivations such as social preferences and tax incentives behind philanthropy. In section 2.3, I detail my empirical strategy, starting with describing the shocks employed in my tests, followed by the design of the difference-in-differences model, where I present the sharp definition of treatment and control groups of US and UK charities, accompanied by summary statistics on these charities. Finally, section 2.4 presents results from the two sets of empirical tests and offers additional analysis on alternative explanations to my main finding.

2.2 Related Literature

Social Preferences and Donations

One underlying assumption of many existing economic and finance theories is that individuals are driven solely by self-interest. Yet economic and psychological studies point out the existence of other-regarding preferences (e.g., Smith, 1759; Becker, 1974). Under unconditional or pure altruistic intentions, agents derive strictly positive utility with respect to material resources others receive (e.g., Andreoni & Miller, 2003). However, a crucial distinction exists between pure and impure altruism Andreoni (1989, 1990). Impure altruism refers to the "warm glow" from the act of giving to others and

[2] The top ten countries are as follows: US, New Zealand, Canada, UK, Republic of Korea, Singapore, India, Russia, Italy, Netherlands (CAF, 2016).

therefore is considered not purely selfless.

Empathy is another important determinant of social preferences (Singer et al., 2004b). In their psychological experiment, Singer et al. (200b) find that when a woman observes painful stimulation received by her partner, an empathetic response is automatically generated in the area of the brain that processes affective rather than sensory components of pain. This experimental result is consistent with an earlier study by Preston and de Waal (2002). They show that empathy does not require physical experience or conscious information processing. The existence of empathy thus undermines individuals' sole pursuit of self-interest and may therefore induce social preferences and behavior such as donating.

Guilt is also a psychological motivation behind prosocial behavior related to empathy. In a randomized field experiment, Andreoni et al. (2017) demonstrate that individuals are aware of the link between empathy and prosocial behavior. Thus, in the authors' experiment, subjects choose to avoid feelings of guilt and empathy by avoiding fundraisers at their door. Similar evidence also shows that guilt can stimulate donation intentions and prosocial behavior (Hibbert et al., 2007; Huhmann & Brotherton, 1997; Basil et al., 2006).

Nonetheless, it is worth noting that empathy is not consistent (Hichri & Kirman, 2007) but rather heterogeneous across individuals (Singer et al., 2006) or even considered a transitory response to specific social contexts or social interactions (Kirman & Teschl, 2010). In this regard, empathy-induced pro-social behavior, such as donating to charities, may also be context-based. In fact, individuals make donation choices based on whether they emotionally connect with a cause or can align their social concerns and ideals via this recipient, but not the amount of welfare improvement achieved (Berman et al., 2018; Andreoni, 2006). . As a result, in events of severe shocks to humanitarian values that spur human compassion and care for others, charities that aim to save lives or relate to humanitarian aid may receive more donations than non-humanitarian-focused charities such as art-related charities. Nonetheless, empathy-induced pro-social actions also depend on the "identifiability" of victims. Agents care more when victims are geographically closer to themselves (Loewenstein et al., 2008) and when disasters have more concentrated than dispersed effects, i.e., affecting 100% of the population instead of 10% of the population (Jennie & Loewenstein, 1997).

Another example of the identifiability effect is that people are more likely to give to charities treating specific diseases if someone in their family is exposed to the same illness (Burgoyne et al., 2005). In addition, fundraisers also exploit this “vividness” effect strategically, for example, by using pictures of kids to draw public attention to monetary appeals to tackling malnutrition (Jennie & Loewenstein, 1997).

Another source of social preferences is reciprocity. Depending on the expected future material outcomes due to the actions of others, individuals may choose to act reciprocally towards different “types” of agents (Levine, 1998). The concept of reciprocity can be understood and modeled with classic game theory or Prisoner’s Dilemma under one or repeated interactions between agents (Fehr & Schmidt, 2006; Gintis, 2000; Axelrod & Hamilton, 1981). Reciprocal behavior ensures that kind or altruistic agents are “rewarded”; unkind or unfair agents are “punished”. For example, individuals will withdraw from sacrificing for social welfare when others are not concerned (Charness & Rabin, 2002). Most participants from a sequential Prisoner’s Dilemma game appear fonder of fair instead of unfair players (Signer et al., 2004b).

Signaling and pro-social behavior

Many economic and psychological experiments on social preferences discussed above are conducted in the presence of experimenters who observe the subjects’ actions. Levitt and List (2007) predict that this experiment set-up creates incentives for selfish individuals to act as having social preferences for social image concerns. Hoffman et al. (1994) support this theory: when researchers do not observe subjects, the probability of giving to others is reduced from 46% to 16%, which indicates a taste for being highly regarded by others. Andreoni and Bernheim (2009) also report that the resource split between experiment participants tilts towards equality (50 to 50) as more scrutiny is in place. Similar evidence is provided by Haley and Fessler (2005) and Harris and Munger (1989). Therefore, reputational concerns or reputation signaling can be used to explain conformity with pro-social behavior such as donation and generosity (Bénabou & Tirole, 2006; Ariely et al., 2009). In this Chapter, I avoid such potential overstatement of individuals’ social preferences by using actual donation flows and revealed preferences, not experimental data. I also explicitly test that the reputation effect is unlikely to drive my results for one of my shocks.

Glazer and Konrad (1996) consider a wealth-signaling motivation for philanthropy. In their theoretical model, individuals want their donations to be noticed by others but not to benefit others. The authors use university alumni donation reports and find the rareness of anonymous compared to named donations, which conforms with the signaling motivation. Harbaugh (1998) also identifies the taste for prestige as a critical driver behind philanthropy using law school alumni donation data and notes that this taste may be heterogeneous across individuals and even stronger for public figures.

Monetary incentives and disincentives for giving

Beyond social preference and signaling incentives, other extrinsic motivations can also affect pro-social behavior, such as giving and blood donation (Ariely et al., 2009; Mellström & Johannesson, 2008). For example, the monetary driver behind charitable giving is its tax shield, whereby donation flows are elastic relative to the “tax price” of making contributions (List, 2011; Duquette, 2016). Among existing literature, the median of the estimated tax elasticity of charitable donation is -1.2 (Peloza & Steel, 2005), and it is reactive to the provision of tax credits and tax deductions (Auten et al., 2002). It is therefore crucial in this chapter to also test how individuals’ pro-social behavior reacts to such monetary incentives.

Reece and Zieschang (1985) find evidence that a 20 percent tax credit on donations or a 30 percent credit may induce increased donations but display declining numerical incentive efficiencies as income increases. Auten et al. (2002) model donation behavior using a dynamic consumption model under the presence of income taxation. They test the model using data from the US Internal Revenue Service (IRS) from 1979 to 1993 on the amount of tax deductions claimed by individuals. The authors observe several generalizable patterns from the data. First, contributions from individuals increase with income: the average contribution from donors increases monotonically with family income. Second, a relatively small number of individuals account for the bulk of all giving within each income level. Thirdly, the sample is stratified with oversampling of high-income individuals; these wealthy donors donate a disproportionate fraction of the total contributions. The panel data sample also allowed the authors to separate transitory and permanent aspects of the price and income effects from taxation and income changes. They find that taxes affect the individual level of contributions through both price and income effects, each with a permanent and transitory component. Individual donation behavior is most significantly affected by a permanent price effect. For example, for an individual facing a marginal tax rate of 30%, a flat-tax proposal would lead to a 25% to 36% decrease in charitable contributions.

After the January 2010 Haiti earthquake, the Canadian province of Quebec implemented a policy allowing local taxpayers to claim tax credits on their donations for the tax year 2009 rather than 2010. Under this quasi-experimental setting, Hickey et al. (2015) find that in the tax year 2009, Quebec taxpayers donated relatively more to charities than the rest of Canada, where the policy was not in place. In addition, when there are tax incentives, donating money may also become more attractive compared to other forms of giving, such as volunteering (Duncan, 1999; Feldman, 2010), indicating that the importance of private giving as a means (instead of, e.g., volunteering) to achieve social preferences for individuals can be sensitive to the financial incentives of doing so.

As such, a shock to tax policies affecting donation deductions may serve as a plausible shock to disentangle financial and non-pecuniary motives (e.g., social preference-based motives for "doing good"). However, there are also significant discrepancies and inconsistencies in definitions of charitable giving and tax-deductible charitable giving across the globe (List, 2011). These difficulties limit the feasibility and comparability of conducting cross-country analysis using natural experiment settings with tax policy shocks. Nonetheless, a tax policy shock within a single country may still generate insight into whether donation reacts to price incentives.

In addition to contributions by the public, charities and non-profit organizations may receive government grants. Under the assumption that individuals are unconditionally altruistic, individuals are indifferent between the ultimate sources of donation funds, and therefore increased giving by the government completely crowds out private giving (Bruni & Zamagni, 2013). However, consistent with the "warm glow" (Andreoni, 1989; Gruber & Hungerman, 2007) and signaling motives behind donations (Glazer & Konrad, 1996), empirical evidence suggests that government crowding out is only partial (Steinberg, 1989; Andreoni, 2006; Glazer & Konrad, 1996). For example, Andreoni and Payne (2003) quantify the effect of government grants on fundraisers; they find that an additional $1000 in government grants decreases fundraising expenditures to various extents for different types of non-profit organizations. In addition, social service organizations incur an average decline of 35% in fundraising expenses. Moreover, Okten and Weisbrod (2000) and Payne (2001) suggest that givers will view government grants as financed by the income tax they paid, therefore making fewer voluntary donations: increasing federal research funding by one dollar depresses private donations by 9% to non-research universities and 45% to liberal arts colleges (Payne, 2001).

Government grants may also lead to fewer fund-raising efforts, resulting in reduced public contribution, an outcome like the one resulting from direct crowding-out of donations. I account for this potential crowding out effect of government grants in my tests by controlling for these grants received by charities.

List (2011) documents significant cross-country differences in tax policies and government funding for non-profit organizations. For example, the United States allows income tax deductions of charitable contributions for up to 50%, whereas the number in European countries such as Belgium is capped at 35%. Meanwhile, by mid-2007, 30.5 percent of total monetary support received by US charities came from government grants, while the UK and Belgium received much higher government support, accounting for 40.6% and 76.8% of total monetary support received, respectively (List, 2011). Similar to studies discussed earlier, List (2011) argues that different tax policies and subsidy plans explain a major part of cross-sectional differences in charitable donations. Tocqueville (1835; 1840) also documents that the US and UK represent two distinct institutional and cultural settings about charitable giving by private donors versus the state.

Charity efficacy and reputation

Outcome efficacy and the reputation of charities can also affect donation decisions (LeClair, 2019). Although researchers acknowledge that it is often impossible for the public to thoroughly understand the operations and outcomes of their charitable contributions, outcome efficacy still matters to donation intentions (Cheung and Chan, 2000). For example, Chen (2009) found that charities meeting specific standards receive higher public support than those that do not. This finding is supported by empirical test results from Gordon et al. (2009) where donation positively correlates with changes in third-party charity ratings.

By contrast, Berman et al. (2018) show that when discussing past donation choices being ineffective in improving welfare, participants become discouraged and reduce donations to the ineffective hypothetical "charity". Michel and Rieunier (2012) examine non-profit organization brands and suggest that brand image, indicative of charity efficiency, could potentially explain up to 31% of donor intention to give money and around 24% of their intention to give time. These findings are also consistent with earlier studies on the impact of identifiability in dictator games, where dictators are more willing to donate if provided with more information about recipients (Bohnet & Frey, 1999) and if recipients are well-established charities such as the American Red Cross compared to anonymous individuals (Eckel & Grossman, 1996).

Overall, existing literature on social preferences and prosocial behavior identifies various intrinsic and extrinsic motivations behind other-regarding intentions and actions such as charitable giving: pure and impure altruism stemming from empathy (Singer et al., 2004b) or "warm glow" (Andreoni, 1989;1990), reciprocity (Levine, 1998), wealth and social image signaling (Levitt & List, 2007; Bénabou & Tirole, 2006), and monetary benefits in the form of tax breaks (Hickey et al., 2019; Andreoni & Payne, 2003). In this chapter, I examine whether social preferences manifest as changes in charitable donations when exogenous shocks such as humanitarian disasters and environmental catastrophes occur. I also test if tax policy changes result in different donation behavior. Finally, given the relevance of outcome efficacy and charity reputation in private donation decisions, this chapter empirically tests if shocks to these two factors are reflected in donation outcomes. The following section describes the shocks studied and my empirical methodology.

2.3 Empirical Analysis

Consider the different preference orderings of individuals among charitable causes: humanitarian, climate change, social, animal welfare, arts, and more. An exogenous shock like a natural disaster or a humanitarian crisis would trigger compassion and enhanced altruistic behavior. Thus, according to these preference orderings, I expect a stronger demand for donating to charities with relevant causes.

2.3.1 Exogenous shocks

Humanitarian shocks

In January 2010, Haiti suffered an earthquake leading to an estimated 250,000 death toll and 5 million displaced. Based on Hickey et al. (2019), donations in Quebec, Canada, increased relative to the rest of the country after the 2010 Haiti earthquake. However, the relative increase was due to government policy increasing tax incentives for potential donors. It may also be insightful to examine the sensitivity of donations to a natural disaster without any tax incentive. Therefore, I use the same Haiti earthquake shock as Hickey et al. (2019) with UK charity income data. In 2010, there was no policy change in the UK targeting charitable donations, preventing potential shock contamination. Additionally, since the earthquake took place in January 2010, potential donors would have enough time to react to this shock before the end of the financial year 2010. Therefore, 2010 is also considered a post-shock year in this analysis.

Another humanitarian shock I examine is the 2015 European refugee shock. In September 2015, the picture of an 8-year-old Syrian refugee Aylan Kurdi was released on multiple news sources in the UK and across the globe, including BBC News, the Guardian, the Independent, Wall Street Journal, Reuters, and more. The shock stands as an impactful humanitarian shock to the public. Some of the headlines at the time read as “Image of Drowned Syrian Boy Echoes Around World” (WSJ, 2015), “Refugee aid charities see a surge in donations after image of drowned Syrian toddler Aylan Kurdi moves the nation” (Independent, 2015), “Britons rally to help people fleeing war and terror in the Middle East” (The Guardian, 2015), “Photo of dead Syrian boy boosts fundraising 100-fold: study” (Reuters, 2017). Some news agencies at the time posted direct links to large charities that operate rescue boats and provide foreign aid. For example, the Guardian published an article on 3rd September 2015 containing information about the names and details of UK charity that helps refugees and asylum seekers (The Guardian, 2015). Even in 2017, one of the Guardian's articles on the refugee crisis was on how individuals can help refugees in Britain, and the article also contained specific charities' information (The Guardian, 2017). Hence, it is reasonable to expect that donations to relevant UK charities will increase from 2015 onwards.

Environmental shocks

In this Chapter, I also examine environmental shocks. The first one is related to the BP oil spill in April 2010. The BP-operated oil well site leaked around 4.9 million barrels of oil into the Gulf of Mexico. The oil spill had an enormous and lasting environmental impact, especially on marine flora and fauna. In Louisiana alone, 4.6 million pounds of oiled materials were removed from beaches in 2013 (Elliot, 2013). White et al. (2012) show that coral colonies near the oil spill site presented significant signs of stress. In addition, the spill incurred substantial settlement and compensation costs for BP (Mason, 2010). Similarly to Bialkowski and Starks (2019), who argue that this event affected investment towards environment-focused mutual funds, I hypothesize that the BP oil spill shock triggered social preferences that lead to charitable giving towards environmental and animal-focused charities.

Another environment-related shock I consider is the unexpected election of Donald Trump in 2016 – a well-known climate change skeptic who repeatedly signaled his intent to pull the United States from the Paris Agreement if elected. Trump did not include any environment-related topics in his platform at the time (Harrington, 2016). His coming to power can be further seen as an environmentally damaging event due to his announcement of intending to nominate Scott Pruitt, an opponent of the Obama administration's measures to tackle climate change, as the director of the Environmental Protection Agency (Volcovici & Shepardson, 2016). This nomination led to environmental concerns; for example, 447 former EPA employees wrote a joint letter opposing the nomination (Henry, 2017). It is worth noting that the 2016 election might not have been unexpected for all investors, especially those who voted for the Republican party. However, even Republican voters may not have expected the election outcome since the democratic candidate's consistent lead in the national popular vote based on pre-election polls is usually a precursor to winning (Erikson & Wlezien, 2012). Major prediction sites such as the New York Times, FiveThirtyEight, and seven others all predicted a Clinton administration (Katz, 2016; Wright & Wright, 2018), while Pennsylvania, Michigan, and Wisconsin broke unexpectedly for Trump with a combined margin of only 0.56% (Kennedy et al., 2018). Overall, based on the available information from pre-election polls, it is unlikely that most voters anticipated the election outcome.

The third environmental shock I investigate is the announcement made by the United States President on 1st June 2017 that the US would cease all participation in the Paris Agreement on climate mitigation. The announcement spurred various national and international reactions. For example, the governors of 13 states, including New York, California, and Washington, formed the United States Climate Alliance to pledge to uphold the Paris Agreement. United Nations and several leaders from other countries also commented on the US decision as a “disappointment” (Fox News, 2017). The withdrawal announcement also led to protests and petitions in various states across the United States. I hypothesize that environmental charities will receive increased flows as the public may wish to distance themselves and offset the impact of the US government decision.

Reputation shock

Conversely, charitable donations may decrease when charities experience scandals related to its finance or conduct. In October 2011, Oxfam Great Britain’s director in Haiti resigned after an internal investigation regarding six staff members breaching Oxfam’s behavioral code of conduct. This scandal received wide publicity and brought Oxfam’s name into disrepute. On Factiva, the news article search engine, in 2011 alone, over 110 news pieces were published in the UK relating to the Oxfam scandal. Due to the wide publicity, Oxfam may witness a significant decrease in its voluntary income post-scandal year as donors become aware of the scandal and lose faith in the charity. The shock also affected a large, influential charitable institution in the UK and negatively impacted the whole charity sector (Green, 2018). In 2013, there was another shock to an England-based charity, Save the Kids Company (known as Kids Company), where the charity was investigated for misuse of government grants and poor financial management. The scandal received wide media attention, and the charity was de-registered in 2013. In the same year, the US charity Cancer Fund of America was reported as one of America’s worst charities by a US news agency and the Centre for Investigative Reporting. The charity was shamed for its high management costs, high director salaries, and lack of effectiveness in delivering its charitable purpose. The outbreak of the scandal was followed by formal investigations that led to the charity's dissolution in 2015. I examine whether US charities with medical purposes received lower donations due to the scandal.

Tax policy shock

The last shock I consider is regulatory reform to tax policy in the US that negatively affects the attractiveness of charitable giving in its ability to shield income from taxes. Using Canadian data, Hickey et al. (2019) show that allowing taxpayers in Quebec to claim tax credits earlier than in other provinces led to 9% more donations. This result is suggestive of the impact of a favorable tax policy. In December 2017, the Trump administration signed the Tax Cut and Jobs Act (TCJA) into law. Under the new tax policy, the standard deduction increased from \$6,350 to \$12,000 (and from \$12,700 to \$24,000 for couples). However, taxpayers can only claim charitable contributions through itemized taxable income reductions, not standard deductions. The doubling of the standard deduction removes the benefit of itemizing deductions and the earmarking of certain funds for philanthropy by individuals. According to IRS income tax return data for the tax year 2017, the top three itemized deduction categories (medical expenses, state tax, and interest paid) accounted for 26% of annual gross income (AGI), while charitable contributions were the fourth highest category with 5%. Consider a taxpayer at the lower bound of the median tax bracket (\$38,700 – unchanged by TCJA, but the upper bound and tax rate are reduced from \$93,700 to \$82,500 and from 25% to 22%, respectively). Before TCJA, she would claim \$10,062 for the top three itemized deduction categories (38,700 x 0.26) and \$1,939 for charitable contributions (38,700 x 0.05). This amounts to \$12,000.60 in total deductions, which is \$5,650.60 above the standard deduction of \$6,360 and provides an additional tax shield of \$1,412.70 (\$5,650.60 x 0.25), which incentivizes her to itemize. After TCJA, her total deductions are fully covered by the increased \$12,000 standard deduction, and there is no longer an incentive to itemize. According to a report by Giving USA (2019), inflation-adjusted individual giving declined by 3.4% in 2018 after the implementation of the policy.

I scrutinize the years around all shocks and ensure there are no contaminating events that would compromise the clean definition of pre- and post-shock years. I perform a difference-in-differences (DiD) analysis with the US and UK charity income data. Due to data limitations for US charities in years prior to 2013, I exclude the BP oil spill from the US analysis.

Table 2.1 Shocks to altruism

Date	Country	Charity Name	Shock/Scandal	News Agency
A. Natural/Man-caused Disasters				
20th April 2010	US	N/a	BP Oil Spill	
12th January 2010	Haiti	N/a	Earthquake	
2nd September 2015	UK	N/a	The photo of a refugee boy, Aylan Kurdi, found lifeless on a beach, was spread worldwide	The Guardian, Independent Wall Street Journal Reuters
1st June 2017	US	N/a	US announced withdrawal from Paris Agreement	The Guardian The New York Times BBC Reuters
B. Charity Scandals				
6th September 2011	UK	Oxfam International	Allegations of managers' 'inappropriate behavior'	Guardian
6th June 2013	US	Cancer Fund of America	Misuse of donations	The New York Times Campa Bay News Advisory Board
2013	UK	Kids Company	Misuse of government funding Abuse/ Sexual assault allegations	Financial Times BBC News The Guardian
C. Tax policy shock				
22nd December 2017	US	N/a	The announcement of the Tax Cut and Jobs Act 2017, to be implemented for the tax year 2018	The New York Times Fortune

Source: Factiva

For both US and UK, I run the following DiD model:

$$Inc\_vol_{i,t} = \alpha + \beta_0 Treat_i + \beta_1(Post_t \times Treat_i) + \beta_2 Post_t + \Sigma_\gamma X_{i,t} + \varepsilon_i \quad (2.1)$$

In Equation (2.1), the dependent variable $Inc\_vol_{i,t}$ refers to the inflation-adjusted voluntary income (total contributions) received by the UK (US) charity *i* in year *t*. Given the close timing of Haiti earthquake and BP oil spill, one may argue that if both shocks trigger (or do not trigger) changes to donations, it is impossible to separate their individual effects. To address this potential issue, I exploit differences in charity objectives and use their classification to define the treatment and control charities. The definition of $Treat_i$ reflects the classes of charities closest to the set of values most affected by each of my shocks. For example, after environmental shocks, individuals may become more passionate about environmental protection, thus donating more to environment-focused charities. In this case, I define $Treat_i$ to be equal to one for environmental and animal protection charities and zero for art and culture-related charities, as their objectives are in stark contrast to those of the treated charities.

I employ the UK Charity Commission (CC) charitable purpose classification and the US National Taxonomy of Exempt Entities (NTEE) IRS activity codes to identify charity objectives and define $Treat_i$. In the UK CC database, charities report three-digit classification codes in the data field: "classno" to indicate: their charitable objectives, such as "saving of lives", "overseas aid", and "art" with classification codes 103, 106, and 109; whom they serve, such as "children" with 201, or "the general public/mankind" with 207; how they serve, such as "make grants to individuals", "provide services", with codes 301 and 306. In total, 34 classification codes are split between the three types of the information above and stored as text in "classno". Any charity can have more than one 3-digit code in its "classno" entry as it sees fit. For example, a charity may have "classno" entry: "103; 106; 207; 301; 305". When defining my treatment and control group, I ensure that the charities do not contain unrelated purposes to avoid contamination. However, many charities in the CC database hold the classification code "101", which is defined as "General charitable purposes" along with other more specific purposes such as "saving lives". I do not exclude such charities to ensure a sufficiently large charity sample. I present the classification number reference in my appendix.

In the UK dataset, charities with the "saving of lives" objective contain medical research bodies such as Cancer Research UK, whose objectives may not be as relevant during the humanitarian value crisis I study. I identify medical-research-related charities within the "saving of lives" classification by examining whether they have certain words such as "cancer", "heart", and "Alzheimer's" in their objective statements.

In the US, each charity is assigned a unique NTEE code for its primary purpose, and the NTEE code is contained in the IRS Form 990 annual extract dataset. The NTEE code classifies charities into 26 major groups by their primary charitable purpose coded with one alphabet letter from A to Z. Within major NTEE groups, organizations are further defined by decile level codes and centile level codes (2nd and 3rd digit), subdividing organizations into the major groups by specific non-profit activities and specific organization types. In the IRS F990 annual dataset, charities have one “ntee1” entry for the major group and one “ntee1_3” entry for detailed subgroups. For example, an “Environment” charity could be classified with an “ntee1” code: C, and a “ntee1_3” code: C36 for the specific “Forest Conservation” subgroup. Charities only have one “ntee1” and one “ntee1_3” entry; I use both to ensure no potential data entry mistakes and to define my treatment group for US charities. The complete list of NTEE codes is available in my appendix.

However, large charities operating with multiple charitable purposes will only have a single NTEE classification category. Oxfam America, for example, is classified as “International Development, Relief Services” with NTEE code Q30, yet its website also clearly discloses its work on climate change issues. The lack of accurate classification would be more problematic for larger charities as they are more likely to be multi-objective. To minimize this concern, I exclude the top quartile of US charities by total revenue. In addition, I exclude the smallest US charities in the bottom size quartile for two reasons: 1) they represent only 0.3% of the total revenue of all charities in the US sample, and 2) they do not receive third-party ratings and therefore have extremely low visibility (Krasteva & Yildirim, 2016). To maintain consistency, I focus on the same two quartiles of UK charities. For both countries, I report results from robustness tests with charities of all sizes, i.e., including the top and bottom 25th percentile charities in Appendix 1.1.

Additional to defining treated and control groups based on their purposes, I conduct covariate balancing tests for each shock examined. I first run logit regressions based on a list of 1-year lagged charity financial variables: total revenue, total expense, total assets, total government grants, fundraising expenses, and payroll tax for US charities; total income, charitable activity expenditure, total assets, number of volunteers and number of employees for UK charities. I then match each treated charity with the closest control charity based on the estimated propensity scores. Covariate balance and parallel trend tests are shown in Appendices 2.2 and 2.3.

The definition for the post-shock year $Post_t$ differs by shock and country of analysis. For most shocks, I employ a two-year post-shock period after the event. $X_{i,t}$ is a matrix of control variables that includes income and the number of volunteers as proxies for charity size. Andreoni and Payne (2003, 2011) show that a $1,000 increase in government grants leads to a significant drop in charitable contributions; thus, government grants are included as a control variable. I also control fundraising expenses, which List (2011) documents to vary substantially among charities. Detailed descriptions of charity classification codes, definition $Treat_t$ and $Post_t$ under different shocks, and definitions of control variables are presented in the appendix.

In addition, although it is not possible to conduct a DiD analysis with the scandals of Oxfam and Kids Company because the treated group contains a single institution, my time-series plots of $inc\_vol_{i,t}$ for the two charities show sharp drops in donations received after the scandals (see Section 2.4).

2.3.2 Data & Sample

I obtain UK charity data from the UK Charity Commission, an official regulatory body governing the registration, reporting, and regulation of charities in England and Wales. The commission's dataset contains detailed time-series data on charities registered in England and Wales, including name, registration number, classification by charitable purpose, and financial data. In addition, I obtain charity financials data from CC over the years 2005-2018. At any given year, if a charity has a total annual income greater than £500,000, it must report its detailed financial statement to the CC, including different income sources, expenses, and the charity's assets. Among all charity financial data reported, I choose "voluntary income" as my dependent variable, which CC defines as total donations received by the charity from the public.

In addition, the CC database includes non-profit organizations that are not charities, such as universities, schools, hospitals, workers' unions, research centers, churches, and local councils. These organizations are excluded from my sample of analysis. For each shock examined, I ensure that (UK) charities are over three years old as of the year before the shock to allow pre-shock covariate balancing and avoid potential bias due to new charities receiving large lump sums of donation at inception.

In the US, charity data are available from the US Internal Revenue Service (IRS). The IRS publishes annual data extracts of 501(c) non-profit organizations containing information from their required Form 990 filings. These annual extract data are available from the National Bureau of Economic Research (NBER) and the IRS website for 2009 to 2019. I show summary statistics for both countries for the period 2009-2018, but due to limited availability between 2009 to 2012, I use years after 2013 for the US. All charity financial data are deflated with CPI indices for UK and US and expressed in thousands of pounds/dollars.

Table 2.2 describes the UK (Panel A and B) and US (Panel C and D) charity samples. Since there is no exact match of variable names and definitions between the US and UK data, I manually screen through the variable descriptions to find the pairs of UK and US variables with the most closely related definitions (see Data Definition Table in Appendix). For example, the variable reflecting individual donation flows in the UK data is voluntary income, while in the US dataset, it is referred to as total contributions. These two variables will be used as the main dependent variable for analysis in this chapter. They essentially carry the same underlying meaning: the amount of income charities obtained via donations from the public (e.g., one-off donations from a public member of society). It is worth noting that both of the two variables may contain regular donations such as direct debits to charities, but the exact amount of such regular payments are not reported in neither US nor UK database. In addition, a further difference between US and UK charity data is that I do not observe the number of volunteers or charitable activity expenditure in the US dataset, while fundraising expenses and lobbying fees are unavailable in the UK database. Nontheless, I still include these variables as controls in US and UK regression models respectively. Although one should note that these discrepancies makes direct comparisons between US and UK results difficult.

Table 2.2 Charity summary statistics

Panel A. UK charity sample (in thousands of £) – mean by year

	2011	2012	2013	2014	2015	2016	2017	2018
Voluntary income	591.9	638.9	659.9	641.1	616.3	583.4	630.8	657.5
Total income	1,536.3	1,591.5	1,607.0	1,631.8	1,623.1	1,648.6	1,640.5	1,613.9
Charitable activity expenditure	1,181.6	1,241.7	1,247.8	1,277.6	1,239.2	1,336.9	1,284.8	1,265.6
Total assets	3,704.4	3,889.7	4,347.8	4,728.8	4,271.1	4,859.8	5,171.5	6,112.4
Number of volunteers	60.5	215.7	200.6	158.9	266.2	285.4	359.6	1,146.5
Number of employees	23.0	22.6	22.0	21.3	24.3	25.3	25.0	24.5
Number of charities	403	438	467	489	469	493	523	101*

Panel B. UK Charity Sample (in thousands of £) – overall summary statistics

	Mean	Median	Min	Max	Standard Deviation
Voluntary income	603.6	448.9	0.0	3,749.3	633.7
Total income	1,573.1	1,339.7	660.5	3,796.7	732.9
Charitable activity expenditure	1,220.1	1,052.9	0.0	9,792.2	814.8
Total assets	4,230.1	1,059.5	0.0	146,906.0	9,744.2
Number of volunteers	234.5	6.0	0.0	75,000.0	2,131.2
Number of employees	24	18	0	179	22
Number of charities			990		

Panel C. US charity sample (in thousands of $) – mean by year

Total contributions	132.0	142.6	146.3	153.2	158.2	158.6	165.5	171.3
Total revenue	215.2	224.8	233.2	241.2	245.6	249.8	260.4	265.3
Total expense	208.7	207.4	218.6	224.7	224.2	229.9	236.1	247.6
Total assets	389.8	346.3	330.7	360.3	375.2	355.3	369.7	390.8
Total government grants	574.9	587.0	564.5	617.5	640.1	680.3	665.5	731.4
Fundraising expense	6.8	7.6	6.4	7.5	7.4	6.5	6.2	7.3

Payroll Tax	3.1	3.8	3.7	3.8	3.9	4.1	4.1	4.2
Number of charities	77	538	638	719	780	849	935	972

Panel D. US Charity Sample (in thousands of $) – overall summary statistics

	Mean	Median	Min	Max	Standard Deviation
Total contributions	158.5	148.4	0.0	493.7	99.0
Total revenue	248.8	240.5	34.3	474.7	84.9
Total expense	231.5	222.3	0.0	1,022.9	94.1
Total assets	359.1	156.4	-22.6	8,306.2	611.6
Total government grants	672.7	589.7	0.0	9,576.0	520.0
Fundraising expense	6.9	0.0	0.0	429.0	24.6
Payroll Tax	4.1	0.0	-0.0	49.5	6.1
Number of charities			849		

Panel A (C) reports the means of financial variables for each tax year for UK (US) charities that are ranked in the 2nd and 3rd quartile by charity size measured by total income (total revenue). Panel B (D) reports the overall summary statistics for all UK(US) charities of different size groups. The sample covers only the charities selected for my diff-in-diff analysis of charity flows environmental charities (treatment group) and art charities (control group). In the US, the IRS yearly financial data extract of non-profit organizations is publicly available starting in the tax year 2012; no separate records are available for prior years. As a result, my US sample size is much smaller in the year 2011. In the dataset obtained from the UK Charity Commission, there is a limited number of charities that have reported for the tax year 2018, which is reflected in the number of observations in the last column of Panel A.

* Note that the number of observations for year 2018 is limited to only 108 because at the time of this analysis, charity income data for financial year 2018 was not released by the Charity Comissions UK. The datasets downloaded and used in this analysis, where charity financial data were compiled on an annual basis, were de-commissioned in September 2020, and no longer available on Charity Commisions' website. The replaced data service employs new variable names and definitions and reporting frequency (i.e., on a rolling basis). Therefore, data available from the current Charity Commissions' database is unlikely to be fully comparable with legacy data used in this Chapter. However, I still try to compile summary statistics for financial year 2018 with new data definitions (see appendix)

In the full charity sample, I have data from 2008 to 2018 for 152,634 unique US charities and 14,372 unique UK charities, for which I present summary statistics by year and size quartile in the appendix. For my baseline analysis, I focus on the middle two quartiles by size, representing 68,345 US and 9,127 UK charities, out of which I filter the treated (environmental) and controls (arts)[3]. The resulting sample, for which I show summary statistics in Table 2.2, is used in my diff-in-diff regressions.

It is worth noting that the US dataset does not contain consistent financial data for years before 2013. Therefore, data for earlier years may be used for reference but not for analytical purposes, especially for the shock before 2013 in my study, i.e., the BP oil spill. The lack of data is also reflected the sharp increase in the number of US charities from 77 to 538 between 2011 and 2012 in Panel D of Table 2.2. Similarly, as of 2018, a small proportion of UK charities had submitted filings, and the coverage for that year is thin. Panel A of Table 2.2 shows that, on average, voluntary income received by UK charities accounts for around 30% to 38% of their total income; in the US (Table 2.2 Panel C), this ratio remains around 17-18% between 2013 and 2018 (including government grants).

2.4 Triggers to donation flows test results

The first shock I analyze is the Haiti earthquake shock in January 2010. Potential donors would have had enough time to react to this shock before the end of the tax year 2010. In Table 2.3, the post-shock years are 2010 and 2011, and for those years, I find positive but insignificant donations for charities with humanitarian or aid-related objectives relative to charities unrelated to humanitarian values. The result is robust to different specifications of control variables, and across charity size groups, the increase in voluntary income is insignificant. The Syrian refugee crisis intensified in 2015 and attracted elevated and sustained media attention through 2017. However, I expect donation towards relevant humanitarian-value or life-saving charities will increase more for the year 2015 than for the following two years due to this shock. This is due to the Brexit referendum taking place in the UK in 2016, which is likely to also have impact on individuals' willingness to spare income for donation as uncertainty spikes. Therefore in this analysis, I define $Post_t$ equal to 1 for the year

[3] The rationale for concentrating in the middle two quartiles is explained in Section 2.3.1

2015 only to avoid contaminating the treatment effect. As shown in columns (iii) and (iv) of Table 2.3 I do not find such a relationship through my diff-in-diff regressions. Overall, Table 2.3 below shows that humanitarian shocks seem to be a poor candidate for identifying individual social preferences.

Table 2.3 UK charity voluntary income flow – Humanitarian shocks

	Haiti Earthquake		Syrian Refugee Crisis	
	(i)	(ii)	(iii)	(iv)
$Post_t$	77.19	79.18	15.29	18.76
	(226.92)	(227.02)	(131.35)	(130.73)
$Treat_i$	250.53	240.23	-53.09	-45.36
	(230.21)	(230.66)	(128.64)	(128.04)
$Post_t \times Treat_i$	47.75	45.55	27.65	20.16
	(330.75)	(330.89)	(186.38)	(185.51)
$Total\_income_{i,t}$	0.28^{***}	0.27^{***}	0.22^{***}	0.21^{***}
	(0.02)	(0.02)	(0.01)	(0.01)
$Volunteers_{i,t}$	1.00	1.06	-0.27^{***}	-0.26^{***}
	(0.89)	(0.89)	(0.09)	(0.09)
$Fixed\_assets_{i,t}$		0.00		0.02^{***}
		(0.01)		(0.01)
Number of observations	424	424	1393	1393
Adjusted R-squared	0.424	0.423	0.253	0.260

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Haiti Earthquake shock: $Post_t = 1$ if year = 2010, 2011 and $Post_t = 0$ if year = 2009. Syrian Refugee Crisis shock: $Post_t = 1$ if year = 2015, and $Post_t = 0$ if year = 2014. My results are based on one-to-one propensity score matched treated and control charities using their pre-shock covariates. The treated and control charities are ranked between the 25th to 75th percentile of all UK charities as of pre-shock years 2009 and 2016. Definitions of control variables are available in the appendix.

Table 2.4 UK charity voluntary income flow – Environmental Shocks

	BP Oil Spill		US Paris Agreement Withdrawal	
	(i)	(ii)	(iii)	(iv)
$Post_t$	30.73	37.53	-158.02^{**}	-150.32^{**}
	(53.20)	(52.76)	(61.32)	(60.29)
$Treat_i$	2.34	10.38	80.89	105.14
	(91.15)	(90.37)	(95.46)	(93.95)
$Post_t \times Treat_i$	-53.43	-60.61	300.93^{**}	289.75^{**}
	(113.23)	(112.25)	(131.10)	(128.88)
$Total\_income_{i,t}$	0.37^{***}	0.38^{***}	0.62^{***}	0.64^{***}
	(0.03)	(0.03)	(0.02)	(0.02)
$Volunteers_{i,t}$	0.68^{***}	0.67^{***}	-0.93^{***}	-1.05^{***}
	(0.19)	(0.18)	(0.14)	(0.14)
$Fixed\_assets_{i,t}$		-0.01^{***}		-0.01^{***}
		(0.00)		(0.00)
Number of observations	831	831	739	739
Adjusted R-squared	0.215	0.229	0.551	0.566

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. BP oil spill shock: $Post_t = 1$ if year = 2010, 2011 and $Post_t = 0$ if year = 2009. US Paris Agreement withdrawal shock: $Post_t = 1$ if year = 2017, 2018, and $Post_t = 0$ if year = 2016. My results are based on one-to-one propensity score matched treated and control charities using their pre-shock covariates. The treated and control charities are ranked between the 25th to 75th percentile of all UK charities as of pre-shock years 2009 and 2016. Definitions of control variables are available in the appendix.

Table 2.4 shows that the BP oil spill did not influence UK donations. After the spill, BP agreed to administer a $20 billion response fund and suspend dividend payments, and the US federal government was also engaged in dealing with the aftermath of the disaster. Therefore, the pressure to help reverse the damage on non-profit organizations may have been weaker, which may explain the absence of significance in my baseline results. Furthermore, although the event received global attention, contamination due to the spill affected states in the US and the Gulf of Mexico rather than the UK; therefore, individuals in the UK may have felt less emotionally attached and may not have changed their donation behavior. As explained in the previous section, I could not conduct the same test for US charities due to lack of data.

The other environmental shock I examine, i.e., the Paris Agreement withdrawal, shows evidence of charitable flow reactions in the UK. Table 2.4 shows that after the announcement of the US withdrawing from the Paris Agreement, donors in the UK responded to the shock by donating £300,000 more to environmental and animal-focused charities than art and culture-related charities. This treatment effect is statistically significant even when tested using all size groups, including the smallest and largest 25 percent of charities (see appendix). In the US, there is also strong evidence for donations responding to the same shock (see Table 2.5). However, different from the UK, the size of this donation response is much smaller in absolute terms at $21,070, consistent with the smaller average size of US charities in the data. The relative sizes of the treatment effects for the UK and US are 46% and 14% of the average annual donations in each country (line one in Panels A and C of Table 2.2). Further, an alternative event date to the Paris Agreement withdrawal (i.e., the date of the announcement of the results of the 2016 US presidential election) leads to similar results.[4]

Table 2.5 US charity voluntary income flow

	2016 US presidential election		US Paris Agreement withdrawal		Tax Cut and Jobs Act	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	-23.52***	-22.87***	-14.43***	-14.24***	-2.20**	-1.80*
	(6.93)	(6.89)	(4.69)	(4.62)	(1.02)	(1.02)
$Treat_i$	18.88	18.82	6.01	10.71		
	(12.13)	(12.07)	(8.57)	(8.46)		
$Post_t \times Treat_i$	28.05*	28.24*	21.07*	21.37**		
	(15.41)	(15.32)	(10.82)	(10.66)		
$Total\_assets_{i,t}$	-0.02***	-0.02***	-0.01*	-0.01**	-0.00***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Total\_revenue_{i,t}$	0.49***	0.48***	0.72***	0.72***	0.58***	0.59***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.00)	(0.00)
Total functional	-0.25***	-0.22***	-0.32***	-0.27***	-0.24***	-0.21***
$expense_{i,t}$	(0.02)	(0.02)	(0.03)	(0.03)	(0.00)	(0.00)
Total government	0.14***	0.14***	0.11***	0.11***	0.09***	0.10***

[4] Note that because 2016 overlaps with the Brexit referendum in the UK, I do not test this alternative event date in the UK sample.

grants$_{i,t}$	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Fundraising expense$_{i,t}$	0.13	0.15	-0.02	0.00	0.22***	0.20***
	(0.11)	(0.11)	(0.11)	(0.10)	(0.02)	(0.02)
Payroll Tax$_{i,t}$		-1.72***		-2.26***		-1.69***
		(0.36)		(0.30)		(0.05)
Number of observations	1932	1932	1870	1870	101230	101230
Adjusted R-squared	0.618	0.622	0.685	0.694	0.458	0.463

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. US Paris Agreement withdrawal: Post$_t$ = 1 if year = 2017, 2018 and Post$_t$ = 0 if year = 2016. 2016 US presidential election shock: Post$_t$ = 1 if year = 2016, 2017 and Post$_t$ = 0 if year = 2015. Tax Cut and Jobs Act shock: Post$_t$ = 1 if year = 2018 and Post$_t$ = 0 if year = 2016, 2017; all variables are detrended and winsorized at top and bottom 2.5%. My results for the first two shocks are based on one-to-one, propensity score matched treated and control charities using their pre-shock covariates. The treated and control charities are ranked between the 25th to 75th percentile of all US charities as of pre-shock years.

Next, I consider the shock to the reputation of charitable institutions. Additional to using a humanitarian shock and a natural disaster to examine donation sensitivity, another useful exercise is to examine how donors react when there is a shock to charities' reputations. Donation is often made on the belief that the charity does good to society. If there is a scandal outbreak or an event that brings disrepute to the charity, it could undermine the credibility of charities, thus resulting in fewer donations.

Since the Oxfam scandal shock involves only a single institution, it was not possible to construct a treated sample of charities for a DiD analysis. Instead, I construct a time-series plot in Figure 2.1 and observe that the scandal is associated with a subsequent decrease in Oxfam's voluntary income in 2012. Another large charity, the British Red Cross, with similar objectives and size as Oxfam, also received less voluntary income in 2012. This indicates that donations that might be withdrawn from Oxfam did not move to a similarly large charity. Voluntary income of all other UK charities also decreased in the same year, which is in line with a generally negative effect in the whole sector. I argue that there was an overall decrease in donations in the charity sector after the scandal.

Figure 2.1 and 2.2 show the plot of total voluntary income received by different charities after the Oxfam 2011 scandal and the Kids Company 2013 scandal. In figure 2.1, the Oxfam misconduct scandal was followed by a subsequent decrease in its voluntary income in 2012. Another large charity, the British Red Cross, similar in objectives and size to Oxfam, witnessed a decrease in its voluntary income in 2012. The Salvation Army, however, has a smaller operation and more focused objectives. Therefore, it is used in the graph as a control for the other two.

In Figure 2.2, the scandal of misusing financial resources was reported on Kids

Company in 2013; the charity aims to support children and young people in the UK. The financial record for Kids Company is not available for years after 2013, but the voluntary income plot for another large children-focused charity, Bernado's, is shown on the graph. It appears that Bernado's was not affected substantially by the scandal of its peer. In the US, I conduct the DiD analysis using the Cancer Fund of America scandal as a shock affecting healthcare or medicine-focused charities; however, I do not find evidence that US donors changed their donation behavior after the shock.

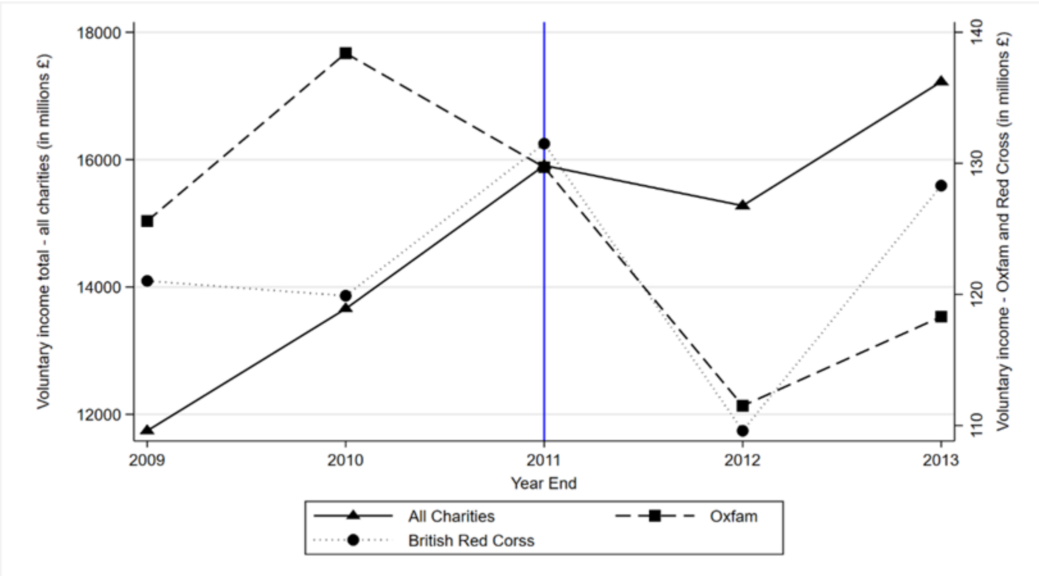


Figure 2.1 Total voluntary income received by Oxfam, British Red Cross, and all other UK charities after the 2011 Oxfam Scandal

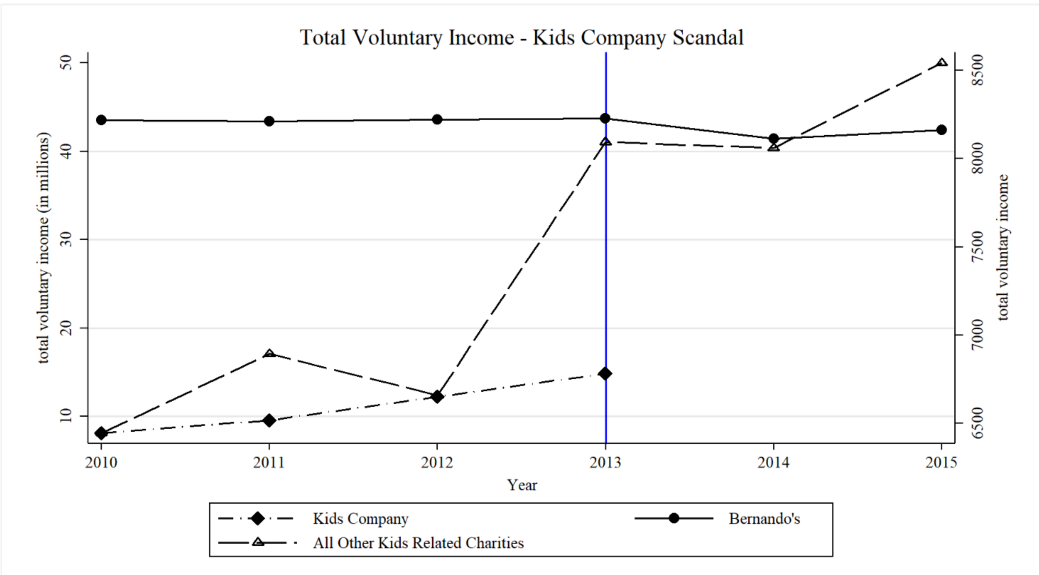


Figure 2.2 Total voluntary income received by three charities after the Kids Company Scandal in 2013

Finally, I assess the reaction of philanthropy to changes in tax exemptions following the US Tax Cut and Jobs Act 2017. As the policy affects the entire charity sector, it becomes impossible to separate charities into treatment and control groups based on their charitable purpose. Therefore, the analysis of this shock is approached from two directions. First, in Figure 2.3, I plot US individuals' annual adjusted gross income and their charitable contribution tax deductions over time using data from US Individual Income Tax Return Form 1040. I observe a clear drop in total charitable contributions claimed by taxpayers (in the form of itemized tax deductions). I further note that both the contribution by taxpayers in their annual tax filings (left-hand side axis) and the total adjusted gross income from returns filed (right-hand side axis) decreased dramatically after 2017. Second, using charity voluntary income data, I show consistent evidence that TCJA reduced donations received by US charities for the year 2018 when TCJA comes into effect. (see the last two columns of Table 2.5). Overall, Figure 2.3 and the last two columns of Table 2.5 illustrate that the TCJA had a clear negative impact on philanthropy: there was a sizable reduction in donations made by US individuals after its implementation.

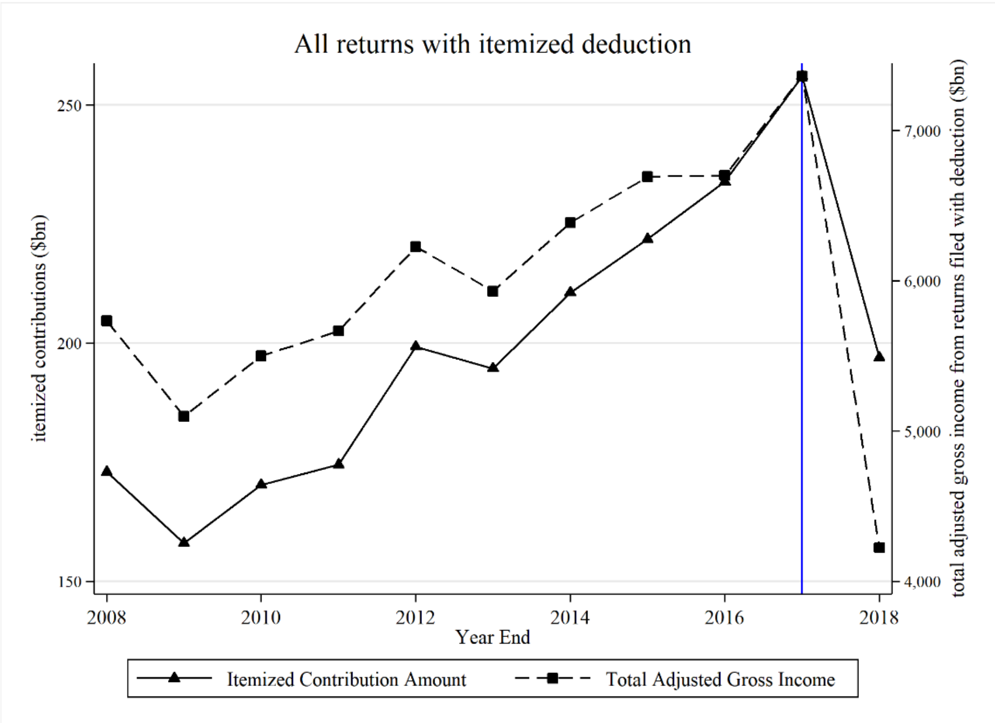


Figure 2.3 Itemized charitable contributions filed by US taxpayers over time

.

To summarize, I have shown that donation behavior reacts to environmental shocks even under my stringent specifications. In response to reputation scandals, affected charities lose donation income, which is not redirected to other charities, resulting in an overall donations decrease in the entire sector. Finally, I find that donations are strongly sensitive to changes in tax incentives.

From the results above, it is vital to conclude that charitable giving is not sticky and reacts to events such as a natural disaster and possibly a shock to a charity's reputation. Moreover, as the results from Table 2.4 and 2.5 show, charitable giving changes may still occur after a natural disaster without any tax policy changes.

2.5 Robustness and Additional Tests

Robustness

In my baseline models, treatment and control charities are implicitly matched by their purposes. However, I also conduct robustness checks on alternative treatment groups. For the UK refugee crisis shock 2015, I changed my definition to $Treat_i = 1$ to include charities that appeared as direct links on the Guardian News for people to access and contribute. The estimation results remain robust.

With regards to the two environmental shocks, I define environmental and animal-focused charities as treated in the baseline models. It may be possible that donating to local stray-dog, homeless pets, and animal charities do not necessarily imply strong concerns over global climate change issues or environmental disasters in general. As a result, animal-focused charities may dilute the treatment effect. To test this possibility, I consider an alternative definition for $Treat_i$ where I exclude animal-focused charities and include only environmental charities. After rerunning the diff-in-diff model for both US and UK, I do not find significance in either of the shocks (see Appendix). In addition, I conduct covariate matching on charities using charities' financial data; the baseline results remain largely unchanged (see Appendix).

Wealth effect on donations

As shown earlier, the cross-country DiD model did not pass the first stage shock strength test under the charity scandal shock and the natural disaster shock. However, there appears to be no significant change to measures of country-level donation and overall generosity. An intuitive explanation for donation being sticky is that it is dependent upon wealth. Wealthier individuals or countries would inherently donate a higher absolute amount than others; thus, unless there is a sudden shock to personal wealth, it is unlikely for individuals or households to change their donation behavior drastically.

Between 2006 and 2008, the US housing crisis significantly reduced the wealth of individuals in the US as their housing assets experienced substantial depreciation (IMF, 2008). By 2008, average US housing prices had declined by over 20% from their peak in mid-2006 (The Economist, 2008); as housing prices declined, homeowners’ equity continued to decrease towards zero or negative figures (Liebowitz, 2009). This crisis could be an exogenous and potentially informative shock to analyze how personal or household wealth changes affect donation.

Specifically, I show evidence on whether changes in the housing price index affect state-level donation. I first collect US Internal Revenue Service individual tax returns data where individuals disclose the deductions they claim for their charitable donations between 1997 and 2016 for all US states. In addition, Consumer Pricing Index (CPI) and state-level housing price index data are obtained from the United States Federal Reserve Bank (FRED) Economic Research Database. The S&P/Case-Shiller U.S. National Home Price Index (HPI) has been available monthly and seasonally adjusted since January 1987; CPI data is available annually from 1st Jan 1960. The base year is 2000 for HPI and 2015 for CPI. I adjust HPI data using CPI data to account for the different bases and inflation over time; donation deduction data at the state level is also inflation adjusted.

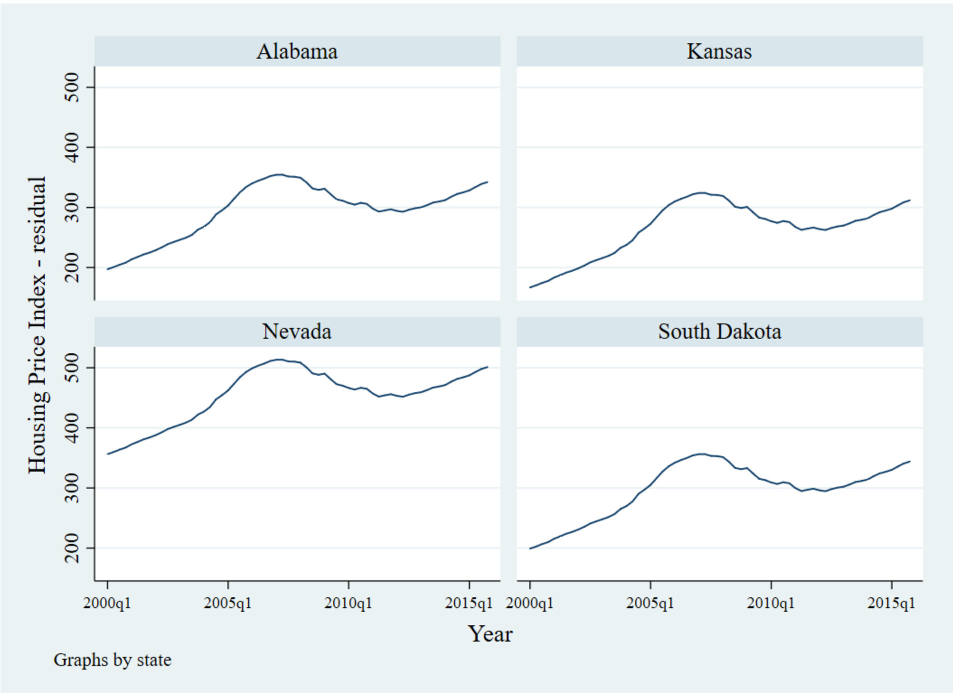


Figure 2.4 Housing Price Index (detrended) in four states with higher donations, 2000-2015

After removing time trends and state-specific effects, HPI shows a similar pattern across several states (see Figure 2.4 above). After a continuous rise since 1997, HPI experienced a decrease around 2006-2007 during the housing crisis. However, there does not appear to be a clear shock to wealth, and no significant difference in HPI trend over time between states with higher and lower donations. Thus, I will not take the 2006-2007 wealth shock further to analyze charitable giving changes.

Reputation effect

An alternative explanation to my shock strength test result, which shows charities receive more donations after the examined shocks, is a reputation effect. Instead of altruism, the desire for a better "reputation" could also drive more donations after salient shocks. To address this concern, I examine whether donation flows are driven by reputation instead of altruism around the US Paris Agreement withdrawal. Andreoni (1990) and Harbaugh (1998) show that a taste for prestige affects giving. The authors examined predicted donations made to a non-profit organization before and after it decided to report the names of donors publicly. My reputation test follows this spirit and studies anonymous versus named donations to a large charity: Oxfam America. Oxfam America publishes a donor list for the period 2015-2018, which reports the names of all donors, plus the number of anonymous donors in seven groups of annual donation amounts in US dollars (<5K; 5K-9,999; 10K-24,999; 25K-49,999; 50K-99,999; 100K-499,999; 0.5-1mil; >1mil). I count the number of named donors (foundations, estates, and other charitable organizations) and anonymous donors, then estimate the total donation amount for both donor types by multiplying the number of donors times the midpoint of the annual donation amount band. Then I run a standard difference-in-difference specification, with outcome variables: total donation amount or the number of donors; independent variables: dummy variable for named donors, dummy variable $Post_t$ equal to one for years 2017 and 2018, and zero for years 2015 and 2016; and their interaction.

Table 2.6 Test for Reputation Effect

	Total Donation Amount	Number of donors
$Post_t$	-1035.02	4.73
	(2278.71)	(37.99)
$Treat_i$	-3272.56	29.00
	(2449.54)	(39.42)
$Post_t \times Treat_i$	1671.17	62.29
	(3345.56)	(53.73)
Number of observations	56	52
R-squared	0.042	0.152

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. US Paris Agreement withdrawal shock: $Post_t = 1$ if year = 2017, 2018, and $Post_t = 0$ if year = 2016. Observations are group total amount and number of donors in each of seven groups by annual donation in US dollar amounts to the charity: Oxfam America (<5K; 5K-9,999; 10K-24,999; 25K-49,999; 50K-99,999; 100K-499,999; 0.5-1mil; >1mil). Treat = 1 for groups of named donors and 0 for anonymous donors. The number of observations differs because the number of donors for the first group (<5K) is unknown, and the total amount is deduced from the annual report minus the total of all other groups.

In Table 2.6, I show that having named donors does not suggest more donations or more donors, which indicates no reputation effect. It is worth noting that the above reputation test has limitations because the mid-point of the donation amount band approximates the donation amount by named and anonymous donors. One may find a reputation effect if the exact average contribution from named donors is higher than anonymous donors within the same donation amount band. However, to the best of my knowledge, Oxfam America does not provide data on exact amount contributed by its named, or anonymous donors, although such information would help bring more insight into the reputation effect hypothesis.

2.6 Conclusion

Charitable contributions are an important source for the provision of many public goods (Bernheim, 1986). In this chapter, I examine the mechanism and the extent to which charitable contributions are motivated by individuals' social preferences. Using granular US and UK charity data and sharp definitions of treatment and controls, I show that individuals demonstrate their social preferences after environmental catastrophes by increasing donations to environment-related charities compared to charities serving unrelated causes. However, UK and US donations to foreign aid-related charities remain unchanged after the humanitarian crises. In a different set of tests, I show that damage to the reputation of the charity sector and an increase in the tax price of donations are both negatively associated with private giving, which indicates the importance of charity brand image and the pecuniary benefit of giving in shaping donation decisions.

My findings contribute to the existing literature on social preferences and prosocial behavior in individual decision-making related to the allocation of financial resources. Most economic and psychological studies on social preferences, such as altruism and empathy, use lab experiments and incentivized games to gather data. In doing so, researchers face the hawthorn effect that participants without social preference also conform with the "desired" course of action when observed (reference). Therefore, a laboratory setting may not clearly distinguish between other-regarding preferences and other motivations (e.g., social image concerns) behind prosocial behaviors. My empirical design offers an alternative method to identify social preferences by using archival and anonymous charitable contribution data, which represent actual donation decisions made by individuals. In addition, to pinpoint social preferences, I compare shock-relevant and shock-irrelevant charities after the environmental and humanitarian crises, complemented by careful pre-shock covariate balancing to rule out alternative explanations of finding an effect. Nonetheless, I still explicitly test if social image or wealth signaling is one alternative explanation to my result and find that such motivation is unlikely to be present for one of my shocks.

In this chapter, two of the shocks I examine: the Haiti earthquake of 2010 and the Refugee crisis of 2015, do not trigger significant effects on charitable flows as environmental shocks do. My results from the Haiti earthquake test differ from Hickey et al. (2019), who find inflows to Canadian charities after the disaster. However, the authors' identification strategy is based on a shock to local tax deduction policy. Therefore, finding significant inflows implies that monetary incentives alter private giving outcomes. My finding of no effect in the absence of such policy change does not contradict this conclusion.

When examining the effect of the US Tax Cut and Jobs Act, I notice reduced giving to US charities when the tax shield on giving is partly lifted. In addition to finding tax policy relevant to prosocial decisions, I also illustrate using the US Paris Agreement withdrawal shock to illustrate how a lack of commitment to government environmental policies can induce more private provisions of environment-related public goods. This result is robust to controlling for fundraising expenses made by charities and thus rules out the possibility of finding significance due to additional charities' fundraising efforts in response to policy changes. I also use named versus anonymous donation to confirm that my results in the US Paris Agreement withdrawal shock are not driven by social image or prestige signaling effects.

My finding of insignificance in Haiti and the Refugee shock does not strictly rule out the possibility of social preferences. A possible explanation could arise from the literature on the heterogeneity of empathy (Hichri & Kirman, 2007; Singer et al., 2006; Kirman & Teschl, 2010) and studies on how identifiability of the victim affects prosocial behaviors (Burgoyne et al., 2005; Jennie & Loewenstein, 1997; Andreoni, 2006). Results from this chapter suggest that individuals with social preferences act more heterogeneously, leading to insignificant donation changes after the humanitarian crisis than the environmental crisis. This could potentially be due to the more distant nature (in Haiti and Syria instead of the Gulf of Mexico and the United States) of the earthquake and the refugee crisis.

In addition, I offer evidence from the Oxfam scandal that the reputation incident of a single charity may lead to reduced contributions to other similarly purposed charities, and consistent with the argument in LeClair (2019) that non-profits also experience reputation damage spill-over as a whole sector in the same way as for-profit organizations would.

Appendix 2.1 Covariate balancing and Parallel Trends – UK

Table A2.1.1 BP oil spill 2010 – covariate balancing test – UK charities

	Lagged Fixed Assets$_{i,t}$	Lagged Total Income$_{i,t}$	Lagged Charitable Expenditure$_{i,t}$	Lagged Number of Employees$_{i,t}$	Lagged Number of Volunteers$_{i,t}$
Panel A. Before Matching					
Mean-Test					
Control	7,048.38	3,598.14	3,013.84	46.00	36.91
Treated	10,486.03	4,969.96	3,578.60	75.81	626.41
t-stat	-1.03	-1.07	-0.53	-1.67	-2.23
Median-Test					
Control	465.03	1,398.55	1,156.96	18.00	0.00
Treated	2,209.46	1,300.93	877.23	23.00	5.00
z-score	-5.92	0.98	3.26	-1.15	-4.60
Number of control charities	321	321	321	321	321
Number of treated charities	327	327	327	327	327
Panel B. After Matching					
Mean-Test					
Control	7,996.23	2,446.29	1,920.08	43.29	63.73
Treated	6,887.85	2,531.06	1,839.69	41.36	60.00
t-statistic	0.42	-0.26	0.32	0.36	0.39
Median-Test					
Control	699.32	1,518.96	1,132.32	25.00	12.00
Treated	1,897.93	1,132.85	764.17	20.00	2.50
z-score	-5.43	1.49	3.63	-0.07	-3.30
Number of control charities	286	286	286	286	286
Number of treated charities	286	286	286	286	286

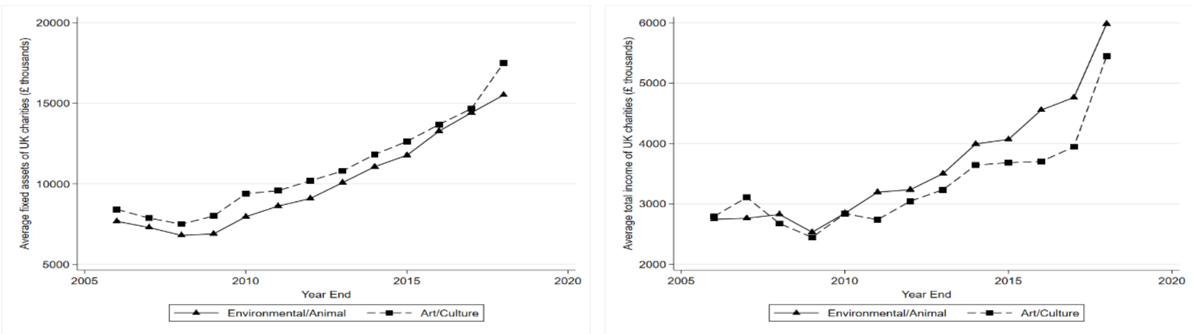
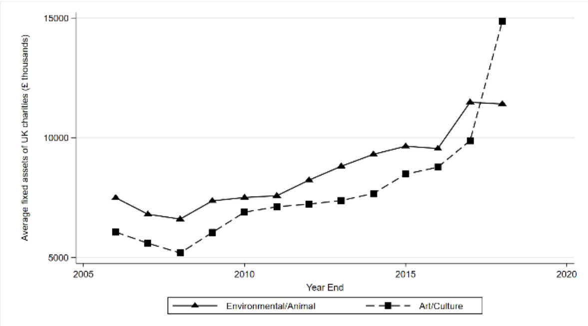


Figure A2.1 Parallel trends of UK treatment and control charities – BP oil spill 2010

Table A2.1.2 US Paris Agreement withdrawal shock – covariate balancing test – UK charities

	Lagged Fixed Assets$_{i,t}$	Lagged Total Income$_{i,t}$	Lagged Charitable Expenditure$_{i,t}$	Lagged Number of Employees$_{i,t}$	Lagged Number of Volunteers$_{i,t}$
Panel A. Before Matching					
Mean-Test					
Control	8,659.08	4,492.81	3,539.03	47.80	65.41
Treated	15,283.22	6,385.98	4,839.32	81.52	731.13
t-stat	-1.68	-1.23	-1.03	-1.93	-3.14
Median-Test					
Control	656.33	1,601.74	1,361.64	17.00	8.00
Treated	3,170.96	1,360.49	994.41	24.00	80.00
z-score	-6.79	1.15	3.03	-2.47	-10.15
Number of control charities	443	443	443	443	443
Number of treated charities	440	440	440	440	440

Panel B. After Matching					
Mean-Test					
Control	8,772.05	3,659.67	2,669.63	50.96	128.44
Treated	9,549.94	3,413.54	2,475.40	46.53	128.27
t-statistic	-0.33	0.47	0.50	0.72	0.01
Median-Test					
Control	1,396.26	2,029.16	1,745.11	23.00	54.00
Treated	2,562.76	1,258.72	925.73	22.00	60.00
z-score	-6.26	2.10	4.09	-0.53	-7.09
Number of control charities	394	394	394	394	394
Number of treated charities	394	394	394	394	394



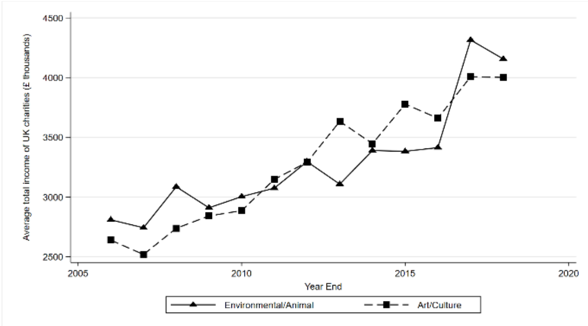


Figure A2.2 Parallel trends of UK treatment and control charities – US Paris Agreement withdrawal shock 2017

Appendix 2.3 Covariate Balancing and Parallel Trends – US

Table A2.3.1 Election Shock 2016 – covariate balancing test – US charities

	Lagged Total Assets$_{i,t}$	Lagged Total Revenue$_{i,t}$	Lagged Functional Expenses$_{i,t}$	Lagged Government Grants$_{i,t}$	Lagged Payroll Tax$_{i,t}$	Lagged Direct Fundraising Expense$_{i,t}$
Panel A. Before Matching						
Mean Test						
Control	337.93	190.30	181.44	866.93	6.64	12.24
Treated	249.20	206.87	190.48	1,016.89	10.20	7.70
t-stat	1.43	-0.79	-0.45	-1.12	-2.39	2.26
Median Test						
Control	45.27	71.72	68.62	369.14	0.00	0.00
Treated	49.92	86.44	80.95	522.98	0.00	0.00
z-score	0.40	-5.15	-4.90	-4.21	-1.77	-1.26
Number of control charities	1,317	1,317	1,317	480	480	480
Number of treated charities	3,005	3,005	3,005	1,080	1,080	1,080
Panel B. After Matching						
Mean Test						
Control	132.56	222.92	198.79	461.62	0.00	0.00
Treated	156.85	230.72	204.40	497.24	0.00	0.00
t-statistic	6.63	3.22	3.11	-3.53	-0.95	-4.68
Median Test						
Control	414.90	299.58	266.25	737.00	5.54	6.34
Treated	377.45	316.03	287.43	743.30	5.58	5.92
z-score	1.09	-1.15	-1.61	-0.17	-0.11	0.47
Number of control charities	1,016	1,016	1,016	1,016	1,016	1,016
Number of treated charities	1,016	1,016	1,016	1,016	1,016	1,016

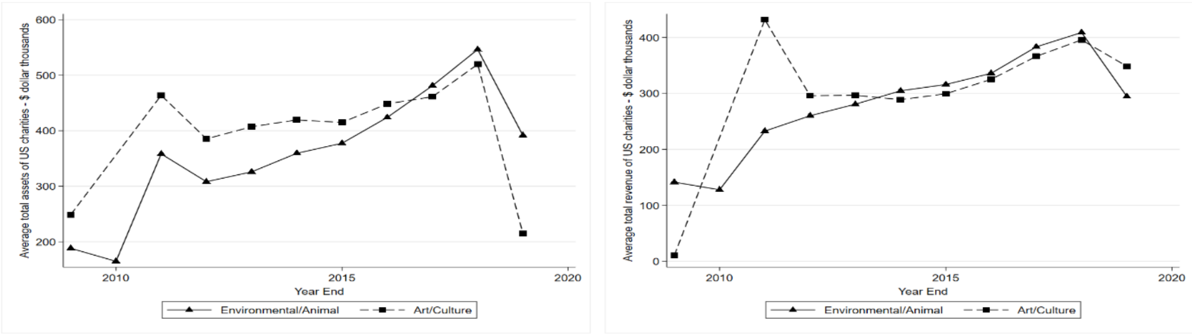


Figure A2.3.1 Parallel trends of US treatment and control charities – Election shock 2016

Table A2.3.2 US Paris Agreement withdrawal shock 2016 – covariate balancing test – US charities

	Lagged Total Assets$_{i,t}$	Lagged Total Revenue$_{i,t}$	Lagged Functional Expenses$_{i,t}$	Lagged Government Grants$_{i,t}$	Lagged Payroll Tax$_{i,t}$	Lagged Direct Fundraising Expense$_{i,t}$
Panel A. Before Matching						
Mean Test						
Control	308.37	182.79	173.47	883.72	7.20	10.73
Treated	275.83	212.25	195.93	998.17	10.57	8.33
t-stat	0.61	-1.64	-1.27	-1.22	-2.22	1.31
Median Test						
Control	43.56	70.64	66.37	395.99	0.00	0.00
Treated	51.85	88.22	82.63	548.35	0.00	0.00
z-score	-1.79	-7.02	-6.62	-4.52	-3.01	-3.09
Number of control charities	1,487	1,487	1,487	540	540	540
Number of treated charities	3,147	3,147	3,147	1,157	1,157	1,157
Panel B. After Matching						
Mean Test						
Control	441.01	265.09	248.51	623.07	3.11	5.06
Treated	399.49	268.27	243.01	667.99	2.98	7.16
t-statistic	0.85	-0.19	0.33	-1.20	0.49	-2.23
Median Test						
Control	123.65	205.10	180.54	369.25	0.00	0.00
Treated	146.79	212.01	187.71	488.40	0.00	0.00
z-score	2.52	0.24	0.39	-6.60	-5.24	-6.02
Number of control charities	951	951	951	951	951	951
Number of treated charities	951	951	951	951	951	951

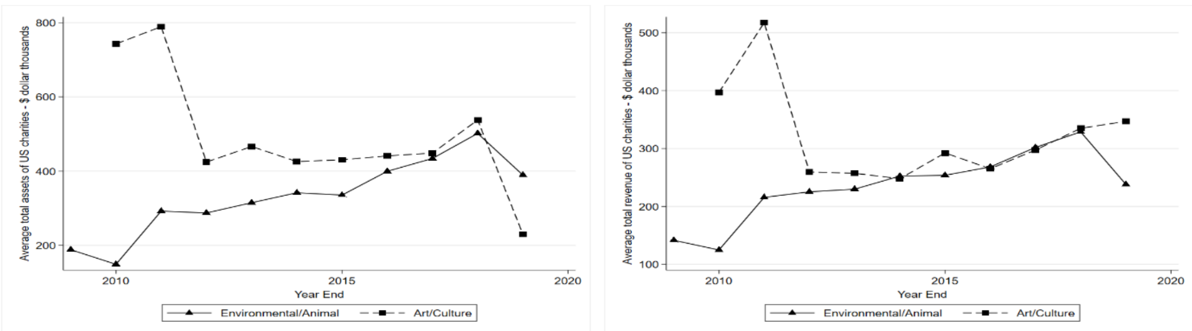


Figure A2.3.2 Parallel trends of US treatment and control charities – US Paris Agreement withdrawal shock 2016

Chapter 3 Is sustainable investing driven by social preferences?

3.1 Introduction

Sustainable investing (SI) has grown enormously over the last decade: global assets under management (AUM) reported by the signatories of the Principles of Responsible Investing stood at $103.4 trillion as of March 2020 – a 16-fold increase from $6.5 trillion in 2006 (PRI, 2020). The portion of capital invested in SI funds reached a record 1.7 trillion in 2020, marking a 50 percent growth over just one year (Mooney & Mathurin, 2021). Recently, investors further demonstrate their preference for sustainability in an unprecedentedly successful proxy challenge at Exxon (Baer & Lim, 2021) and in overwhelming support for climate strategy resolutions at Shell (Bousso, 2021) and Unilever (Cavale, 2021).

While there is some suggestive evidence that the interest in sustainability is at least in part driven by social preferences such as altruistic motives (Hartzmark & Sussman, 2019; Barber et al., 2021; Riedl & Smeets, 2017), surveys among SI fund managers show a wide range of estimates for the importance of social preferences in SI investing. For example, 77% of respondents say social or moral considerations drive them in a Merrill Lynch study, while 27% choose altruistic values in a BNP study (Boffo & Patalano, 2020). In this Chapter, I examine whether social preferences drive SI flows by analyzing market-wide equity mutual fund investment flow data. My approach allows us to mitigate the attitude-behavior gap typically found in surveys, which is especially severe in social marketing (De Pelsmacker et al., 2005).

It is essential to empirically establish which motives dominate overall market movements because successful policy interventions, necessitated by climate change and post-Covid global challenges, must be attuned to investor preferences. I show that SI is not likely driven by social preferences in market-wide data of equity mutual funds in the US and the UK.

To identify the role of social preferences in SI investing, I have shown in Chapter 2 that philanthropy, which is measured as charitable donations and may be driven by various aspects of social preferences such as altruism, does react to shocks that activate social preferences (for example a natural disaster shock, severe environmental damage due to corporate actions). In this chapter, I aim to pinpoint whether the same response is realised in market-wide flow to SI. In the case of SI flow not reacting, one may argue that SI investors with social preference simply do not consider these shocks to be relevant for SI. As a result, I assess a second set of shocks that holds social preferences constant. These shocks negatively affect charity sector reputation or reduce philanthropy attractiveness through tax policy change. As shown in Chapter 2, flow to charities decreased, and I show in this chapter that SI flow remains unaffected.

This chapter is structured as follows, Section 3.2 introduces the concept of ESG and recent trends in sustainable investing and discusses existing studies on investor demand for investing in SI and their social preferences. Section 3.3 illustrates the difference-in-differences methodology and mutual fund data used to examine investor social preferences for SI investing. My main results are presented and discussed in Section 3.4, where I also show additional robustness tests.

3.2 Related Literature

In this chapter, I define social preferences in the investment decision-making context as supporting environmental and/or social and/or governance (ESG) causes, even if this means sacrificing returns (LeClair, 2014).

The terminology around ESG, SI, and corporate social responsibility (CSR) has been evolving over the past few decades, and awareness of CSR has begun in at least the 1970s (Ullman, 1985). Previous research made an extensive effort to scrutinize whether better CSR performance translates into financial outperformance at the firm level (for example, Cochran & Wood., 1984; Wood & Jones., 1995; Orlitzky et al., 2003; Luo & Bhattacharya, 2006; Guenster et al., 2011; Krueger, 2015). These studies suggest that CSR may be associated with reduced cost of capital and decreased systematic risk for companies (for example, Orlitzky et al., 2003; Albuquerque et al., 2019).

In recent years, the topic of CSR, ESG, and sustainability in business and finance has become increasingly relevant and of interest to investors. Increasing social awareness of climate change, pollution, and inequality has materialized in a series of behavioral shifts among consumers and investors, for example: opting for fair-trade coffee and organic foods, traveling by train or purchasing CO2 offsets when flying, and moving investments from conventional to SI instruments. Pro-SI regulatory measures that mandate incorporating ESG aspects in corporate and investment decision-making (Eccles & Klimenko, 2019) also fuelled greater availability of SI investment instruments and products. According to the Global Sustainable Investment Alliance (GSIA) and the Principles of Responsible Investing (PRI), the size of assets under SI management has grown two times over seven years, from $13.3 trillion in 2011 to $30.7 trillion at the start of 2018 (GSIA, 2018) and twelve times from $6.5 trillion in 2006 to $103.4 trillion at the start of 2020 (PRI, 2020). GSIA (2018) also reports that client demand is one of the leading reasons for the emergence of more SI products. Initially, strategies used by SI funds were predominantly based on exclusionary screening, where fund managers exclude certain “sin” stocks, such as those in the tobacco and alcohol industry. However, the industry has witnessed significant growth in the use of alternative SI strategies, such as ESG integration strategies, best-in-class screening, and impact investing. The value of assets managed under the three strategies grew by 69%, 125%, and 79%, respectively, between 2016 and 2018 (GSIA, 2018).

Whilst SI instruments are often marketed as a superior investment option in terms of risk/return, most theoretical (Heinkel et al., 2001; Baker et al., 2022) and empirical studies (Geczy et al., 2021; Renneboog et al., 2008; Barber et al., 2019) conclude that socially-driven investors accept SI instruments performing at a discount compared to non-SI, or conventional investment instruments. Flows toward SI mutual funds are also less sensitive to poor past returns compared to flows toward conventional, non-SI funds (Bollen, 2007; Renneboog et al., 2011). These studies suggest additional non-pecuniary motives behind SI investment but do not attempt to identify them.

For investors who see SI primarily as a profitable long-term investment strategy, social preferences such as altruism are a secondary motivation (20% in the Merrill Lynch survey and 52% in the BNP survey choose a risk-return motivation as a primary driver). In addition, there is some empirical evidence that investment strategies based on certain aspects of ESG, such as high employee satisfaction (Edmans, 2011), can improve investment returns and that institutional investors with better sustainability performance also hold longer investment horizons, which generates higher risk-adjusted performance (Gibson et al., 2021). Such findings are consistent with initial mispricing of ESG-related characteristics of the firm and expected future performance by investors, but not social preferences.

Fama and French (2007) compare socially responsible assets to consumption bundles satisfying both monetary motives and investor tastes unrelated to returns. Recent theoretical work studies investor preferences for sustainable as opposed to conventional assets and derives market equilibrium results with a new sustainable priced factor (or an ESG factor) and lower expected returns for sustainable assets (or "green" assets) (Heinkel et al., 2001; Baker et al., 2022; Pástor et al., 2021; Pedersen et al., 2021; Bauer et al., 2021; Angelis et al., 2020). These models derive optimal portfolio allocations with tilts towards SI depending on the strength of ESG taste. In this chapter, I attempt to isolate social preferences as one of the drivers behind ESG tastes by focusing on unexpected shocks to investors' awareness and sensitivity to ESG issues that manifest in increased philanthropy flows.

In the context of consumer choice, many theoretical models consider a product with a sustainability component equivalent to a regular product plus a charitable donation, such as fair-trade coffee (Dragusanu et al., 2014) or carbon offset programs (Marron & Toder, 2014). In a Modigliani-Miller framework, Zivin and Small (2005) show that to satisfy their mix of monetary and social preferences, investors will be indifferent between a combination of conventional investments and philanthropy on the one hand (Friedman, 1970), and socially responsible assets (Vogel, 2006), on the other. One can see a share in an SI fund as a charity-investment bundle by extending their paradigm from firms to portfolios. However, if substitutability is imperfect because purely profit-maximizing corporate activities are inseparable from imposing negative social externalities, then socially conscious investors would lean in favor of SI funds (Hart & Zingales, 2017). On the other hand, if some investors are skeptical about the ability of corporations to deliver altruistic goals credibly, they would prefer the conventional-philanthropy bundle (Kotchen, 2006). The presence of socially responsible investors with heterogeneous preferences means that a sudden increase in the desire to contribute to a cause, for example, environmental protection, would be reflected in both environmentally oriented philanthropy and SI because they are alternative vehicles (even if not equivalent) in achieving pro-social objectives.

However, these studies do not specify the source of investor preferences for sustainability. Recent studies that attempt to isolate social preferences are based on an experiment with MBA respondents (Hartzmark & Sussman, 2019), survey data (Bauer et al., 2021), and surveys combined with experimental evidence from the

Netherlands (Riedl & Smeets, 2017). In Hartzmark and Sussman (2019), the authors design an experiment with MBA students to capture their preference over ESG to explain their findings that funds rated as high sustainability by Morningstar received net inflows of over $24 billion, while those rated as low sustainability received net outflows of over $12 billion. The authors find that a significant difference in flows between SI and conventional funds is only observed among participants who consider ESG factors when making investment decisions and conclude that investing in sustainability is driven by social preferences, which may be altruistic or other-nonpecuniary motives. In this chapter, I focus on social preferences in a non-experimental setting and the broader investor universe, whereby I use a set of exogenous shocks and study their effect on SI flows.

Overall, the literature suggests that SI investors allocate their money to sustainable mutual funds for nonpecuniary reasons, which can be considered as social preferences such as altruism, warm-glow, or making a social impact. If social preferences are the driver behind SI, then shocks that affect philanthropy may also drive changes in flows toward SI.

A lack of significance in SI fund flow after shocks to social preference may be related to institutional and informational differences between philanthropy and SI. One might argue that pension funds and regulated institutional investors are important potential contributors to SI investing (Eurosif, 2016) and make investment decisions differently from contributors to philanthropy, who are mostly individuals. There is little overlap between the two groups of investors. Firstly, I note that during the period of my analysis, these large regulated institutions both in the US and UK had to be careful not to sacrifice returns for their investors in the name of ESG considerations (there are no ESG funds available in the vast majority of 401(K) retirement plans in the period of my analysis (Quinson, 2020). Therefore, the investors holding SI funds that I analyze are less likely to be such regulated institutional investors. Still, I include interactions with retail versus institutional indicator variable to address this concern. Second, the US Federal Reserve Survey of Consumer Finance shows that the value of employer-sponsored pension savings (which are subject to a restriction on using ESG criteria) is minuscule relative to overall investments in equity funds and stocks, where investors have discretion over the nature of the characteristics of their chosen instruments. In the UK, however, institutional investors have been the leading force behind SI, and I include interactions with an institutional dummy variable to address this potential concern.

Another difference between philanthropy and SI is that the former is often done in an episodic or one-time manner (CAF, 2019), while SI could be more akin to regular pension saving. When analyzing SI flows, I address this concern by performing robustness tests with different numbers of post-shock years. Alternatively, one may be concerned that the lack of effect on SI flows could be due to the annual frequency of my baseline analyses, whereas an effect might exist on a shorter-term scale. Again, I show that this is not the case using a daily frequency event study.

Visibility and search costs may also differ between SI and philanthropy. While charitable giving is often encouraged by well-publicized appeals (Bendapudi et al., 1996), SI funds prospectuses may be less accessible to investors. For example, it may be less costly and time-consuming for individuals to search for charities with environmental protection goals through news agencies, charity websites, or word of mouth than to search for sustainable investment funds with an environment focus through a proprietary database, financial advisors, or individual fund websites. Search costs are relevant to mutual fund investor decisions (Hortaçsu & Syverson, 2004; Roussanov et al., 2021), and therefore, fund flows (Sirri & Tufano, 1998). I minimize this concern by performing tests both before and after a sudden increase in the visibility of SI funds. The Morningstar database includes an SI classification for the entire period of my analysis. However, the designation of sustainability became especially detailed at the start of 2016 with the introduction of their globes ratings, which resulted in a surge in SI flows, as shown by Hartzmark and Sussman (2019). By including two shocks that take place after the end of 2016, I confirm that my results are unchanged during this recent period of higher visibility for SI resulting from the issuance of Morningstar globe ratings.

Overall, unlike the charity flows effect found in Chapter 2, SI fund flows in all my tests never show a significant reaction to the environmental shocks I examined. Thus I argue that social preferences are unlikely to be a strong determinant of SI flows. Results from my second set of shocks based on the US Tax Cut and Jobs Act and a charity scandal in the UK strengthen my conclusion as donations are not reallocated to SI. My findings suggest that it is unlikely for social preferences to be the primary determinant for SI fund flows.

My work contributes to existing research on investor preferences for SI. Classic asset pricing theory extended to incorporate taste-based factors (Fama & French, 2007) or “green” preferences (Heinkel et al., 2001; Baker et al., 2022; Pástor et al., 2021; Pedersen et al., 2021; Hong et al., 2020) predicts that SI should earn lower returns because they satisfy an additional non-monetary preference of investors. I show that altruism, as an obvious candidate for a non-monetary motive, is likely not an instrumental driver of SI fund flows, at least in the data from 2008 to 2018. If SI flows differ from conventional flows, it is for reasons other than whether they are motivated by altruism. This finding is consistent with Ramelli, Wagner et al. (2021), who show that, following the 2016 US presidential election, if investors rewarded climate-responsible firms, it was more likely for strategic reasons rather than concerns over global warming. This result can inform the policy debate around reversing the restriction on regulated investment institutions incorporating ESG factors at a financial cost to their members, introduced by the US Department of Labor in October 2020. Since most other motives behind philanthropy and SI, except altruism, can be linked to a financial benefit, the lack of effect of altruism means that financial considerations must already drive the investment decision-making process, and these restrictions may be unnecessary.

Furthermore, suppose policymakers consider introducing higher ESG mandates following Hong et al. (2020) recommendations and want to garner greater political and societal support for them. In that case, my results suggest that appealing to altruism may not be as productive in inducing redirection in investment flows as risk/return-based incentives.

Some SI skeptics question whether SI is a fundamentally different type of investment compared to a conventional portfolio, as opposed to merely a marketing ploy and a well-sounding label (Agnew et al., 2022). These doubts are justified by evidence that most widely used ESG ratings are inconsistent (Gibson et al., 2022), and some have been adjusted ex-post to predict future returns better (Berg et al., 2021). However, recent regulatory initiatives in the EU, US, and UK towards greater transparency and compliance on ESG dimensions will likely result in better discernment between SI and conventional instruments. In this case, SI may show worse risk-return performance in the short term. My findings suggest that many existing SI investors may abandon these assets because their main motivation is not pro-social. This possibility mirrors the augment that the business case for board gender diversity creates false expectations of women directors (Adams, 2016).

3.3 Empirical Analysis

3.3.1 Survey Data of investors and donors in the US

Before I move to my main analysis using aggregated market-wide data, I examine an individual-level survey that reveals investor preferences and how they manifest in savings and investment decisions. The US Federal Reserve Survey of Consumer Finance (SCF) is a representative triennial cross-sectional survey of U.S. families. It interviews heads of households and collects information on their balance sheets, pensions, income, and demographic characteristics. Notably, the SCF reveals the overlap between the individuals who invest and who give to charity because it responds to the same participants regarding owning stocks and/or mutual funds and donating. Thus, I observe the distribution of preferences for different kinds of investments and philanthropy in the population and can deduce whether my hypotheses about the complementarity and substitutability between investments and philanthropy are supported.

I classify as investors in my context, respondents who hold stocks either directly or through pension funds or mutual funds. The SCF provides the proportions of different instruments that are in stocks, allowing us to compute their corresponding dollar amounts.

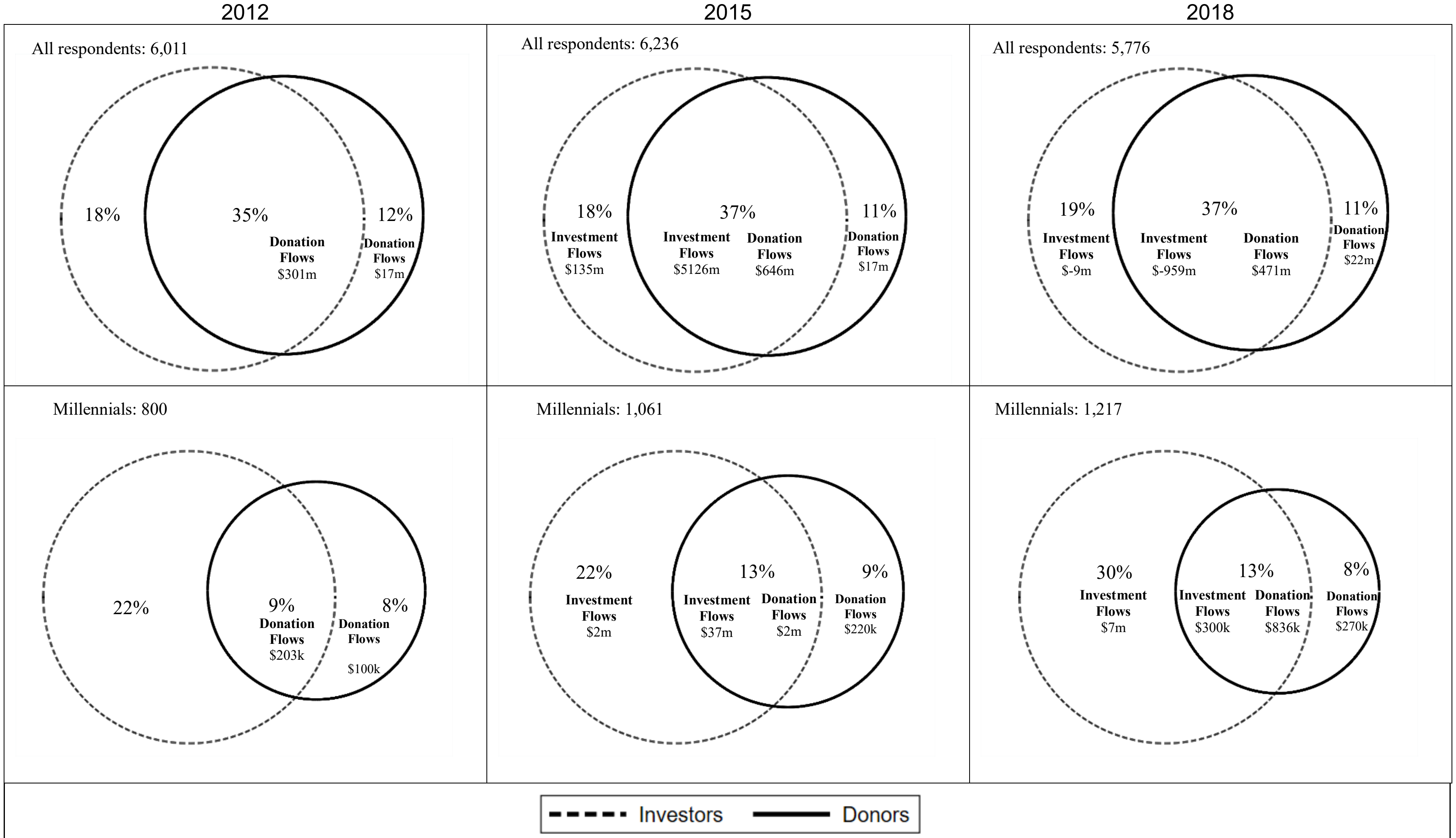


Figure 3.1 Proportion of respondents who invest and donate among all US families from the Survey of Consumer Finance.

The graph shows the proportions of investors and donors among the respondents of the Survey of Consumer Finance (SCF) executed by the Federal Reserve every three years among US households. The intersection between the dashed and solid circles represents those who invest in the stock market and donate to charities. Inside the circles, I report the dollar amounts of flows to charity and investment instruments. The investment flows are missing in the year 2012 because they are computed as changes between two consecutive survey years. Source: US Survey of Consumer Finance

Figure 3.1 above illustrates the proportions of investors, donors and their overlap in three survey years. The top row shows that the proportion of respondents who both invest and donate is the highest at 35%-37%, while the groups who only invest (18%-19%) or only donate (11%-12%) are much smaller. The SCF does not distinguish SI among all the savings and investment instruments respondents report. SI can be part of both groups: invest-only and invest-and-donate.

On the one hand, respondents who hold SI and do not donate may likely see SI as a better alternative to the bundle of conventional investing and philanthropy. On the other hand, respondents who hold SI and donate simultaneously may not see SI as a perfect substitute for philanthropy. I obtain an indication about the former group by focusing on the subset of respondents who have been repeatedly shown to be most interested in SI – the Millennials (among many others, see Morgan Stanley Institute for Sustainable Investing, 2020). On the bottom row of Figure 3.1, I note that a larger proportion of Millennial respondents belong to the invest-only group (22% going up to 30% in 2018). This is suggestive evidence consistent with my hypothesis that some investors see SI as an alternative to the bundle of conventional investing and philanthropy.

In Chapter 2, I show a significant shift in philanthropy in response to some of the shocks I consider. To link donation flows shift to a corresponding shift in investments, it is necessary for the donors behind it also to be investors. I see this in the total dollar amounts reported below the percentages of respondents in the three groups in Figure 3.1. The donation flows in the invest-and-donate group are 18 to 37 times higher than those in the donate-only group. I see a similar result for Millennials in the bottom row, where the dollar amounts of donations inside the invest-and-donate bundle are 2 to 9 times higher than those in the donate-only group. This reassures us that most of the donation dollar amounts indeed come from investors.

Table 3.1 US Survey of Consumer Finance - Respondent Investment and Donation Behaviour

	Survey year	Millennials	Gen X	Baby	Golden Gen
INVESTOR ONLY					
Investment	2012	3.7	22	23	83
	2015	5.7	40	40	179
	2018	12.6	59	73	106
	Δ flows '15 to '18	**5.0**	**1**	**16**	**-169**
INVESTOR AND DONOR					
Investment	2012	5.4	1,385	1,814	7,542
	2015	42.7	1,206	3,208	11,378
	2018	43.0	2,147	3,133	9,590

	Δ flows '15 to '18	**-37.0**	**1,119**	**-1,469**	**-5,624**
Donation flows	2012	0.2	17	53	231
	2015	1.8	22	102	520
	2018	0.8	38	79	353
	Δ flows '15 to '18	**-1.0**	**16**	**-23**	**-167**
		DONOR ONLY			
Donation flows	2012	0.1	1.2	12.7	2.9
	2015	0.2	0.7	1.3	15.1
	2018	0.3	1.4	3.7	16.1
	Δ flows '15 to '18	**0.1**	**0.7**	**2.4**	**1.0**

This table presents the number of respondents (Panel A), investment and donation amounts (Panel B), and proportions (Panel C) by age group as reported in the Survey of Consumer Finance (SCF), executed by the Federal Reserve every three years among US households. The generation groups are constructed based on respondent birth year: Millennials (1981 to 1996); Gen X (1980 to 1965); Baby Boomers (1956 to 1964); and Golden Generation (1928 to 1945). To obtain the amounts in Panel B, I add up the total dollar value of all investments by household. To obtain the donation amount for each respondent, I add the value of the contribution (X5823) with the value of the charitable trust or foundation (X7661). A respondent is in the Invest&Donate group if he/she is both an investor and a donor as defined above; the Invest (Donate) Only group includes those who are investors (donors) but not donors (investors).

In Table 3.1, I report the total dollar amounts in the three groups: invest-only, invest-and-donate, and donate-only, split by the four commonly defined generations: Millennials (born after and in 1981); Generation X (born from 1980 to 1965); Baby Boomers (born from 1956 to 1964); and the Golden Generation (born before or in 1945). Unlike my main tests in the next section, where I have annual data and can better isolate changes around particular shocks, SCF data only allows us to observe flow changes between every two survey years that are three years apart. The shifts I find in the SCF data cannot be reliably attributed to a given shock of my study but are useful as preliminary evidence. In bold, I report the changes in flows between 2015 and 2018 for the different generations. There are many sizable shifts in and out of the different age and invest-donate groups—notably, Millennials and Baby Boomers exhibit outflows from the invest-and-donate bundle and inflows to the invest-only group. Moving to aggregate market data with annual frequency in the following section, I try to pinpoint these shifts to particular shocks to social preferences.

3.3.2 Data and Summary Statistics

I construct samples of UK and US open-end equity funds from Morningstar Direct. It classifies funds into Sustainable Investing or not and offers further details on the sustainable investing focus based on the fund prospectus. For example, the category "Sustainable Investment – Environmental" defines environmentally focused SI funds. Following Renneboog et al. (2008, 2011), Pástor et al. (2021), and Pástor and Vorsatz (2020), my sample excludes index funds, bond funds, money market funds, real estate funds, and funds with less than 75% allocation to equity. I exclude international and non-domestic funds from the US sample. Since the UK was still part of the European Union at the end of my sampling period (2018), I include UK-domiciled funds with a European focus and UK domestic equity funds. I also obtain fund financial data, including annual net flow, return, expense ratio, and inception date.[5]

Table 3.2 describes the fund data used in this Chapter. The full sample contains 1,013 UK SI funds (out of which Morningstar identifies 568 as environment-focused) and 7,639 UK conventional funds. My US sample contains 404 SI funds, of which 246 had an environmental focus and 12,715 US conventional funds. Fund age in Table 3.2 is calculated as of year-end 2018; SI funds tend to be younger than conventional ones in both the US and the UK. The overall mean and median of fund flows are comparable between SI and conventional funds in both the UK and US samples (Panels B and D of Table 3.2). However, cross-sectional flow and return variation in both countries are higher for conventional funds than for SI funds over the sampling period.

[5] Mutual funds issue multiple share classes over time; different share classes hold the same underlying assets but have different fee structures for different types of investors. Therefore, fund flows differ at the share class level (Franzoni & Schmalz, 2017). In this thesis, I abstain from aggregating multiple share classes as our purpose is to study flows.

Table 3.2 Fund summary statistics

Panel A. UK fund sample (million $) – mean by year

		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
SI Funds	Flow	1.9	-4.1	-9.2	-4.4	-4.8	-6.0	-0.6	-2.2	1.2	-1.5
	RTN	38.5	9.8	-12.8	20.9	26.9	-5.1	0.2	-3.5	25.4	-15.4
	Fund size	517.1	450.9	359.0	380.2	408.5	312.9	285.1	272.2	326.1	265.1
	Expense Ratio	1.5	1.6	1.5	1.5	1.5	1.7	1.5	1.5	1.5	1.3
	α	-0.1	-0.1	0.1	-0.0	-0.0	0.0	-0.0	-0.1	-0.2	-0.1
	β_{mkt}	0.9	0.9	0.9	1.0	1.0	1.0	0.9	0.9	1.0	0.9
	β_{smb}	0.5	0.5	0.3	0.2	0.2	0.2	0.1	0.3	0.2	0.4
	β_{hml}	-0.1	-0.0	-0.1	-0.1	-0.2	-0.2	-0.2	-0.3	-0.2	-0.1
	β_{mom}	-0.0	-0.0	-0.0	0.0	-0.0	0.0	-0.0	-0.2	-0.1	-0.1
	Age	8.3	8.7	9.0	7.9	7.9	8.0	8.1	8.4	7.9	7.3
	Number of funds					1,031					
Conventional Funds	Flow	3.5	-1.2	-5.3	-0.9	11.1	4.2	8.9	-9.5	0.2	-4.7
	RTN	42.1	12.1	-11.0	22.0	27.8	-5.1	1.7	-3.9	25.7	-16.2
	Fund size	566.9	572.5	499.0	608.9	859.3	853.9	995.9	753.4	955.6	705.5
	Expense Ratio	1.6	1.6	1.5	1.5	1.6	1.5	1.6	1.5	1.5	1.4
	α	-0.0	0.0	0.2	0.1	0.1	0.1	0.0	-0.0	-0.1	-0.2
	β_{mkt}	0.9	0.9	0.9	1.0	1.0	1.0	0.9	0.9	0.9	0.9
	β_{smb}	0.5	0.5	0.4	0.2	0.1	0.2	0.2	0.3	0.3	0.4
	β_{hml}	-0.1	-0.0	-0.1	-0.2	-0.3	-0.2	-0.2	-0.2	-0.1	-0.1
	β_{mom}	-0.1	-0.0	-0.0	0.0	-0.0	0.0	-0.0	-0.2	-0.2	-0.1
	Age	8.0	8.1	8.1	7.4	7.0	7.2	7.4	7.9	8.2	8.3
	Number of funds					7,639					
Environment-focused SI funds	Flow	10.1	-8.2	-11.9	-6.0	-7.4	-4.5	-4.1	-1.3	2.1	-2.2
	RTN	38.2	9.1	-11.9	19.2	27.4	-3.9	-1.4	-4.9	25.7	-14.1
	Fund size	943.5	651.6	608.9	559.5	538.3	418.6	312.8	327.5	397.9	311.2
	Expense ratio	1.3	1.3	1.3	1.3	1.6	1.9	1.6	1.7	1.8	1.4
	α	-0.2	-0.2	0.1	0.0	0.0	0.0	0.0	-0.1	-0.2	-0.1
	β_{mkt}	0.9	0.9	0.9	0.9	1.0	0.9	0.9	0.9	0.9	0.9
	β_{smb}	0.6	0.6	0.4	0.2	0.2	0.3	0.2	0.3	0.3	0.4
	β_{hml}	-0.1	-0.0	-0.1	-0.1	-0.2	-0.3	-0.2	-0.3	-0.2	-0.2
	β_{mom}	-0.1	-0.0	-0.0	0.0	-0.0	0.0	-0.0	-0.3	-0.2	-0.2
	Age	9.4	9.6	10.2	7.7	7.8	7.9	7.8	8.5	8.0	7.4
	Number of funds					568					

Panel B. UK fund sample (million $) – overall summary statistics						
		Mean	Median	Min	Max	Standard Deviation
SI Funds	Flow	-1.8	-0.0	-443.4	568.5	45.9
	RTN	5.4	3.4	-61.6	82.7	22.5
	Fund size	406.2	160.5	0.2	6,394.0	732.3
	Expense Ratio	1.5	1.5	0.1	3.2	0.7
	α	-0.1	-0.1	-0.9	1.2	0.3
	β_{mkt}	1.0	1.0	0.6	1.3	0.1
	β_{smb}	0.3	0.2	-0.6	1.5	0.4
	β_{hml}	-0.1	-0.1	-0.8	0.6	0.3
	β_{mom}	-0.1	-0.0	-0.8	0.6	0.2
	Age	7.9	5.0	0.0	61.0	8.3
	Number of funds			1,031		
Conventional Funds	Flow	-0.3	-0.0	-589.7	1,049.5	58.7
	RTN	6.0	4.3	-61.6	82.7	22.5
	Fund size	749.7	259.5	0.2	9,006.0	1,203.4
	Expense Ratio	1.5	1.6	0.1	3.2	0.6
	α	-0.0	-0.0	-0.9	1.2	0.3
	β_{mkt}	1.0	1.0	0.6	1.3	0.1
	β_{smb}	0.3	0.2	-0.6	1.5	0.4
	β_{hml}	-0.1	-0.2	-0.8	0.6	0.3
	β_{mom}	-0.1	-0.0	-0.8	0.6	0.2
	Age	7.7	5.0	0.0	85.0	8.5
	Number of funds			7,639		
Environmental SI Funds	Flow	-0.7	0.0	-443.4	568.5	55.2
	RTN	5.2	2.8	-61.6	82.7	22.5
	Fund size	563.0	215.4	3.5	6,394.0	962.6
	Expense Ratio	1.4	1.5	0.1	3.2	0.7
	α	-0.1	-0.1	-0.9	0.9	0.3
	β_{mkt}	0.9	0.9	0.6	1.3	0.1
	β_{smb}	0.4	0.3	-0.6	1.4	0.3
	β_{hml}	-0.2	-0.2	-0.8	0.6	0.3
	β_{mom}	-0.1	-0.0	-0.8	0.6	0.2
	Age	8.2	6.0	0.0	35.0	8.0
	Number of funds			568		

Table 3.2 Fund summary statistics – continued											
Panel C. US fund sample (million $) – mean by year											
		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
SI Funds	Flow	-0.7	-7.1	5.7	-11.2	7.3	3.0	-8.5	-7.6	-6.8	-6.3
	RTN	31.8	16.6	-2.1	14.5	33.6	9.5	-1.4	9.8	19.5	-6.7
	Fund size	537.3	571.2	575.2	562.0	752.1	834.3	726.9	667.8	713.7	648.9
	Expense Ratio	1.5	1.4	1.4	1.4	1.3	1.3	1.2	1.2	1.1	1.1
	α	-0.0	-0.1	-0.2	-0.2	-0.2	-0.1	-0.1	-0.1	-0.2	-0.2
	β_{mkt}	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	1.0
	β_{smb}	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	β_{hml}	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1
	β_{mom}	-0.0	-0.0	-0.0	-0.0	-0.0	-0.1	0.0	-0.0	-0.0	-0.1
	Age	7.8	8.6	9.1	9.4	10.1	10.5	9.9	9.9	10.4	11.0
	Number of funds	404									
Conventional Funds	Flow	-0.8	-1.0	-2.8	-6.4	11.2	0.5	-8.3	-21.7	-20.5	-18.5
	RTN	33.5	18.8	-2.3	15.0	32.1	7.8	-2.1	11.3	19.1	-8.1
	Fund size	1,491.1	1,678.1	1,583.5	1,664.7	2,124.1	2,146.7	1,932.7	1,954.3	2,228.7	1,955.5
	Expense Ratio	1.3	1.3	1.2	1.2	1.2	1.1	1.1	1.1	1.0	1.0
	α	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.2
	β_{mkt}	1.0	1.0	1.0	1.0	1.0	1.0	0.9	0.9	0.9	0.9
	β_{smb}	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1
	β_{hml}	-0.1	-0.1	-0.1	-0.0	-0.0	-0.0	-0.0	-0.0	-0.1	-0.0
	β_{mom}	-0.0	-0.0	-0.0	-0.0	-0.0	-0.1	-0.0	-0.0	-0.0	-0.1
	Age	9.1	9.6	10.0	10.3	10.6	10.9	11.1	11.4	11.8	12.4
	Number of funds	12,715									
Environment-focused SI funds	Flow	10.7	-3.6	11.2	-10.7	13.4	5.8	-2.0	-5.1	-4.7	-3.5
	RTN	33.6	17.2	-3.0	14.5	33.7	8.9	-1.3	10.0	18.9	-6.6
	Fund size	542.4	523.1	570.6	527.7	700.0	797.2	676.6	628.3	673.6	651.6
	Expense Ratio	1.5	1.4	1.4	1.4	1.4	1.3	1.2	1.2	1.1	1.1
	α	0.0	-0.1	-0.2	-0.2	-0.2	-0.1	-0.1	-0.1	-0.2	-0.2
	β_{mkt}	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.9	0.9	0.9
	β_{smb}	0.3	0.2	0.3	0.2	0.2	0.2	0.2	0.1	0.1	0.1
	β_{hml}	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.0	-0.1	-0.0
	β_{mom}	-0.0	-0.0	-0.0	-0.0	-0.0	-0.1	0.0	-0.0	-0.0	-0.1
	Age	7.6	8.4	8.8	8.7	9.3	9.5	8.8	9.0	9.7	10.6
	Number of funds	246									

Table 3.2 Fund summary statistics – continued

Panel D. US fund sample (million \$) – overall summary statistics						
		Mean	Median	Minimum	Maximum	Standard Deviation
SI Funds	Flow	-3.1	0.0	-553.4	627.9	87.8
	RTN	8.5	10.4	-51.4	65.3	17.6
	Fund size	661.6	196.7	2.6	16,157.3	1,238.9
	Expense Ratio	1.3	1.2	0.1	2.5	0.5
	α	-0.1	-0.1	-0.8	0.7	0.2
	β_{mkt}	1.0	1.0	0.5	1.4	0.1
	β_{smb}	0.1	0.0	-0.4	1.0	0.3
	β_{hml}	-0.1	-0.0	-0.8	0.5	0.2
	β_{mom}	-0.0	-0.0	-0.5	0.7	0.1
	Age	9.2	8.0	0.0	47.0	7.2
	Number of funds			404		
Conventional Funds	Flow	-5.6	-0.0	-553.4	627.9	114.7
	RTN	8.5	9.8	-51.4	65.3	18.3
	Fund size	1,853.2	547.9	2.6	20,121.9	3,533.6
	Expense Ratio	1.2	1.1	0.1	2.5	0.5
	α	-0.1	-0.1	-0.8	0.7	0.2
	β_{mkt}	1.0	1.0	0.5	1.5	0.1
	β_{smb}	0.2	0.1	-0.4	1.0	0.3
	β_{hml}	-0.1	-0.1	-0.8	0.6	0.3
	β_{mom}	-0.0	-0.0	-0.6	0.7	0.1
	Age	10.3	9.0	0.0	94.0	9.5
	Number of funds			12,715		
Environment-focused SI funds	Flow	0.4	0.0	-553.4	477.7	92.9
	RTN	8.6	10.3	-51.4	65.3	17.7
	Fund size	604.4	164.7	2.6	16,157.3	1,232.4
	Expense Ratio	1.3	1.2	0.1	2.5	0.6
	α	-0.1	-0.1	-0.8	0.7	0.2
	β_{mkt}	1.0	1.0	0.5	1.4	0.1
	β_{smb}	0.2	0.1	-0.4	1.0	0.3
	β_{hml}	-0.1	-0.0	-0.8	0.5	0.2
	β_{mom}	-0.0	-0.0	-0.5	0.7	0.1
	Age	8.7	7.0	0.0	34.0	7.0
	Number of funds			246		

This table shows the cross-sectional mean of each fund-level variable listed for the full sample (winsorized at the top and bottom 2.5%). Fund size is valued at each fiscal year-end, and the numbers are in millions of USD (GBP). The variable Flow (fund flow) is expressed in millions of USD (GBP).

3.3.3 Sustainable investing test design

In Chapter 2, I identified shocks that trigger social preferences, e.g., those that lead to the expected changes in donation flow. In this section, I turn to my baseline analysis of these shocks on SI flows.

Consistent with my shock strength tests in Chapter 2, my tests on SI flow to estimate the following DiD model:

$$Flow_{i,t} = \alpha + \beta_0 Treat_i + \beta_1(Post_t \times Treat_i) + \beta_2 Post_t + \Sigma_\gamma X_{i,t} + \varepsilon_i \quad (3.1)$$

where $Flow_{i,t}$ is the annual net flow towards fund *i* at the end of year *t*, obtained from Morningstar.

In Equation (3.1), the chosen shocks are exogenous because they are unexpected; however, in an econometric sense, they affect flows through other channels in addition to altruism, for example, risk. Therefore, I include in the control vector $X_{i,t}$, all factor betas and alpha from the four-factor Fama-French-Carhart model and other exhaustive controls.

As part of my design, I ensure pre-shock differences in fund characteristics do not drive my results; I conduct covariance balance tests on the full sample using the following variables winsorized at top and bottom 2.5%: one-year lagged fund-level Fama-French four-factor betas: β_{mkt}, β_{smb}, β_{hml}, β_{mom} and alpha, one-year lagged holding period returns, and fund age (Appendix 3.2). I also only include funds over three years old at the year of each shock to avoid incubation bias (Evans, 2010). At the fund level, I first calculate each fund's monthly excess returns, then run a Fama-French four-factor model using the past 24 months of data. Factor returns and the risk-free rate in the US and UK markets are obtained from Kenneth R. French's website.

I then proceed with logit propensity score matching based on pre-shock fund financial characteristics and fund-specific sustainability attributes classification relevant to the different shocks. The matching process is described under the table notes in Appendix 3.2.

For the two environmental shocks, withdrawal from the Paris Agreement and the 2016 US presidential election, I define treatment and control groups as follows:

Treat$_i$ = 1 if fund *i* is identified as an environment-focused SI fund, and 0 if fund *i* is identified as a non-environment-focused SI fund.

An exception is the BP oil spill shock which happened in 2010 when the number of SI funds, particularly environment-focused SI funds, is limited. In this case, I resort to a broader definition: $Treat_i$ = 1 if fund *i* is identified as an SI fund, and 0 if fund *i* is identified as a conventional fund.

Because my second set of shocks (the Oxfam scandal and the US Tax Cut and Jobs Act policy) is selected to capture events affecting the overall attractiveness of the entire charity sector, I cannot separate SI funds into shock related versus non-related subgroups. . Instead, I define $Treat_i$ = 1 if a fund *i* is identified by Morningstar as an SI fund, and 0 if fund *i* is identified as a conventional fund. $X_{i,t}$ is a matrix of control variables for fund *i*. which includes fund age and other variables to control for risk and return of funds: one-year lagged returns and contemporaneous Fama-French four-factor betas and alpha. I also include institutional dummy variable in my regression models. I follow prior literature, such as Chevalier and Ellison (1997) and Hartzmark and Sussman (2019), on the selection of controls. All data on control variables except for betas and alpha are obtained from Morningstar.

$Post_t$ is a dummy variable equal to zero for pre-shock years and one for post-shock years. If a shock takes place at the start of a year, I expect changes in fund flows or donations to realize by the end of the same year. If a shock takes place towards the end of a year, flow or donation changes may also be witnessed in the following year. For example, when examining the US presidential election shock, which took place in November 2016, my baseline model uses:

$$Post_t = 1 \text{ for years 2016 and 2017 and } = 0 \text{ for year 2015}$$

In Appendix 3.1, I show evidence that all matched samples satisfy the parallel trend assumptions; thus, the results are not driven by differences in pre-shock flow trends between treated and control groups.

3.4 Sustainable investing test results

3.4.1 Main results

Tables 3.3 through 3.7 show my UK and US fund flow analysis results. I begin with the BP oil spill to provide continuity with previous work, but in light of the charity results shown in Chapter 2, I do not consider it evidence of changes to social preferences.[6] In addition, I could not test its effect on philanthropy in the US due to data limitations, but I still include it in the SI analysis for both countries for consistency purposes. In fact, if my hypothesis is correct, I should find no effect on SI flows after these shocks as well. After the BP oil spill, UK and US SI funds did not exhibit significant flow changes compared to conventional funds (Tables 3.3 and 3.4). Only in one specification for the UK, where the institutional dummy is not included, I observe outflow (significant at 10%). One could consider the BP oil spill as a falsification test, whereby no effect on philanthropy should correspond to no effect on SI flows. Out of four specifications tested on BP oil spill shock, the three specifications where I find no significance are in line with the argument. However, the significant effect in one of the specifications is against this reasoning. The only result where the effect is positive, even if insignificant, is the 2016 US presidential election shock. To assess the economic size of the coefficient (USD 6.6 million), I compare it to the average annual flow to US SI funds over the ten years up to 2019 of USD 6.3 billion (Hale, 2021). Therefore, the effect of the 2016 US presidential election shock is only 0.1% of the average annual flow to US SI funds.

Table 3.3 UK SI fund flow – environmental shocks

	BP Oil Spill		US Paris Agreement withdrawal shock	
	(i)	(ii)	(iii)	(iv)
Post$_t$	14.50	12.26	-1.77	-1.02
	(18.07)	(18.20)	(4.45)	(4.59)
Treat$_i$	-0.60	-3.14	5.90	6.52
	(6.70)	(7.36)	(6.54)	(6.74)
Post$_t$ × Treat$_i$	-15.10*	-12.53	-2.44	-4.92
	(7.83)	(8.63)	(4.04)	(4.40)
Lagged_RTN$_{i,t}$	0.07	0.06	0.40*	0.42**
	(0.17)	(0.17)	(0.21)	(0.21)
RTN$_{i,t}$	0.34**	0.35**	0.34**	0.35**
	(0.17)	(0.17)	(0.15)	(0.15)
Age$_i$	-0.44*	-0.44*	-0.58***	-0.59***
	(0.23)	(0.23)	(0.15)	(0.15)
Lagged Fundsize$_{i,t}$	-0.02***	-0.02***	0.00	0.00
	(0.00)	(0.00)	(0.01)	(0.01)
Expense Ratio$_{i,t}$	-1.88	-2.04	2.18	2.05
	(3.19)	(3.21)	(1.63)	(1.64)
Lagged $\alpha_{i,t}$	11.75**	11.35**	-0.91	-1.34
	(5.37)	(5.40)	(6.27)	(6.29)
Lagged $\beta_{mkt,i,t}$	-2.49	-1.63	-3.21	-3.63
	(15.41)	(15.53)	(10.95)	(10.96)
Lagged $\beta_{smb,i,t}$	-0.05	-0.28	-8.93	-8.87
	(5.35)	(5.37)	(6.87)	(6.88)
Lagged $\beta_{hml,i,t}$	12.55	12.25	-13.36	-14.16*
	(9.07)	(9.10)	(8.34)	(8.37)
Lagged $\beta_{mom,i,t}$	-30.75**	-30.22*	6.28	6.52
	(15.35)	(15.39)	(8.59)	(8.60)

[3] Bialkowski and Starks (2016) use the BP oil spill and the Fukushima nuclear disaster together in a single specification at monthly frequency and find a positive effect on flows.

Institutional	N	Y	N	Y
Fund Family Fixed Effects	Y	Y	Y	Y
Number of observations	430	430	555	555
Number treated funds	87	87	94	94
Number of control funds	63	63	97	97
Adjusted R-squared	0.135	0.132	0.135	0.135

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. BP oil spill shock: $Post_t = 1$ if year = 2010, 2011 and $Post_t = 0$ if year = 2009. US Paris Agreement withdrawal shock: $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2016. In BP oil spill shock, I employ 1(treated)-to-1(control) matching for SI funds (treated) and conventional SI funds (control) with replacement. In the US Paris Agreement withdrawal shock, I employ 1(treated)-to-1(control) matching for environmental SI (treated) and non-environmental SI (control) funds without replacement. Control variables include 1-year lagged holding period return ($Lagged\_RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size ($Lagged\ Fundsize_{i,t}$, in millions), expense ratio ($Expense\ Ratio_{i,t}$) and fund age (Age_i). Institutional dummy and its interaction term with $Post_t \times Treat_i$ are included in my specification; the coefficients are insignificant. The number of treated and control funds reflects those used in the estimation sample of each model. Results and estimation sample size from models without any control variables are shown in the appendix.

Table 3.4 US SI fund flow – environmental shocks

	BP Oil Spill		US Paris Agreement withdrawal shock		2016 US presidential election	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	56.04**	52.83**	0.23	-3.58	10.94	-1.57
	(24.52)	(25.21)	(20.06)	(21.54)	(18.32)	(19.65)
$Treat_i$	2.22	4.21	-69.18**	-98.06***	-9.53	-33.09
	(14.67)	(15.50)	(33.64)	(35.60)	(24.09)	(25.49)
$Post_t \times Treat_i$	-6.10	-10.10	-9.53	-11.25	-11.12	-6.84
	(11.34)	(12.94)	(18.61)	(21.94)	(18.08)	(20.74)
$lagged\_RTN_{i,t}$	-0.65*	-0.62*	1.42	1.45	0.29	0.24
	(0.34)	(0.34)	(0.95)	(0.94)	(0.73)	(0.72)
$RTN_{i,t}$	0.08	0.05	0.25	0.27	0.67	0.67
	(0.26)	(0.26)	(0.54)	(0.54)	(0.67)	(0.66)
Age_i	-5.30***	-5.31***	-4.40***	-4.12***	-4.94***	-4.82***
	(0.51)	(0.51)	(0.81)	(0.81)	(0.82)	(0.81)
$Lagged\ Fundsize_{i,t}$	-0.01**	-0.01**	-0.03***	-0.03***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Expense\ Ratio_{i,t}$	-10.43	-10.50	27.59	31.36*	41.03***	43.57***
	(7.56)	(7.57)	(16.95)	(16.83)	(15.69)	(15.48)
Lagged $\alpha_{i,t}$	68.16***	67.90***	49.13	47.63	21.23	18.67
	(15.92)	(15.94)	(47.66)	(47.41)	(40.52)	(39.96)
Lagged $\beta_{mkt,i,t}$	11.03	10.65	24.61	1.79	4.16	-32.44
	(33.48)	(33.70)	(113.10)	(112.20)	(90.70)	(90.01)
Lagged $\beta_{smb,i,t}$	23.50*	23.46*	-108.48**	-101.73*	-68.53	-55.20
	(12.04)	(12.10)	(54.45)	(53.95)	(41.61)	(41.19)
Lagged $\beta_{hml,i,t}$	19.58	19.66	81.14	73.10	93.89*	89.70*
	(20.47)	(20.57)	(60.21)	(59.74)	(48.21)	(47.58)
Lagged $\beta_{mom,i,t}$	60.55*	62.38*	-41.75	-19.23	-62.49	-42.67
	(33.80)	(34.05)	(193.91)	(192.50)	(82.63)	(81.65)
Institutional	N	Y	N	Y	N	Y
Fund Family Fixed Effects	Y	Y	Y	Y	Y	Y
Number of observations	688	688	354	354	385	385
Number of treated funds	113	113	62	62	68	68
Number of control funds	102	102	62	62	64	64
Adjusted R-squared	0.343	0.342	0.258	0.273	0.356	0.374

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. BP oil spill shock: $Post_t = 1$ if year = 2010, 2011 and $Post_t = 0$ if year = 2009. US Paris Agreement withdrawal shock: $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2016. 2016 US presidential election shock: $Post_t = 1$ if year = 2016, 2017 and $Post_t = 0$ if year = 2015. In BP oil spill shock, I employ 1(treated)-to-1(control) matching for SI funds (treated) and conventional funds (control) with replacement. In the US Paris Agreement withdrawal shock and 2016 US presidential election shock, I employ 1(treated)-to-1(control) matching for environmental SI (treated) and non-environmental SI (control) funds without replacement. Control variables include 1-year lagged holding period return ($Lagged\_RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size ($Lagged\ Fundsize_{i,t}$, in millions), expense ratio ($Expense\ Ratio_{i,t}$) and fund age (Age_i). Institutional dummy and its interaction term with $Post_t \times Treat_i$ are included in my specification; the coefficients are insignificant. The number of treated and control funds reflects those used in the estimation sample of each model. Results and estimation sample size from models without any control variables are shown in the appendix.

My results are different from those in Bialkowski and Starks (2016) (BS) for several possible reasons: 1) BS use monthly data frequency and aggregate flows of SI and conventional funds and combine two environmental shocks in the same analysis – the BP oil spill and the Fukushima nuclear disaster. Given my annual frequency, I cannot isolate the two disasters as they happen too closely. To address possible contamination by Fukushima, I show consistent results with 2010 as the only post-shock year (Table 3.5 – Panel B). It is also possible that these two shocks affected very different ideologies regarding renewable energy among environmentally conscious investors, and therefore their impacts may be different and offsetting. 2) Because BS aggregate fund flows into the two categories - SI and conventional- they cannot control for fund-level risk and return variables as I do. This means that the significant effect in BS could be a result of risk/return concerns. BS use matched conventional funds as controls, which I also report here for comparability. In addition, the granularity of data available on each fund's specific sustainability focus also allows for comparison between environmental versus non-environmental SI funds (Table 3.5 – Panel A).

Table 3.5 BP oil spill 2010 – matched sample results

Panel A. Baseline $Post_t$ definition

	Environmental vs. Non-Environmental SI funds			SI vs. Conventional funds		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	254.69***	237.75***	232.24***	43.42	51.98	54.37
	(60.05)	(57.61)	(57.94)	(37.19)	(36.78)	(37.08)
$Treat_i$	66.08***	45.58***	46.81***	22.54	23.18	29.87*
	(16.28)	(16.19)	(16.25)	(14.72)	(14.78)	(15.79)
$Post_t$ × $Treat_i$	-61.76***	-30.13*	-33.18*	-33.92*	-39.63**	-43.63**
	(17.70)	(16.61)	(16.95)	(17.45)	(17.47)	(18.62)
Lagged $RTN_{i,t}$	-1.33***	-1.60***	-1.57***	-0.04	-0.15	-0.15
	(0.50)	(0.45)	(0.46)	(0.35)	(0.34)	(0.34)
$RTN_{i,t}$	3.10***	2.79***	2.74***	0.50	0.43	0.44
	(0.60)	(0.57)	(0.57)	(0.33)	(0.33)	(0.33)
$Age_{i,t}$	0.40	-0.11	-0.13	-1.04*	-1.06*	-1.03*
	(0.61)	(0.55)	(0.55)	(0.55)	(0.55)	(0.55)
Lagged $Fundsize_{i,t}$	-0.08***	-0.10***	-0.10***	-0.09***	-0.09***	-0.09***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$Institutional_i$	-13.89	8.53	-6.34	-11.45	14.62	41.38
	(14.54)	(12.85)	(20.73)	(12.12)	(12.75)	(32.26)
Lagged Expense $Ratio_{i,t}$		43.31***	43.68***		30.12***	30.38***
		(5.26)	(5.28)		(5.33)	(5.34)
$Post_t$ × $Institutional_i$			22.10			-14.24
			(24.16)			(37.17)
$Treat_i$ × $Institutional_i$						-55.11
						(45.08)
$Post_t$ × $Treat_i$ × $Institutional_i$						30.63
						(52.84)
Lagged $\alpha_{i,t}$	35.55**	37.13**	36.89**	14.81	6.50	8.79
	(18.05)	(15.88)	(15.89)	(11.08)	(11.17)	(11.31)
Lagged $\beta_{mkt,i,t}$	69.13	178.64***	177.31***	-8.25	-45.24	-52.21
	(57.43)	(57.80)	(57.84)	(32.57)	(33.76)	(34.15)
Lagged $\beta_{smb,i,t}$	50.30**	105.90***	107.00***	21.51*	26.44**	28.26**
	(20.24)	(18.37)	(18.42)	(11.24)	(11.18)	(11.27)
Lagged $\beta_{hml,i,t}$	-29.11	46.03*	44.07*	-54.00***	-50.67***	-49.97***
	(26.34)	(24.17)	(24.28)	(17.09)	(16.90)	(16.93)

Lagged $\beta_{mom,i,t}$	70.99	-125.94***	-130.96***	-4.51	-31.12	-31.22
	(50.74)	(48.13)	(48.45)	(28.36)	(29.44)	(29.49)
Intercept	-296.10***	-474.25***	-469.44***	-3.76	-24.26	-23.04
	(72.20)	(77.32)	(77.52)	(41.23)	(41.24)	(41.37)
N	315	307	307	448	424	424
Adjusted R^2	0.307	0.493	0.493	0.358	0.411	0.410

Panel B. Alternative $Post_t$ definition

	Environmental vs. Non-Environmental SI funds			SI vs. Conventional funds		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	-237.46***	-205.82***	-231.22***	-53.30	-45.77	-45.17
	(78.09)	(73.19)	(73.43)	(49.60)	(45.54)	(45.66)
$Treat_i$	98.47***	67.85***	70.89***	22.02	20.87	25.74*
	(14.59)	(14.40)	(14.33)	(14.56)	(13.74)	(14.74)
$Post_t$ × $Treat_i$	-118.00***	-58.50***	-69.75***	-20.06	-27.18	-25.79
	(20.05)	(19.24)	(19.76)	(19.49)	(18.22)	(19.65)
Lagged $RTN_{i,t}$	5.14***	3.39***	3.70***	1.22**	1.08**	1.10**
	(0.89)	(0.83)	(0.83)	(0.54)	(0.50)	(0.50)
$RTN_{i,t}$	5.19***	2.66***	2.73***	1.37**	1.15**	1.26**
	(0.91)	(0.88)	(0.87)	(0.56)	(0.52)	(0.53)
$Age_{i,t}$	0.16	-0.89	-0.89	-1.28*	-1.11*	-1.04
	(0.67)	(0.60)	(0.60)	(0.69)	(0.65)	(0.65)
Lagged $Fundsize_{i,t}$	0.05***	-0.00	0.00	-0.08***	-0.09***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
$Institutional_i$	-5.37	14.58	-10.43	-13.64	20.89	43.57
	(16.05)	(14.18)	(18.27)	(14.56)	(14.20)	(29.44)
Lagged Expense $Ratio_{i,t}$		39.61***	39.56***		38.34***	38.74***
		(5.28)	(5.23)		(5.41)	(5.41)
$Post_t$ × $Institutional_i$			49.72**			2.95
			(23.26)			(38.04)
$Treat_i$ × $Institutional_i$						-43.49
						(41.25)
$Post_t$ × $Treat_i$ × $Institutional_i$						-11.11
						(54.82)
Lagged $\alpha_{i,t}$	-29.29	-25.48	-28.73*	8.74	-5.13	-2.38
	(18.96)	(16.83)	(16.73)	(13.99)	(13.08)	(13.21)
Lagged $\beta_{mkt,i,t}$	-3.90	106.92	101.25	2.93	-31.67	-42.76
	(69.87)	(66.00)	(65.40)	(39.47)	(37.59)	(38.11)
Lagged $\beta_{smb,i,t}$	-58.43**	39.15	35.29	-1.37	4.48	6.03
	(26.62)	(26.25)	(26.06)	(15.95)	(14.78)	(14.87)
Lagged $\beta_{hml,i,t}$	-88.67***	-12.90	-20.41	-75.98***	-67.56***	-66.51***
	(29.55)	(27.65)	(27.60)	(19.12)	(17.65)	(17.65)
Lagged $\beta_{mom,i,t}$	124.38**	-30.94	-37.54	22.51	-16.81	-16.61
	(50.76)	(48.99)	(48.60)	(31.76)	(30.60)	(30.60)
Intercept	-3.29	-155.08*	-135.64	23.21	-11.30	-8.79
	(83.81)	(85.62)	(85.26)	(47.25)	(43.89)	(43.91)
N	196	192	192	287	273	273
Adjusted R^2	0.593	0.698	0.704	0.372	0.484	0.485

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Panel A: $Post_t = 1$ if year = 2010, 2011 and $Post_t = 0$ if year = 2009. In Panel B: $Post_t = 1$ if year = 2010 and $Post_t = 0$ if year = 2009. Additional test results using $Post_t = 1$ if year = 2010, 2011, and $Post_t = 0$ if year = 2009, 2010 are available upon request. In Column (iii), interaction terms: $Post_t$ × $Treat_i$ × $Institutional_i$ and $Treat_i$ × $Institutional_i$ is omitted from the regression due to lack of observations. In Columns (i) to (iii), I employ 1(treated)-to-2(control) matching by pre-shock fund characteristics for environmental SI funds (treated) and non-environmental SI funds (control); in columns (iv) to (vi), I employ 1(treated)-to-2(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$), and fund age (Age_i).

The results I rely on to test whether pro-social motives drive SI flows are the ones that trigger philanthropy: the Paris Agreement withdrawal (third and fourth columns of Tables 3.3 and 3.4) and the 2016 US presidential election (the last two columns of Table 3.4). However, I do not report the 2016 US presidential election shock for the UK sample, as the shock year is 2016, which overlaps with another important UK event: the Brexit referendum. In this case, a lack of significance could be due to pro-social investors expecting the UK to pass pro-SI regulation much faster than the EU, which results in a lack of urgency to invest privately in SI. Therefore, I rely on the cleaner US-based test.

While the Paris Agreement Withdrawal was shown to affect altruism slightly in the UK, the reaction to both shocks was strong in the US (see Chapter 2). Yet again, after controlling for risk and return factors, I do not find any impact on environmental SI fund flows in two distinct institutional settings of the UK and the US. Both shocks took place after the introduction of the Morningstar globes ratings, which dramatically improved visibility and reduced search costs, as confirmed by Hartzmark and Sussman (2019). Therefore, the lack of significance cannot be explained by low visibility. Ramelli, Wagner et al. (2021) perform firm-level event study analysis following the election and the nomination of Scott Pruitt to head the EPA and show that the cumulative abnormal returns were higher for both carbon-intensive firms and firms with climate-responsible strategies. In their further analysis, the authors suggest that institutional investors value a firm's climate responsibility for strategic reasons rather than for altruistic motives. While their results and ours are not comparable because they use individual firm stock returns and holdings, I analyze flows to mutual funds. Their finding that the main driving force behind the market reactions following the 2016 US presidential election is strategic rather than altruistic is in line with my results.

I address the fact that during the period of my analysis, certain institutional investors are prohibited from investing in SI funds by law (US pension funds), and this might result in a lack of an effect by including an institutional dummy variable and its interactions with the $Post_t$ and $Treat_i$ variables (columns (ii) and (iv) of Table 3.3 and columns (ii), (iv) and (vi) of Table 3.4). In these specifications, the baseline $Post_t \times Treat_i$ interaction captures the effect on retail-investors-focused funds and remains insignificant. Therefore, retail investors, who have complete control over the allocation of their resources, do not respond to these shocks. Tables 3.3 and 3.4 show that, overall, I do not find a positive and significant effect of the two environmental shocks on the flows to environmental SI funds.

My insignificant results in Tables 3.3 and 3.4 could be due to the nature of the shocks examined. One could argue that while altruism is still a common motive for both, the 2016 US presidential election and the US Paris Agreement withdrawal shock are not as relevant to SI as they are to philanthropy. To address this, I then examine two other shocks, which focus on two common benefits of philanthropy and SI: credibility in achieving pro-social goals and monetary advantage and monetary advantage. In the first case, when the charity sector loses its reputation for delivering altruistic goals, those who place a sufficiently high value on these goals might then be to redirect resources to SI funds instead of the philanthropy-conventional fund bundle. Regarding the monetary advantage of philanthropy, a sudden drop in donations' tax shield ability can make charitable giving less attractive relative to SI. In this case too, if investors value altruism, they are more likely to choose to invest in SI instead. Both of these mechanisms fit the Modigliani-Miller type theoretical framework in Zivin and Small (2005) and Hart and Zingales (2017), where investors are indifferent between SI and a philanthropy-conventional bundle because both satisfy a combination of risk/return and social preferences.

Table 3.6 UK and US SI fund flows – Oxfam scandal and TCJA

	UK – Oxfam Scandal		US – Tax Cut and Jobs Act	
	(i)	(ii)	(iii)	(iv)
$Post_t$	16.45	23.40*	-24.58*	-24.25
	(12.41)	(13.03)	(14.26)	(15.31)
$Treat_i$	-10.20	-3.83	18.21	21.86
	(8.62)	(9.17)	(15.52)	(16.86)
$Post_t \times Treat_i$	3.80	-0.43	-8.88	-15.29
	(8.05)	(8.70)	(13.92)	(16.20)
$lagged\_RTN_{i,t}$	0.52**	0.58**	3.22***	3.23***
	(0.25)	(0.25)	(0.68)	(0.69)
$RTN_{i,t}$	0.49**	0.53***	1.09***	1.10***
	(0.19)	(0.19)	(0.39)	(0.39)
Age_i	-1.13***	-1.07***	-5.22***	-5.20***
	(0.38)	(0.38)	(0.63)	(0.63)
Lagged $Fundsize_{i,t}$	-0.02***	-0.02***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Expense $Ratio_{i,t}$	-6.95	-7.15	12.54	12.66
	(5.04)	(5.03)	(10.63)	(10.66)
Lagged $\alpha_{i,t}$	9.76	8.72	-20.07	-20.67
	(7.08)	(7.20)	(26.54)	(26.62)
Lagged $\beta_{mkt,i,t}$	30.46	30.01	-82.89*	-84.01*
	(20.74)	(20.69)	(43.62)	(43.69)
Lagged $\beta_{smb,i,t}$	-23.45**	-24.90**	3.99	4.18
	(9.94)	(9.94)	(21.20)	(21.23)
Lagged $\beta_{hml,i,t}$	13.89	12.32	47.39**	46.81**
	(15.01)	(15.05)	(20.36)	(20.40)
Lagged $\beta_{mom,i,t}$	-16.43	-24.41	92.14	91.51
	(28.49)	(28.67)	(60.22)	(60.54)
Institutional	N	Y	N	Y
Fund Family Fixed Effects	Y	Y	Y	Y
Number of observations	408	408	967	967
Number of treated funds	79	79	170	171
Number of control funds	85	85	161	161
Adjusted R-squared	0.254	0.258	0.250	0.248

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Oxfam scandal shock: $Post_t = 1$ if year = 2011, 2012 and $Post_t = 0$ if year = 2010. Tax Cut and Jobs Act shock: $Post_t = 1$ if year = 2018 and $Post_t = 0$ if year = 2017. In both shocks, I employ 1(treated)-to-1(control) matching for SI funds (treated) and Conventional funds (control) with replacement. Control variables include 1-year lagged holding period return ($Lagged\_RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i). Institutional dummy and its interaction term with $Post_t \times Treat_i$ are included in my specification; the coefficients are insignificant. The number of treated and control funds reflects those used in the estimation sample of each model. Results and estimation sample size from models without any control variables are shown in the appendix.

As established in Chapter 2, the reputation shock to Oxfam and the entire UK charity sector induced a reduction in donations post-scandal in 2011 and 2012. Keeping social preference constant, would imply a redirection of resources away from the philanthropy-conventional bundle and towards an alternative vehicle like SI. Nonetheless, I do not find increased flow to SI funds versus conventional ones in the UK. The results are similar when I assess the effect of the TCJA on SI. While I illustrated that the tax policy shock negatively impacts philanthropy in Chapter 2: there is a sizable reduction in donations made by US individuals after its implementation, US SI funds did not experience a significant change after TCJA. Controlling for institutionally focused funds does not significantly impact post-shock fund flows for institutional or retail SI funds (Table 3.6). If a redirection from philanthropy to SI was partially funded with a reduction in conventional investing, it would be more likely for my regressions to detect an effect. The lack of significance, therefore, strengthens my conclusions. In fact, when examining the change in flows to conventional funds and SI funds separately in Table 3.7, I find that instead of a reduction, the TCJA shock induced increased investment in conventional funds while having no effect on SI. Therefore, the higher disposable income some investors enjoyed after TCJA went mostly to conventional investing. For comparison, the SCF numbers show an outflow from investments in both the invest-only and the invest&donate groups. However, they cannot be attributed to the TCJA alone since the survey waves are three years apart and reflect multiple economic and political developments. I obtain similar results for conventional fund flows after the Oxfam scandal, showing that the shock does not significantly impact UK SI fund flows.

Table 3.7 UK and US fund flows – Separate Analysis – Oxfam scandal and TCJA

	UK – Oxfam Scandal		US – Tax Cut and Jobs Act	
	Conventional Funds	SI funds	Conventional Funds	SI funds
$Post_t$	25.58***	-28.08	53.51***	13.73
	(7.87)	(28.24)	(6.64)	(39.20)
$lagged\_RTN_{i,t}$	0.19	-0.78	0.97***	2.00***
	(0.22)	(0.70)	(0.15)	(0.75)
$RTN_{i,t}$	1.00***	-0.28	1.76***	1.31
	(0.18)	(0.43)	(0.23)	(1.41)
Age_i	-0.90***	-0.97**	-3.84***	-3.57***
	(0.16)	(0.48)	(0.16)	(0.76)
Lagged $Fundsize_{i,t}$	-0.01***	-0.03***	-0.01***	-0.01
	(0.00)	(0.01)	(0.00)	(0.01)
Expense $Ratio_{i,t}$	0.14	7.73	12.74***	-2.99
	(2.10)	(6.87)	(2.17)	(10.90)
Lagged $\alpha_{i,t}$	22.73***	22.98	54.02***	-12.99
	(4.95)	(17.69)	(5.87)	(37.56)
Lagged $\beta_{mkt,i,t}$	-8.51	-24.45	-17.26**	-24.11
	(11.91)	(30.85)	(8.69)	(39.07)

Lagged $\beta_{smb,i,t}$	-13.79***	26.68*	4.04	-13.83
	(5.25)	(15.92)	(3.69)	(20.97)
Lagged $\beta_{hml,i,t}$	-9.92	18.73	22.75***	-9.65
	(7.14)	(23.58)	(5.15)	(22.79)
Lagged $\beta_{mom,i,t}$	-7.28	-70.64	-5.33	-2.52
	(14.56)	(48.09)	(7.50)	(39.74)
Institutional	Y	Y	Y	Y
Number of observations	1343	119	13224	331
Adjusted R-squared	0.103	0.266	0.170	0.100

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Oxfam scandal shock: $Post_t = 1$ if year = 2011 and $Post_t = 0$ if year = 2010. Tax Cut and Jobs Act shock: $Post_t = 1$ if year = 2018 and $Post_t = 0$ if year = 2017. Control variables include 1-year lagged holding period return ($Lagged\_RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i). Institutional dummy and its interaction term with $Post_t \times Treat_i$ are included in my specification; the coefficients are insignificant. Results and estimation sample size from models without any control variables are shown in the appendix. Additional results using funds from 1(treated)-to-3(control) and 1(treated)-to-4(control) matching for SI funds (treated) and Conventional funds (control) with replacement are available in the appendix.

In terms of coefficients on control variables, consistent with previous literature on fund flows, such as Renneboog et al. (2011) and Bialkowski and Starks (2016), I find that across almost all shocks and both the US and the UK, investors are attracted to younger and smaller funds. When controlling for Fama-French four-factor betas and alpha, I obtain similar results to Benson and Humphrey (2008) – the one-year lagged returns are not strongly and persistently associated with fund flows.

To summarise, I do not obtain empirical evidence that environmental shocks, loss of reputation of the charity sector, or reduced tax benefits of philanthropy translate into increased post-shock period SI flows.

3.4.2 Additional Tests

In my SI flow analysis shown above, a lack of significance of my tests could be attributed to the lack of visibility of SI funds, statistical power, or an ambiguous control group since Morningstar SI designation may not be able to identify conventional funds who also incorporate ESG factors in their investment decisions but do not explicitly label themselves as "sustainable". I address these two possibilities at once using the Morningstar Globes rating instead (Hartzmark & Sussman, 2019). Morningstar Globes rating is a portfolio-weighted measure for mutual funds' ESG performance, and I classify as treated - funds with a 5-globe rating and as controls - funds with 1- and 2-globe ratings. This approach produces a much larger sample of treated and control candidate funds which overcomes the problem of visibility and potentially low statistical power. In Appendix 3.3, I show that my results are not affected when increasing the number of observations to alleviate the concern of low statistical power. Table A3.3.1 shows the findings of the 2016 election shock. In Tables A3.3.2 and Table A3.3.3, I use SI vs. conventional definitions of treated and controls for the UK and US, respectively, and my findings are unaffected.

In addition, my main results are based on a two-year post-shock period, which one may argue dilutes any treatment effect and would explain finding no significance.

The results in this chapter are obtained using annual frequency fund flow data, as available in Morningstar. This is a limitation to the study as monthly frequency fund flow data would have allowed the DiD model to pinpoint precisely the pre- and post-shock months around each shock, leading to a more precise estimation of any treatment effect. However, as a result of the annual frequency data, the definition of post-shock periods is restricted to: either the shock year itself or the shock year and the year next when a shock happens closer to the year-end. I address this potential concern by using an alternative one-year post-shock period definition, and my results remain unchanged. However, this exercise does not rule out possible fund flow changes in the months following the shocks. Therefore, in this chapter, any immediate, short-term changes in fund flows may have been overlooked. Thus, results found in this chapter do not indicate the absolute absence of social preferences among SI investors, especially for months after the shock.

I further use event studies to complement and verify my analysis of SI investor demand. Under an event study setting, if investors demand SI due to altruistic motives after the shocks, the impact should be reflected in the cumulative abnormal returns on the treated, shock-relevant SI funds, thereby verifying my DiD analysis results. The advantage of an event study also lies in that it utilizes daily data as opposed to annual frequency, thus strengthening the test. My event studies exclude the BP oil spill shock due to limited data availability for 2013 and earlier. I first obtain fund net asset value data (NAV) from Bloomberg and construct two equally weighted portfolios of the same treated and control funds used in each baseline test. The two portfolios then form a long-short portfolio. For example, concerning the Paris Agreement withdrawal shock, I calculate the return from an equally weighted portfolio with a long position in environmentally focused SI funds and a short position in non-environmentally focused SI funds. I use the Standard & Poor's 500 Index as the market benchmark. Under a 250-day estimation window and a 30-day event window setting, I estimate the long-short portfolios' abnormal returns and cumulative abnormal returns.[7] Figure 3.2 shows that the difference between abnormal returns of environmental and non-environmental SI funds is not distinguishable from zero (the 95% confidence interval bands are always far from 0). Therefore, there is no relative short-term surge in demand for environmental SI funds.

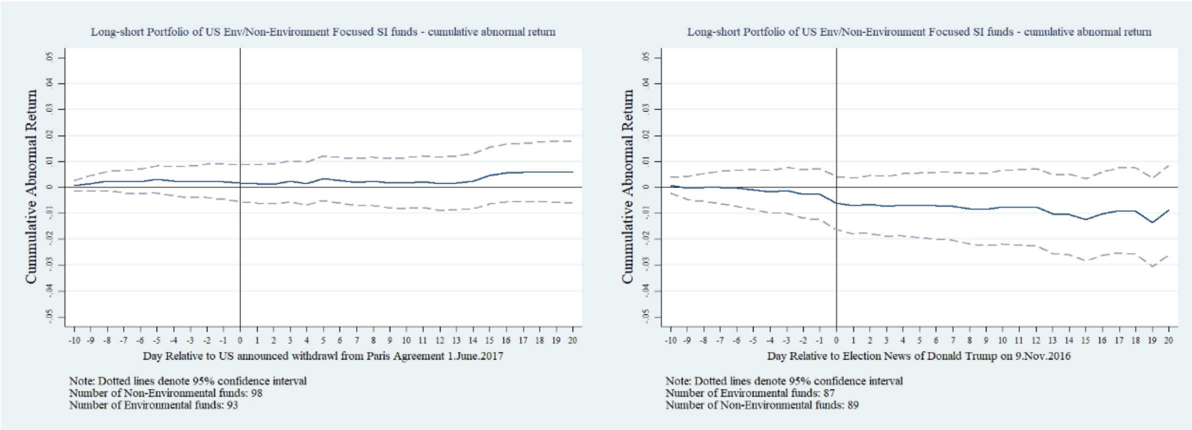


Figure 3.2 – Long-short portfolio CAR – Paris Agreement and 2016 US presidential election shock

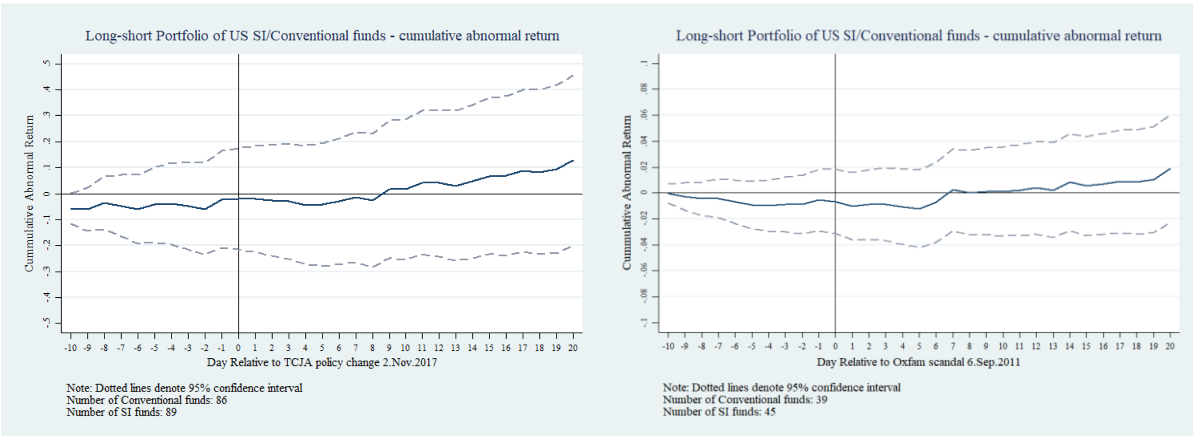


Figure 3.3 Long-short portfolio CAR – Tax Policy Shock 2017 & Oxfam scandal SI and conventional funds

Figure 3.3 demonstrates that, relative to conventional funds, SI funds did not receive significantly higher abnormal returns after the Oxfam scandal or the TCJA shock (0 is always included in the 95% confidence intervals). Both figures confirm my results from the DiD analysis that SI flows do not react to the shocks I study.

[7] I also graph CARs separately for environmental SI funds, non-environmental SI funds, SI funds and conventional funds (see appendix). The main conclusion does not change.

3.5 Conclusion

By assessing the reaction of SI fund flows to environmental shocks that affect social preferences, I show that SI fund flows do not appear to be driven by social preferences. I test the possibility that the lack of SI effect may be due to the low relevance of the specific shocks I choose to SI investors compared to charity donors. However, when I examine shocks that are unrelated to a particular charitable cause – i.e., a shock to the reputation of charitable institutions, which undermines the effectiveness of charities to deliver pro-social goals, as well as a shock to the tax shield of philanthropy – there is still no impact on SI flows, which supports the conclusion that pro-social motives are unlikely to drive SI flows.

The following possibilities for not finding a significant effect on fund flows are also embedded in my research design and setting: (1) different visibility and search costs associated with selecting SI funds versus charitable institutions by considering periods of high and low visibility; (2) the possibility that environment-related shocks affect not just altruism but the risk-return profile of SI funds by analyzing shocks exogenous to firm risk exposures and explicitly controlling for risk and return variables; (3) low statistical power by examining an alternative approach to the treated-control definition that produces a much larger number of observations; (4) different motivations behind individual versus institutional investor SI flows by including retail versus institutional interaction in my regressions; (5) the presence of a short-term effect instead of a long-term one by performing both a daily frequency event study and an annual frequency difference-in-difference regression analysis.

The mechanisms behind my findings that social preferences do not affect SI flows could be either that these preferences are not a strong enough driver for SI flows to be detected or that investors do not see SI as a vehicle for satisfying them. Regardless of the mechanism, if policy makers' objectives are limiting climate change or improving social justice, my results suggest that SI in its present form does not respond to investor preferences in line with these objectives.

Several recent studies examine the drivers of SI theoretically and empirically. The novelty in my approach is to analyze SI in relation to shocks to social preferences instead of to conventional funds, as is the typical approach in existing work. By doing so, I narrow in on altruism as a possible explanation behind the recent surge in SI flows and products, which has been hard to isolate (Hartzmark & Sussman, 2019). My empirical design has the advantage of being able to circumvent some of the econometric challenges that researchers typically encounter, as it benefits from a number of separate exogenous shocks of a different nature that affect altruism: environmental disasters, a misconduct scandal, political events, and tax reform affecting philanthropy. I am also able to perform clean balancing of treated and control samples of charity institutions and SI funds, given the richness and detail of my data.

The rise of SI can be seen as a response by active fund managers to the overwhelming shift toward passive management in recent years (Waite et al., 2019). Fund managers offer increasing numbers of SI products as a way of distinguishing themselves from competitors (Cao et al., 2021). This trend could reflect changing investor tastes, as modeled by Fama and French (2007). I show, however, that social preferences are not the key force behind these changing tastes, at least in the period 2008-2018. For example, Ramelli, Wagner et al. (2021) find that institutional investors choose to hold climate-responsible firms for strategic reasons.

Considering the recommendations in Hong et al. (2020) for higher regulatory ESG mandates in the interest of first-best societal welfare outcomes, my results support the necessity for intervention since altruism alone appears not to be sufficient to drive investment flows.

If the need to care about the environment or reduce inequality is not a driver of SI, then socially responsible consumers mainly rely on institutions that directly transfer resources to these causes rather than using market-based solutions. It has long been recognized that direct money transfers are likely to distort incentives (Economides et al., 2008). Therefore, improving the trust towards SI or reducing the heightened expectations in terms of its risk-return properties while focusing more on its potential welfare implications may encourage socially conscious investors to choose these investments to satisfy their social preferences and contribute to a more sustainable economy.

The transformative changes that the Covid pandemic brought about in 2020 may have resulted in a much greater increase in both preferences for ESG and resulting SI-directed investment flows (Pástor and Vorsatz, 2020; UBS, 2020). The largest asset management companies are taking an active stand: BlackRock's Larry Fink urged companies to disclose how they are preparing for a "net zero world", where net greenhouse gas emissions are eliminated by 2050, while the main stewardship priorities in 2021 for State Street Global Advisors will be the systemic risks associated with climate change and a lack of racial and ethnic diversity on company boards (Fink, 2021; Taraporevala, 2021). It would be interesting to revisit my research question at a future time when the global economy has largely recovered and other policy or natural shocks can help with inference to see if stronger socially responsible preferences are making a difference. For now, the recent enormous growth in SI does not seem to reflect altruism.

In this Chapter, I show that social preferences do not appear to affect SI fund flows. I isolate social preferences by choosing environmental shocks that trigger philanthropy. I test for the possibility that social preferences could still be a common motive for SI and philanthropy, but the lack of an SI effect may be due to the low relevance of the specific shocks I choose. When I examine shocks that are unrelated to a particular charitable cause – i.e., a shock to the reputation of charitable institutions, which reduces the relative effectiveness of the whole charitable sector to deliver social-related goals, as well as a shock to the tax shield of philanthropy, which makes charity-giving less attractive, there is still no impact on SI, which supports the conclusion that social motives are unlikely to drive SI flows.

Appendix 3.1 Covariate Balancing and Parallel Trends – UK

Table A3.1.1 BP oil spill 2010 – covariate balancing test – UK funds

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	38.45	9.22	-0.08	0.94	0.35	-0.01	-0.02
Treated	37.88	11.67	-0.28	0.85	0.65	-0.12	-0.08
t-stat	0.31	-1.93	3.08	3.24	-5.55	3.13	2.60
Median Test							
Control	35.44	7.00	-0.04	0.94	0.26	-0.04	0.04
Treated	37.42	10.00	-0.38	0.81	0.71	-0.08	-0.05
z-score	-1.56	-3.50	3.54	3.76	-5.13	2.11	3.03
Number of control funds	83	100	70	70	70	70	70
Number of treated funds	59	63	59	59	59	59	59
Panel B. After Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	38.35	10.85	-0.07	0.91	0.41	-0.03	-0.03
Treated	37.88	11.80	-0.28	0.85	0.65	-0.12	-0.08
t-stat	0.24	-0.61	2.97	2.29	-4.43	2.50	2.05
Median Test							
Control	35.93	7.00	-0.02	0.90	0.27	-0.06	0.04
Treated	37.42	10.00	-0.38	0.81	0.71	-0.08	-0.05
z-score	-1.24	-2.12	3.39	2.91	-4.22	1.41	2.55
Number of control funds	59	59	59	59	59	59	59
Number of treated funds	59	59	59	59	59	59	59
Panel C. Before Matching (SI vs. Conventional Funds)							
Mean Test							
Control	41.91	10.18	-0.05	0.92	0.43	-0.13	-0.06
Treated	38.21	10.17	-0.17	0.90	0.49	-0.06	-0.05
t-stat	3.34	0.02	3.98	1.96	-1.48	-3.15	-1.57
Median Test							
Control	39.41	7.00	-0.07	0.92	0.42	-0.14	-0.06
Treated	36.30	9.00	-0.22	0.89	0.47	-0.06	-0.03
z-score	3.50	-0.92	3.67	1.99	-2.37	-3.41	-1.90
Number of control funds	1,548	1,802	1,426	1,426	1,426	1,426	1,426
Number of treated funds	142	163	129	129	129	129	129
Panel D. After Matching (SI vs. Conventional Funds)							
Mean Test							
Control	37.87	10.98	-0.14	0.92	0.49	-0.04	-0.05
Treated	37.58	11.21	-0.19	0.89	0.50	-0.06	-0.05
t-stat	0.23	-0.22	0.93	1.37	-0.22	0.64	0.05
Median Test							
Control	36.37	8.00	-0.17	0.92	0.44	-0.04	-0.05
Treated	36.54	9.00	-0.26	0.86	0.47	-0.06	-0.04
z-score	-0.09	-0.92	0.87	1.29	-0.54	0.63	-0.36
Number of control funds	97	97	97	97	97	97	97
Number of treated funds	125	125	125	125	125	125	125

This table shows balancing test results using 1-year lagged covariates as of the shock year 2017. The number of treated and control funds reflects the unique fund count in each group. Matched sample in Panel B is obtained by truncating the before-matching sample (Panel A) to keep observations with a logit propensity score between 0 and 0.33 (inclusive) or between 0.38 and 0.8 (inclusive). Sorting by propensity score, each environmental SI fund is matched by the closest two non-environmental funds with replacement. The sample in Panel D is obtained by truncating the before-matching sample (Panel C) to keep observations with a logit propensity score between 0 and 0.15 (inclusive).

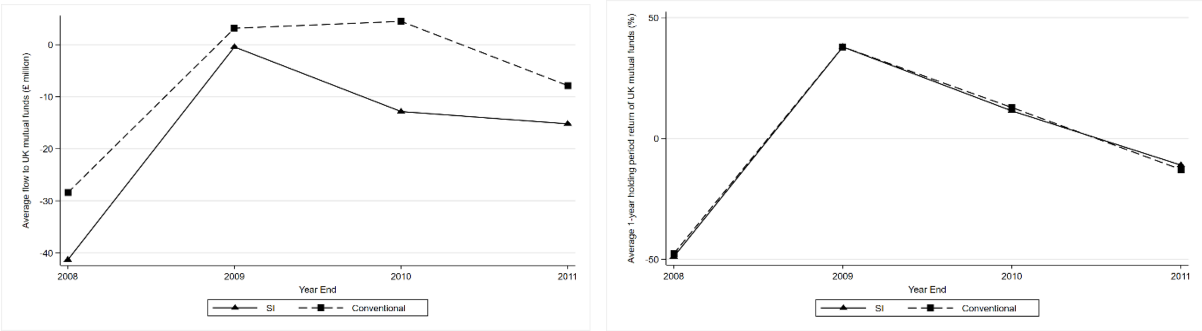


Figure A3.1.1 Parallel trends of UK treatment and control funds – BP oil spill 2010

Table A3.1.2 US Paris Agreement withdrawal shock – covariate balancing test – UK funds

	Lagged RTN$_{i,t}$	Age$_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	-2.77	12.12	-0.05	0.96	0.21	-0.20	-0.16
Treated	-5.92	11.70	-0.12	0.93	0.40	-0.36	-0.28
t-stat	3.71	0.43	3.24	1.76	-3.81	3.81	3.55
Median Test							
Control	-2.65	10.00	-0.05	0.99	0.14	-0.22	-0.09
Treated	-6.67	9.00	-0.11	0.97	0.39	-0.40	-0.28
z-score	3.69	1.17	2.56	2.13	-4.10	4.37	3.65
Number of control funds	134	166	135	135	135	135	135
Number of treated funds	104	131	109	109	109	109	109
Panel B. After Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	-3.49	11.06	-0.06	0.94	0.22	-0.26	-0.20
Treated	-4.67	11.02	-0.10	0.92	0.31	-0.28	-0.23
t-stat	1.18	0.03	1.59	0.78	-1.68	0.49	0.88
Median Test							
Control	-2.83	10.00	-0.07	0.97	0.15	-0.28	-0.13
Treated	-5.58	7.00	-0.11	0.97	0.34	-0.32	-0.28
z-score	1.52	0.79	1.30	0.77	-1.65	1.21	0.96
Number of control funds	97	97	97	97	97	97	97
Number of treated funds	97	97	97	97	97	97	97
Panel C. Before Matching (SI vs. Conventional Funds)							
Mean Test							
Control	-4.21	10.04	-0.05	0.94	0.32	-0.24	-0.21
Treated	-3.62	10.92	-0.07	0.95	0.27	-0.26	-0.20
t-stat	-1.48	-1.84	1.38	-0.98	1.83	0.98	-0.51
Median Test							
Control	-4.39	8.00	-0.05	0.95	0.24	-0.27	-0.14
Treated	-2.92	9.00	-0.07	0.97	0.18	-0.24	-0.13
z-score	-2.03	-2.61	1.51	-1.52	1.31	0.82	-0.11
Number of control funds	3,774	4,349	3,574	3,574	3,574	3,574	3,574
Number of treated funds	274	335	270	270	270	270	270
Panel D. After Matching (SI vs. Conventional Funds)							
Mean Test							
Control	-3.85	11.82	-0.06	0.94	0.26	-0.26	-0.21
Treated	-3.72	11.19	-0.06	0.95	0.28	-0.25	-0.20
t-stat	-0.22	0.73	0.23	-0.81	-0.49	-0.16	-0.44
Median Test							
Control	-4.45	9.00	-0.07	0.94	0.19	-0.30	-0.25
Treated	-3.04	9.00	-0.06	0.97	0.20	-0.25	-0.13
z-score	-0.56	-0.12	0.52	-1.50	-0.66	-0.17	-0.20
Number of control funds	235	235	235	235	235	235	235
Number of treated funds	257	257	257	257	257	257	257

This table shows balancing test results using 1-year lagged covariates as of the shock year 2017. The number of treated and control funds reflects the unique fund count in each group. Matched sample in Panel B is obtained by truncating the before-matching sample (Panel A) to keep observations with a logit propensity score between 0 and 0.33 (inclusive) or between 0.38 and 0.8 (inclusive). Sorting by propensity score, each environmental SI fund is matched by the closest two non-environmental funds with replacement. The sample in Panel D is obtained by truncating the before-matching sample (Panel C) to keep observations with a logit propensity score between 0 and 0.15 (inclusive).

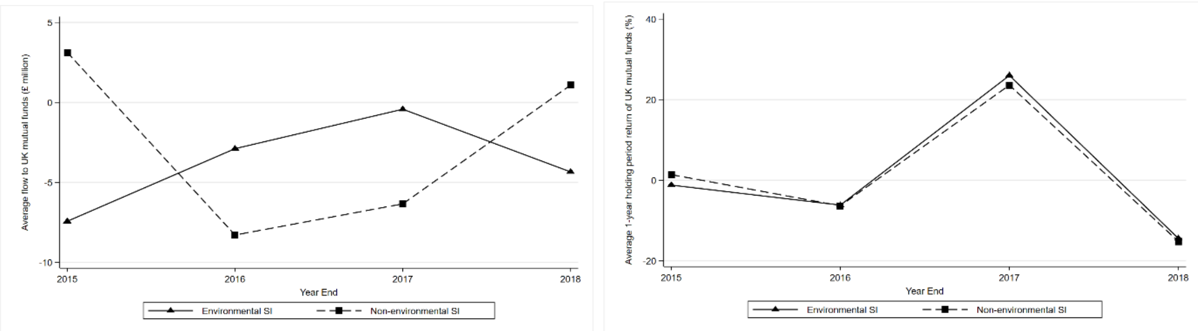


Figure A3.2.2 Parallel trends of UK treatment and control funds – Paris Agreement withdrawal 2017

Table A3.1.3 Oxfam scandal 2011 – covariate balancing test – UK funds

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (SI vs. Conventional Funds)							
Mean Test							
Control	11.73	10.21	0.03	0.94	0.45	-0.03	-0.01
Treated	11.23	10.50	-0.15	0.92	0.50	-0.00	-0.02
t-stat	0.61	-0.43	5.82	2.12	-1.21	-1.87	0.69
Median Test							
Control	10.39	8.00	-0.02	0.94	0.41	-0.04	-0.01
Treated	11.05	8.00	-0.22	0.90	0.41	0.04	-0.00
z-score	-0.10	-1.61	6.38	2.01	-1.86	-3.02	0.46
Number of control funds	1,761	2,045	1,641	1,641	1,641	1,641	1,641
Number of treated funds	153	177	149	149	149	149	149
Panel B. After Matching (SI vs. Conventional Funds)							
Mean Test							
Control	13.08	10.99	-0.02	0.92	0.54	-0.01	-0.01
Treated	10.97	10.81	-0.08	0.94	0.46	-0.03	-0.00
t-stat	1.72	0.16	1.26	-0.60	1.46	1.13	-0.71
Median Test							
Control	11.10	8.00	-0.08	0.92	0.45	0.02	-0.01
Treated	10.42	10.00	-0.13	0.93	0.39	-0.02	0.01
z-score	1.46	-0.57	1.53	-0.77	1.01	0.77	-0.81
Number of control funds	110	110	110	110	110	110	110
Number of treated funds	118	118	118	118	118	118	118

This table shows balancing test results using 1-year lagged covariates as of the shock year 2011. The number of treated and control funds reflects the unique fund count in each group. Matched sample in Panel B is obtained by truncating the before-matching sample (Panel A) to keep observations with a logit propensity score between 0 and 0.25 (inclusive). Sorting by propensity score, each SI fund is matched by the closest conventional fund with replacement.

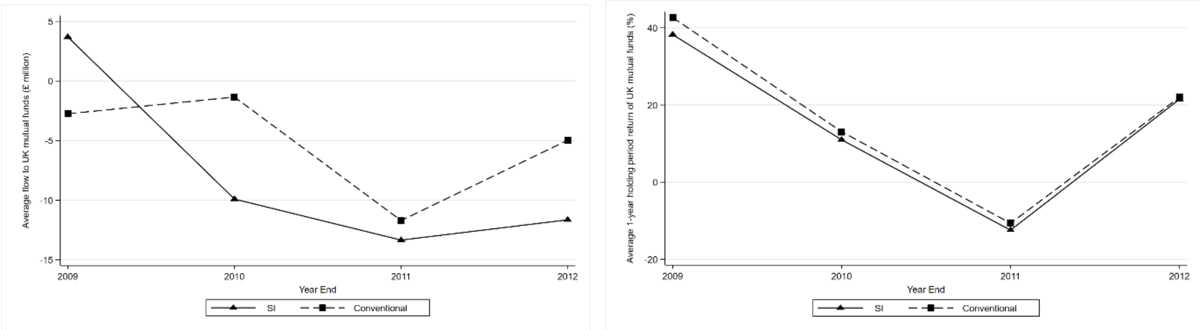


Figure A3.1.3 Parallel trends of UK treatment and control funds – Oxfam scandal shock 2011

Table A3.1.4 BP oil spill 2010 – covariate balancing test – US funds

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	29.58	9.94	-0.04	1.01	0.15	-0.10	-0.00
Treated	32.60	9.73	-0.01	1.01	0.26	-0.15	-0.01
t-stat	-1.59	0.24	-0.70	-0.03	-2.21	1.49	0.32
Median Test							
Control	30.83	10.00	-0.10	1.02	0.07	-0.10	0.01
Treated	31.86	9.00	0.05	1.04	0.16	-0.15	0.01
z-score	-0.66	0.45	-0.22	-1.02	-2.03	1.59	0.62
Number of control funds	63	71	62	62	62	62	62
Number of treated funds	80	88	76	76	76	76	76
Panel B. After Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	29.64	10.15	-0.04	1.01	0.15	-0.10	-0.00
Treated	30.31	10.30	-0.01	1.01	0.26	-0.15	-0.01
t-stat	-0.41	-0.16	-0.70	-0.03	-2.21	1.49	0.32
Median Test							
Control	30.90	10.00	-0.10	1.02	0.07	-0.10	0.01
Treated	30.35	9.00	0.05	1.04	0.16	-0.15	0.01
z-score	0.16	-0.31	-0.22	-1.02	-2.03	1.59	0.62
Number of control funds	61	62	62	62	62	62	62
Number of treated funds	73	76	76	76	76	76	76
Panel C. Before Matching (SI vs. Conventional Funds)							
Mean Test							
Control	33.58	10.76	0.03	1.03	0.19	-0.14	-0.02
Treated	31.27	9.82	-0.03	1.01	0.21	-0.13	-0.00
t-stat	2.33	1.31	2.52	2.11	-0.57	-0.25	-1.68
Median Test							
Control	31.75	9.00	0.02	1.03	0.09	-0.12	-0.02
Treated	30.90	9.00	-0.02	1.02	0.16	-0.13	0.01
z-score	2.28	0.27	2.62	2.06	-1.77	0.06	-2.76
Number of control funds	5,448	6,044	5,386	5,386	5,386	5,386	5,386
Number of treated funds	143	159	138	138	138	138	138
Panel D. After Matching (SI vs. Conventional Funds)							
Mean Test							
Control	31.42	9.52	-0.01	1.02	0.24	-0.17	-0.01
Treated	30.01	10.26	-0.02	1.00	0.21	-0.13	-0.01
t-stat	1.11	-0.80	0.58	0.93	0.74	-1.54	0.06
Median Test							

Control	30.47	7.00	0.02	1.02	0.13	-0.17	-0.02
Treated	30.77	9.50	-0.02	1.02	0.16	-0.13	0.01
z-score	0.81	-3.12	0.97	0.97	-0.14	-1.82	-0.58
Number of control funds	110	110	110	110	110	110	110
Number of treated funds	134	134	134	134	134	134	134

This table shows balancing test results using 1-year lagged covariates as of the shock year 2017. The number of treated and control funds reflects the unique fund count in each group. In Panel B, sorting by propensity score, each environmental SI fund is matched by the closest non-environmental fund without replacement to obtain the matched sample. The sample in Panel D is obtained by truncating the before-matching sample (Panel C) to keep observations with a logit propensity score between 0 and 0.1 (inclusive).

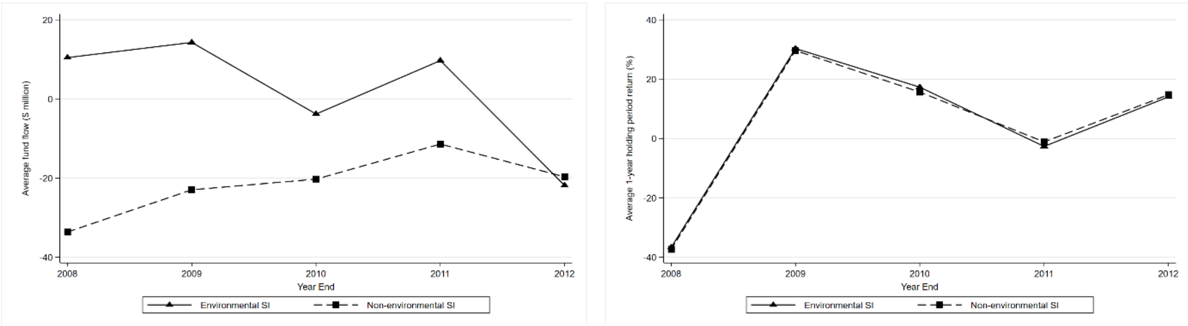


Figure A3.1.4 Parallel trends of US treatment and control funds – BP Oil Spill 2010

Table A3.1.5 US Paris Agreement withdrawal shock – covariate balancing test – US funds

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	8.73	14.51	-0.16	0.99	0.03	-0.09	-0.03
Treated	9.76	12.28	-0.13	0.94	0.15	-0.05	-0.01
t-stat	-1.00	2.54	-0.88	2.97	-3.33	-1.09	-1.02
Median Test							
Control	9.09	13.00	-0.12	0.99	-0.04	-0.02	-0.00
Treated	10.27	11.50	-0.13	0.94	0.01	0.00	0.01
z-score	-1.38	2.96	-0.22	3.22	-2.44	-2.06	-1.94
Number of control funds	62	97	78	78	78	78	78
Number of treated funds	118	150	137	137	137	137	137
Panel B. After Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	8.73	13.95	-0.17	0.99	0.04	-0.08	-0.02
Treated	9.86	13.61	-0.18	0.98	0.01	-0.06	-0.04
t-stat	-1.03	0.25	0.57	0.35	0.88	-0.68	1.02
Median Test							
Control	9.09	12.00	-0.13	0.99	-0.04	-0.02	-0.00
Treated	10.27	12.00	-0.16	0.99	-0.02	-0.01	-0.00
z-score	-1.53	0.53	1.16	0.86	0.88	-1.30	0.10
Number of control funds	62	62	62	62	62	62	62
Number of treated funds	62	62	62	62	62	62	62
Panel C. Before Matching (SI vs. Conventional Funds)							
Mean Test							
Control	11.22	13.60	-0.15	0.94	0.16	-0.04	-0.02
Treated	9.40	13.15	-0.14	0.96	0.11	-0.06	-0.02
t-stat	3.07	0.74	-0.50	-2.13	2.35	1.49	-0.22
Median Test							
Control	9.83	12.00	-0.14	0.97	0.01	-0.05	-0.00
Treated	9.74	12.00	-0.13	0.98	-0.00	-0.01	0.00
z-score	2.29	-0.29	-0.73	-1.52	1.54	0.73	0.25
Number of control funds	7,013	9,257	8,094	8,094	8,094	8,094	8,094
Number of treated funds	180	247	215	215	215	215	215
Panel D. After Matching (SI vs. Conventional Funds)							
Mean Test							

Control	7.83	11.97	-0.13	0.95	0.04	-0.11	-0.01
Treated	8.10	13.16	-0.16	0.96	0.05	-0.08	-0.01
t-stat	-0.43	-1.46	1.49	-0.94	-0.36	-1.31	-0.25
Median Test							
Control	7.26	10.00	-0.14	0.98	-0.03	-0.11	-0.01
Treated	9.11	12.00	-0.14	0.99	-0.02	-0.01	-0.00
z-score	-1.32	-1.52	1.18	-0.82	-0.92	-1.83	-0.52
Number of control funds	143	143	143	143	143	143	143
Number of treated funds	148	148	148	148	148	148	148

This table shows balancing test results using 1-year lagged covariates as of the shock year 2017. The number of treated and control funds reflects the unique fund count in each group. In Panel B, sorting by propensity score, each environmental SI fund is matched by the closest non-environmental fund without replacement to obtain the matched sample. Sorting by propensity score, each environmental SI fund is matched by the closest non-environmental fund with replacement. The sample in Panel D is obtained by truncating the before-matching sample (Panel C) to keep observations with a logit propensity score between 0.02 and 0.07 (inclusive).

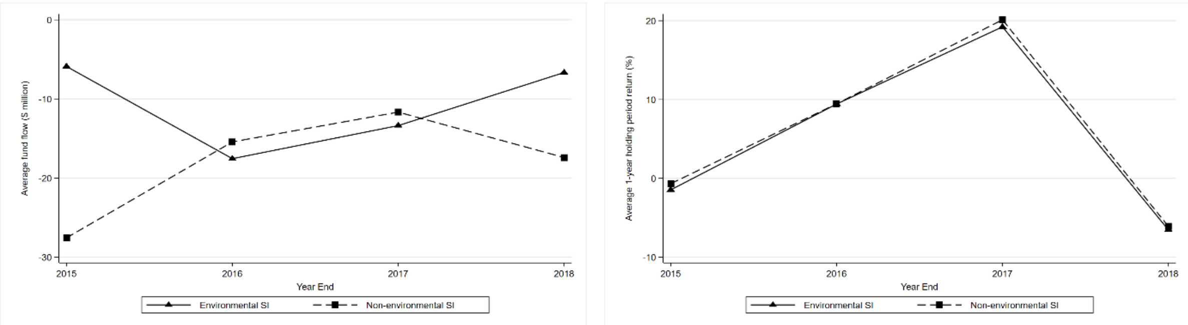


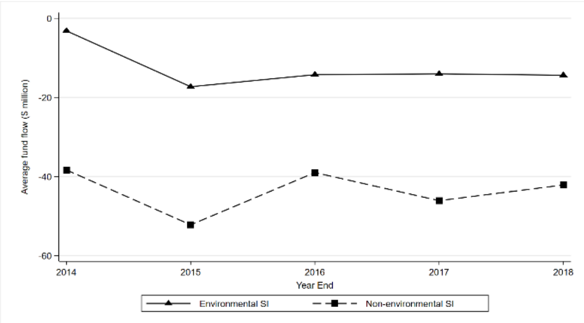
Figure A3.1.5 Parallel trends of US treatment and control funds – US Paris Agreement withdrawal shock 2017

Table A3.1.6 2016 US presidential election shock – covariate balancing test – US funds

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	-1.34	13.48	-0.10	1.00	0.04	-0.04	-0.00
Treated	-0.97	11.67	-0.08	0.95	0.19	-0.07	0.01
t-stat	-0.57	2.08	-0.87	3.91	-3.90	0.88	-1.04
Median Test							
Control	-0.60	12.00	-0.07	1.00	-0.02	0.07	-0.01
Treated	-1.66	10.00	-0.11	0.97	0.04	-0.01	0.01
z-score	1.46	2.56	0.46	3.56	-3.55	0.82	-0.66
Number of control funds	77	97	80	80	80	80	80
Number of treated funds	130	143	125	125	125	125	125
Panel B. After Matching (Environmental vs. Non-Environmental SI funds)							
Mean Test							
Control	-1.34	12.96	-0.11	1.00	0.05	-0.04	-0.00
Treated	-1.21	13.61	-0.11	0.99	0.06	-0.03	0.01
t-stat	-0.18	-0.58	0.05	1.21	-0.55	-0.22	-1.33
Median Test							
Control	-0.60	11.00	-0.07	1.01	-0.03	0.07	-0.01
Treated	-1.23	13.00	-0.11	1.00	0.01	-0.01	0.02
z-score	0.63	-0.40	0.88	1.45	-0.75	-0.13	-1.29
Number of control funds	77	77	77	77	77	77	77
Number of treated funds	77	77	77	77	77	77	77
Panel C. Before Matching (SI vs. Conventional Funds)							
Mean Test							
Control	-2.09	13.32	-0.14	0.94	0.17	-0.05	-0.00
Treated	-1.11	12.40	-0.09	0.97	0.14	-0.06	0.00
t-stat	-2.34	1.54	-3.19	-4.25	1.68	0.51	-0.77
Median Test							
Control	-1.75	12.00	-0.12	0.96	0.03	-0.04	0.00
Treated	-0.87	11.00	-0.09	0.99	0.03	0.02	0.01
z-score	-2.71	0.85	-2.93	-4.36	0.87	0.15	0.26

Number of control funds	7,577	8,652	7,622	7,622	7,622	7,622	7,622
Number of treated funds	207	240	205	205	205	205	205
Panel D. After Matching (SI vs. Conventional Funds)							
Mean Test							
Control	-0.90	12.31	-0.08	0.97	0.12	-0.07	0.01
Treated	-1.12	12.28	-0.09	0.98	0.15	-0.06	0.01
t-stat	0.37	0.04	0.28	-0.84	-0.91	-0.48	-0.06
Median Test							
Control	-0.64	11.00	-0.08	0.99	0.02	-0.00	0.01
Treated	-0.87	10.00	-0.09	0.99	0.03	0.04	0.01
z-score	0.78	-0.16	0.28	-0.41	-1.37	-0.08	0.08
Number of control funds	155	155	155	155	155	155	155
Number of treated funds	159	159	159	159	159	159	159

This table shows balancing test results using 1-year lagged covariates as of the shock year 2016. The number of treated and control funds reflects the unique fund count in each group. In Panel B, sorting by propensity score, each environmental SI fund is matched by the closest non-environmental fund without replacement to obtain the matched sample. The sample in Panel D is obtained by truncating the before-matching sample (Panel C) to keep observations with a logit propensity score between 0.01 and 0.049 (inclusive).



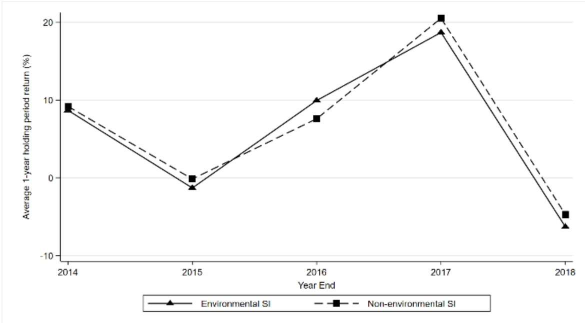


Figure A3.1.6 Parallel trends of US treatment and control funds – 2016 US presidential election

Table A3.1.7 Tax Cut and Jobs Act 2017 – covariate balancing test – US funds

	Lagged RTN$_{i,t}$	Age$_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (SI vs. Conventional Funds)							
Mean Test							
Control	11.22	13.60	-0.15	0.94	0.16	-0.04	-0.02
Treated	9.40	13.15	-0.14	0.96	0.11	-0.06	-0.02
t-stat	3.07	0.74	-0.50	-2.13	2.35	1.49	-0.22
Median Test							
Control	9.83	12.00	-0.14	0.97	0.01	-0.05	-0.00
Treated	9.74	12.00	-0.13	0.98	-0.00	-0.01	0.00
z-score	2.29	-0.29	-0.73	-1.52	1.54	0.73	0.25
Number of control funds	7,013	9,257	8,094	8,094	8,094	8,094	8,094
Number of treated funds	180	247	215	215	215	215	215
Panel B. After Matching (SI vs. Conventional Funds)							
Mean Test							
Control	8.50	12.14	-0.14	0.95	0.08	-0.10	-0.01
Treated	9.33	12.91	-0.15	0.95	0.08	-0.06	-0.02
t-stat	-1.25	-1.02	0.73	-0.16	0.18	-1.71	0.83
Median Test							
Control	7.64	10.00	-0.14	0.98	-0.02	-0.10	-0.01
Treated	9.67	12.00	-0.13	0.98	-0.01	-0.01	-0.00
z-score	-1.80	-1.03	0.41	0.16	-0.89	-2.13	0.14
Number of control funds	166	166	166	166	166	166	166

Number of treated funds	171	171	171	171	171	171	171

This table shows balancing test results using 1-year lagged covariates as of the shock year 2017. The number of treated and control funds reflects the unique fund count in each group. Matched sample in Panel B is obtained by truncating the before-matching sample (Panel A) to keep observations with a logit propensity score between 0.01 and 0.1 (inclusive). Sorting by propensity score, each SI fund is matched by the closest conventional fund with replacement.

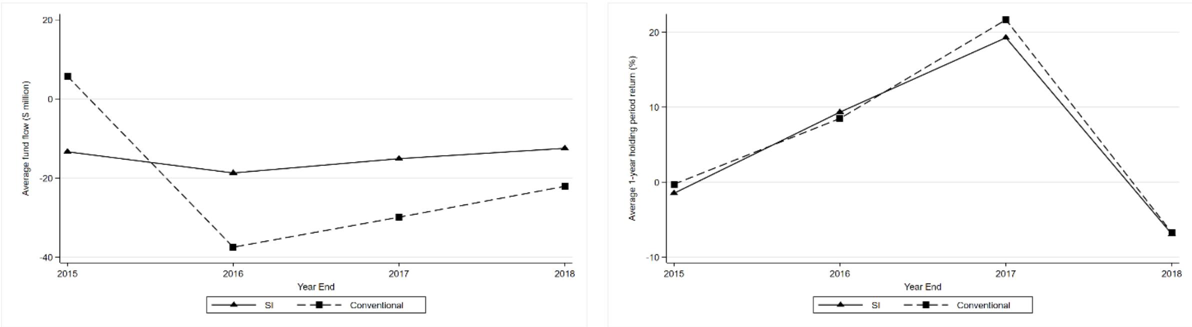


Figure A3.1.7 Parallel trends of US treatment and control funds – Tax Cut and Jobs Act 2017

Appendix 3.2 – Robustness check – fund flows

Table A3.2.1 US high vs. low Morningstar Sustainability Globes rating funds – Election shock 2016

	(i)	(ii)	(iii)	(iv)
$Post_t$	-31.40***	-38.23***	-38.99***	-42.59***
	(8.88)	(8.42)	(8.39)	(8.99)
$Treat_i$	-8.44	1.24	-2.53	0.08
	(8.93)	(8.48)	(8.48)	(9.36)
$Post_t \times Treat_i$	12.90	14.98	14.67	11.20
	(9.62)	(9.11)	(9.08)	(10.29)
$Lagged\_RTN_{i,t}$	0.17	0.27	0.30	0.29
	(0.27)	(0.25)	(0.25)	(0.25)
$RTN_{i,t}$	1.29***	1.28***	1.38***	1.36***
	(0.30)	(0.29)	(0.29)	(0.29)
Age_i	-4.37***	-4.03***	-3.90***	-3.90***
	(0.28)	(0.27)	(0.27)	(0.27)
Lagged $Fundsize_{i,t}$		-0.01***	-0.01***	-0.01***
		(0.00)	(0.00)	(0.00)
Expense $Ratio_{i,t}$			32.63***	32.37***
			(6.61)	(6.63)
Lagged $\alpha_{i,t}$	71.26***	79.66***	89.45***	88.91***
	(14.33)	(13.59)	(13.69)	(13.68)
Lagged $\beta_{mkt,i,t}$	44.88	8.89	17.38	16.21
	(32.50)	(30.87)	(30.81)	(30.81)
Lagged $\beta_{smb,i,t}$	-4.30	-32.58***	-38.79***	-38.76***
	(8.56)	(8.27)	(8.33)	(8.33)
Lagged $\beta_{hml,i,t}$	49.83***	20.88*	21.70**	21.21*
	(11.58)	(11.10)	(11.06)	(11.08)
Lagged $\beta_{mom,i,t}$	-75.03***	-36.54*	-36.49*	-35.22*
	(21.95)	(20.91)	(20.83)	(20.84)
Institutional	N	N	N	Y
Fund Family Fixed Effects	Y	Y	Y	Y
Number of observations	2976	2976	2974	2974
Number of treated funds	507	507	507	507
Adjusted R-squared	0.223	0.303	0.308	0.309

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $Post_t = 1$ for the 6-month period (May to October 2016 before the election in November 2016, and $Post_t = 0$ for the 6-month period (November to April after the election). $Treat_i = 1$ if a fund is rated with "Above Average" or "High" consistently by Morningstar Sustainability Globes rating for the 24 months period starting from August 2018 (when the rating is first accessible); $Treat_i = 0$ if a fund is rated with "Below Average" or "Low" by Morningstar Sustainability Globes rating consistently for the 24 months period starting from August 2018. I employ 1(treated)-to-1(control) matching for treated funds and control funds with replacement. Control variables include 1-year lagged holding period return (Lagged_RTN$_{i,t}$), contemporaneous return (RTN$_{i,t}$), 1-year lagged fund size (Lagged Fundsize$_{i,t}$ in millions), expense ratio (Expense Ratio$_{i,t}$) and fund age (Age$_i$). The institutional dummy, and its interaction term with Post$_t$ × Treat$_i$ are included in column (iv); the coefficients are insignificant.

Table A3.2.2 UK SI fund flow – alternative Treat$_i$ definition

	BP Oil Spill		Paris Agreement Shock	
	(i)	(ii)	(iii)	(iv)
Post$_t$	8.38	7.49	3.97	1.98
	(21.94)	(22.22)	(6.82)	(7.35)
Treat$_i$	-92.72***	-96.27***	-0.90	-2.35
	(32.83)	(33.98)	(5.75)	(6.18)
Post$_t$ × Treat$_i$	10.75	9.22	-7.76	-6.14
	(9.07)	(9.80)	(5.81)	(6.37)
Lagged_RTN$_{i,t}$	-0.15	-0.13	0.52*	0.55*
	(0.21)	(0.22)	(0.29)	(0.30)
RTN$_{i,t}$	0.15	0.15	0.47**	0.49**
	(0.21)	(0.22)	(0.20)	(0.21)
Age$_i$	-0.01	0.06	-1.02***	-1.03***
	(0.42)	(0.42)	(0.23)	(0.23)
Lagged Fundsize$_{i,t}$	-0.06***	-0.06***	-0.01***	-0.01***
	(0.01)	(0.01)	(0.00)	(0.00)
Expense Ratio$_{i,t}$	0.05	-1.03	-0.25	-0.27
	(4.61)	(4.79)	(3.74)	(3.76)
Lagged $\alpha_{i,t}$	22.02**	21.92**	1.48	1.15
	(10.77)	(10.87)	(8.64)	(8.68)
Lagged $\beta_{mkt,i,t}$	-49.27	-49.64	-6.02	-5.01
	(41.75)	(42.13)	(18.57)	(18.69)
Lagged $\beta_{smb,i,t}$	-9.12	-8.92	11.30	10.92
	(14.33)	(14.54)	(7.95)	(7.99)
Lagged $\beta_{hml,i,t}$	2.14	1.10	-13.25	-13.68
	(22.56)	(22.99)	(10.78)	(10.84)
Lagged $\beta_{mom,i,t}$	-5.35	0.78	5.61	5.08
	(34.53)	(35.51)	(13.61)	(13.72)
Institutional	N	Y	N	Y
Fund Family Fixed Effects	Y	Y	Y	Y
Number of observations	202	202	464	464
Number of treated funds	59	59	257	257
Adjusted R-squared	0.420	0.414	0.260	0.255

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. BP oil spill shock: Post$_t$ = 1 if year = 2010, 2011 and Post$_t$ = 0 if year = 2009. US Paris Agreement withdrawal shock: Post$_t$ = 1 if year = 2017, 2018; Post$_t$ = 0 if year= 2016. In both shocks, I employ 1(treated)-to-1(control) matching for SI funds (treated) and Conventional funds (control) with replacement. Control variables include 1-year lagged holding period return (Lagged_RTN$_{i,t}$), contemporaneous return (RTN$_{i,t}$), 1-year lagged fund size (Lagged Fundsize$_{i,t}$ in millions), expense ratio (Expense Ratio$_{i,t}$) and fund age (Age$_i$). Institutional dummy and its interaction term with Post$_t$ × Treat$_i$ are included in my specification and are insignificant.

Table A3.2.3 US SI fund flow – alternative Treat$_i$ definition

	BP Oil Spill		Paris Agreement Shock		2016 US presidential election	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Post$_t$	63.79*	46.81	-17.92	-17.57	56.04**	52.83**
	(37.27)	(38.32)	(16.08)	(17.08)	(24.52)	(25.21)
Treat$_i$	64.05***	47.88**	19.82	23.97	2.22	4.21
	(20.04)	(21.72)	(16.56)	(18.03)	(14.67)	(15.50)
Post$_t$ × Treat$_i$	-23.05	-19.06	-5.92	-13.40	-6.10	-10.10
	(15.30)	(17.08)	(14.70)	(17.03)	(11.34)	(12.94)
Lagged_RTN$_{i,t}$	-0.81	-0.69	3.21***	3.21***	-0.65*	-0.62*
	(0.51)	(0.51)	(0.80)	(0.80)	(0.34)	(0.34)
RTN$_{i,t}$	0.18	0.09	0.92*	0.92*	0.08	0.05
	(0.40)	(0.40)	(0.47)	(0.47)	(0.26)	(0.26)
Age$_i$	-1.99**	-1.82**	-5.40***	-5.38***	-5.30***	-5.31***
	(0.82)	(0.82)	(0.70)	(0.71)	(0.51)	(0.51)
Lagged Fundsize$_{i,t}$	0.01	0.01	-0.01***	-0.01***	-0.01**	-0.01**

	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Expense Ratio$_{i,t}$	1.58	2.54	13.60	13.81	-10.43	-10.50
	(10.81)	(10.76)	(11.58)	(11.63)	(7.56)	(7.57)
Lagged $\alpha_{i,t}$	29.38	26.28	0.23	-0.17	68.16***	67.90***
	(27.62)	(27.51)	(29.21)	(29.34)	(15.92)	(15.94)
Lagged $\beta_{mkt,i,t}$	-75.83	-82.34*	-82.24*	-83.58*	11.03	10.65
	(46.63)	(46.55)	(47.85)	(47.93)	(33.48)	(33.70)
Lagged $\beta_{smb,i,t}$	-1.88	-9.28	19.34	18.86	23.50*	23.46*
	(23.77)	(23.89)	(29.62)	(29.67)	(12.04)	(12.10)
Lagged $\beta_{hml,i,t}$	51.95	48.83	59.36***	58.72***	19.58	19.66
	(37.32)	(37.28)	(22.50)	(22.55)	(20.47)	(20.57)
Lagged $\beta_{mom,i,t}$	-14.48	-33.75	183.41***	182.55***	60.55*	62.38*
	(65.23)	(65.47)	(64.98)	(65.34)	(33.80)	(34.05)
Institutional	N	Y	N	Y	N	Y
Fund Family Fixed Effects	Y	Y	Y	Y	Y	Y
Number of observations	333	333	842	842	688	688
Number of treated funds	76	76	148	148	134	134
Adjusted R-squared	0.312	0.320	0.281	0.280	0.343	0.342

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. BP oil spill shock: $Post_t = 1$ if year = 2010, 2011 and $Post_t = 0$ if year = 2009. US Paris Agreement withdrawal shock: $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2016. 2016 US presidential election shock: $Post_t = 1$ if year = 2015, 2016 and $Post_t = 0$ if year = 2014. In the BP oils pill shock, I employ 1(treated)-to-1(control) matching for environmental SI funds (treated) and non-environmental funds (control) without replacement. For Paris Agreement and 2016 election shock, I employ 1(treated)-to-1(control) matching for SI funds (treated) and conventional funds (control) with replacement. Control variables include 1-year lagged holding period return (Lagged_RTN$_{i,t}$), contemporaneous return (RTN$_{i,t}$), 1-year lagged fund size (Lagged Fundsize$_{i,t}$ in millions), expense ratio (Expense Ratio$_{i,t}$) and fund age (Age$_i$). Institutional dummy and its interaction term with $Post_t \times Treat_i$ are included in my specification, the coefficients are insignificant.

Chapter 4 : SI mutual fund portfolio choices and ESG controversies

4.1 Introduction

In this Chapter, I examine differences in portfolio choices between sustainable investing (SI) mutual funds and conventional mutual funds. In line with existing literature, I show evidence that although labeled differently, SI and conventional funds share similar financial characteristics. Additionally, anecdotal evidence suggests that SI funds do not change their portfolio holdings of firms involved in well-publicized ESG scandals. To examine whether SI funds commit to ESG, I employ a difference-in-differences approach and show that, compared to conventional funds, US SI funds do not appear to substantially increase or decrease holdings of ESG controversial firms and are likely to be motivated by concerns over ESG ratings and pecuniary interests than social preferences. The novelty of this research lies firstly in the use of controversies as an objective measure of failures to ESG standards rather than changes to ESG ratings. This is important because existing research has shown that ESG ratings are potentially noisy and subjective. Secondly, my analysis incorporates the common monetary motive shared by conventional and SI mutual funds by considering the market reaction to controversies. Moreover, whereas conventional funds are only committed to risk/return characteristics, SI managers may have to compromise returns for their mutual fund ESG principles. The research design also addresses endogeneity in portfolio decision making by SI/Conventional mutual funds by using firm-level "ESG compliance" measured by the number of past ESG controversies.

4.2 Related Literature and Anecdotal evidence

Differences between SI and Conventional mutual funds

The debate over whether SI funds differ from conventional funds in financial performance is, at its root, related to the long-standing question of "doing well by doing good" in corporate social responsibility (CSR). Better CSR attributes may be associated with higher firm value through more efficient resource use (Majumdar & Marcus, 2001), higher employee productivity (Korschun et al., 2014) and commitment (Mueller et al., 2012), enhanced corporate reputation (Fombrun & Shanley, 1990; Hong & Liskovich, 2015), increase in corporate goodwill and social capital (Waddock & Graves 1997; Lins et al., 2017). In addition, Hoepner et al. (2016) and Goss and

Roberts (2011) point out that better CSR performance leads to lower costs of capital; on the other hand, firms bear higher loan interest rates if there are high levels of ESG-related concerns (Chava, 2014).

From an investor's perspective, demanding higher CSR performance may generate abnormal returns (Dimson et al., 2015) or reduced level of portfolio risks (Albuquerque et al., 2018; Krueger et al., 2020), and employing an employee satisfaction-tilted portfolio construction strategy could lead to superior performance relative to market benchmarks (Edmans, 2011). Theoretical frameworks and empirical evidence also suggest that stock prices react to investment in CSR activities when some market participants hold pro-social tastes or preferences (Friedman & Heinle, 2016). On average, investors "punish" firms with adverse CSR events (Krueger, 2015) while holding expectations of higher future performance on investment opportunities with better sustainability attributes (Hartzmark & Sussman, 2019).

However, an SI investment portfolio may be less diversified due to the employment of negative screening strategies. Negative screening strategies refer to using CSR or ESG-related constraints in the stock selection process, for example, avoiding "sin stocks" such as tobacco companies' shares. The strategy is used by 69% of US sustainable investment fund managers (US SIF, 2020). According to Markowitz's (1952) portfolio theory, under pro-social constraints, investors face a smaller investable universe than passive investors leading to suboptimal risk-adjusted returns. Consistent with this theory, an average US SI fund earns a negative Fama-French four-factor adjusted alpha of -3.4% per annum (Renneboog et al., 2008), and investing in mutual funds with socially responsible constraints leads to losses ranging from a few basis points to at least 30 basis points per month, depending on investors' belief of the appropriate market benchmarks and their reliance on funds' track record as predictors for future performance (Geczy et al., 2021). When examining US mutual funds, earlier researchers such as Statman (2005) also find that SI fund performance is no different from that of conventional funds. Hong and Kacperczyk (2009) show that "sin stocks" may be initially underpriced due to lower demand from SI investors and are considered relatively risker assets. Thus, non-SI investors require higher returns as compensation. By the same token, the model in Heinkel et al. (2001) shows that avoidance of polluting firms by ethical investors creates downward pressure on stock prices and, thus, higher expected investment returns. A more recent study by Bolton and Kacperczyk (2021) show higher stock returns for firms with higher levels of carbon emissions using a global sample of US, European and Asian companies; in the US, this carbon premium is a result of the divestment of polluting firms by institutional investors.

Overall, the existing literature on socially responsible investment provides some evidence that firm-level CSR may be associated with higher firm value through operational efficiency, better stakeholder relationships, lower risks, goodwill, and reputation. Nevertheless, applying ESG-based screenings or incorporating ESG information in portfolio construction and management does not consistently induce significant differences in investment returns between SI and conventional mutual funds (Liang & Renneboog, 2020).

Following this body of literature, I first show the similarities in the risk-return profiles of US SI and conventional funds over time. I define SI versus conventional using the Morningstar data point: "Sustainable Investment – Overall". I begin by plotting fund total return and fund return standard deviations (fund total risk) by year in Figure 4.1. I show that most SI funds lie within the "efficient frontier" and overlap with many conventional funds, which indicates that, consistent with existing literature, most SI funds may be similar to risk-return optimized compared to conventional mutual funds.

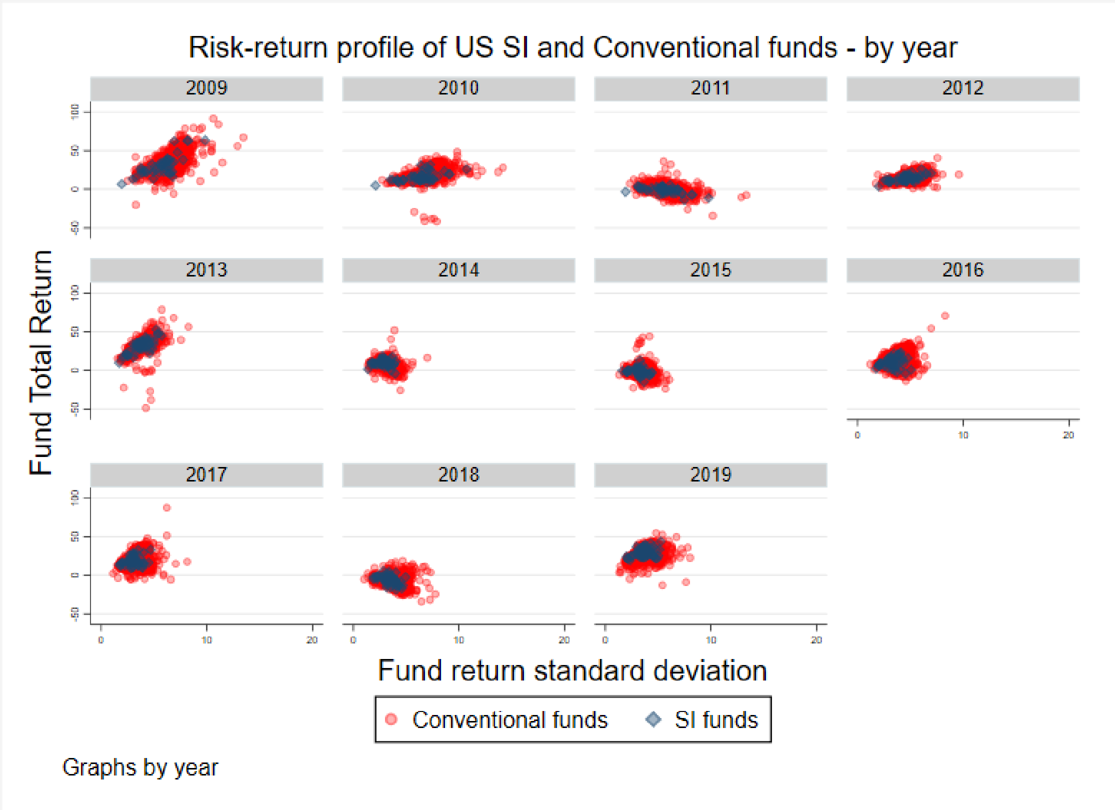


Figure 4.1 Total Return – Standard Deviation Plot of SI vs. Conventional Funds overtime

In addition, I show that SI and conventional funds are not too dissimilar in terms of the distribution of their systematic risks (Figure 4.2) and risk-adjusted returns, as indicated by Sharpe ratios (Figure 4.3). To add further dimensions and formality to the analysis, I run logit regression models to estimate propensity scores of being an SI fund using 1-year lagged fund financial characteristics: fund return, Fama-French four-factor loadings, expense ratio, and turnover rate. Figures 4.4 and 4.5 illustrate that there is substantial overlap in the distribution of SI and conventional funds along the propensity score scales. Therefore, there is a high degree of overlap in the distributions of SI and conventional funds in a single measure combining multiple fund-level financial characteristics in terms of multiple fund-level financial characteristics. The only notable difference is the greater mass of conventional funds in the lower tail, indicating that more conventional funds are predicted to be SI rather than actual SI funds. The above descriptive evidence points to the absence of any notable differences in the characteristics and performance of conventional and SI funds.

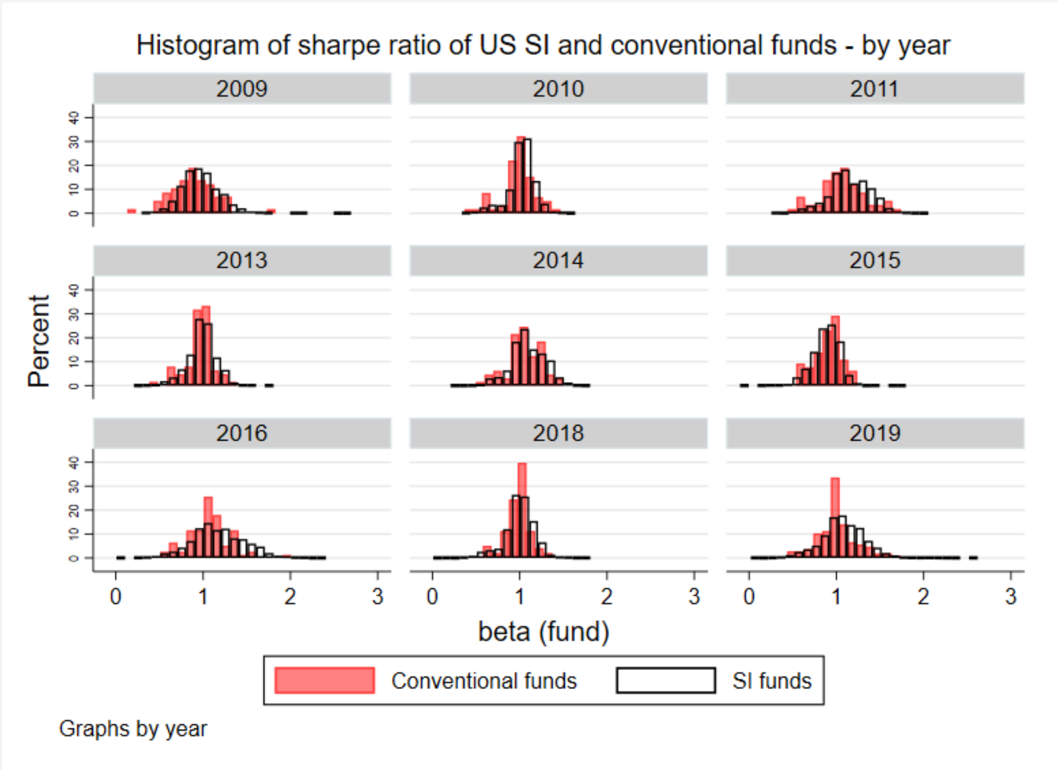


Figure 4.4 Distribution of beta of US SI and conventional mutual funds

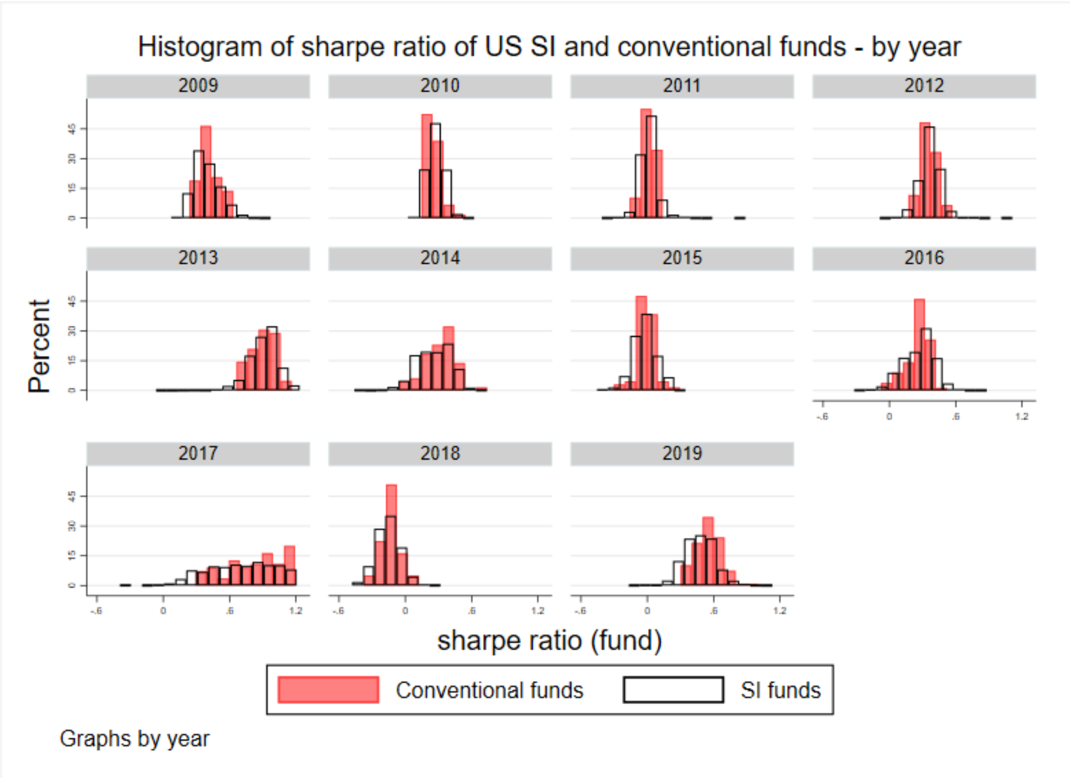


Figure 4.5 Distribution of Fund Sharpe Ratio of US SI and conventional mutual funds

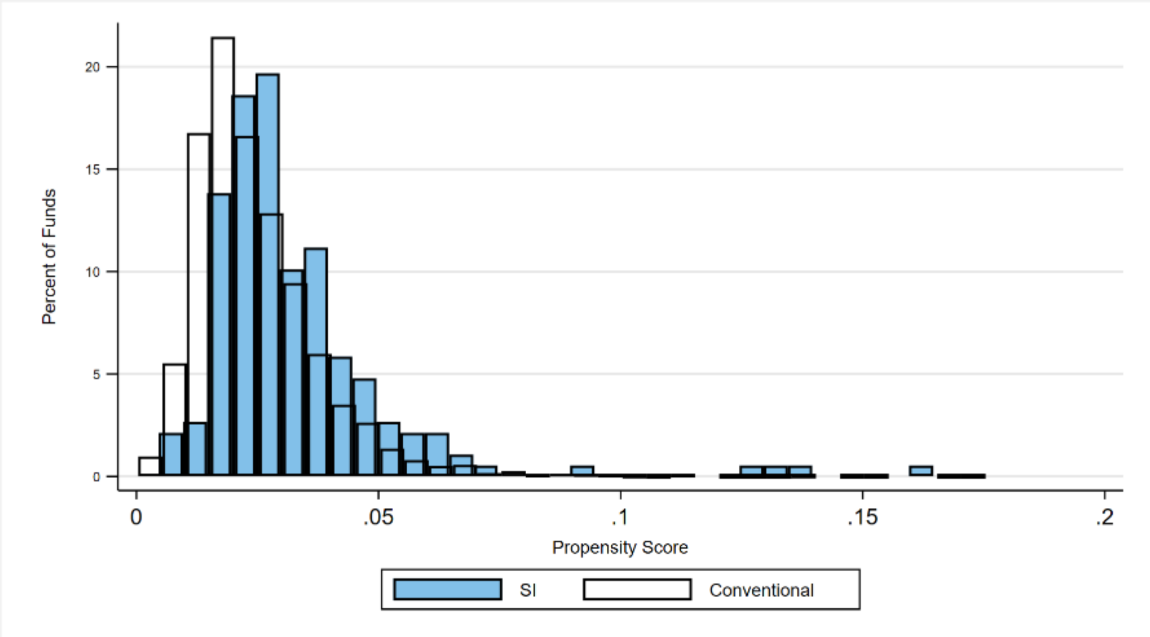


Figure 4.6 Logit propensity score matching of SI and conventional funds using: 1-year holding period return, FF 4-Factor loadings, turnover, expense ratio, as of 2016

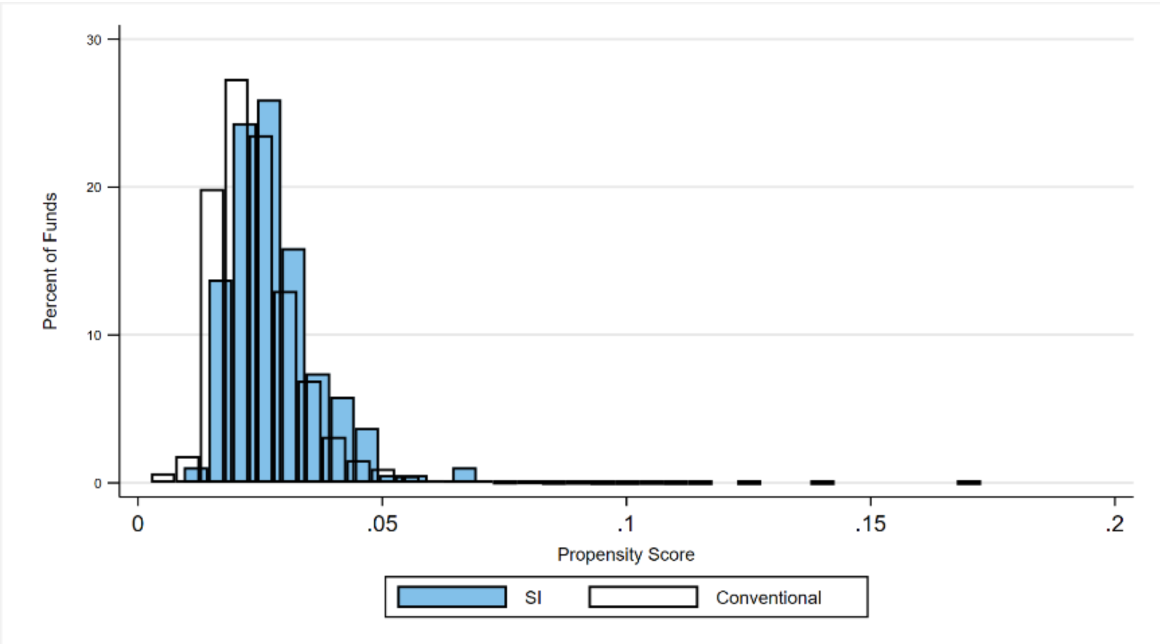


Figure 4.7 Logit propensity score matching of SI and conventional funds using: 1-year holding period return, FF 4-Factor loadings, turnover, expense ratio, as of 2017

Although the financial characteristics of SI and conventional funds are similar, one cannot capture how the exact holdings of SI funds and conventional funds differ and whether there are notable differences in their portfolio rebalancing decisions. In this Chapter, I conduct formal tests using firm-level ESG controversies. Before presenting my methodology, I first discuss existing literature on holdings of SI funds.

Holdings of SI funds and ESG performance

The inconclusive debate about SI performance relative to conventional funds could result from identical underlying assets held by SI and conventional funds. However, as discussed earlier, SI funds have a history of employing exclusionary screening for religious considerations (Statman, 2005) and avoiding controversial industries such as tobacco and alcohol to align portfolio composition with social values or beliefs (Hong and Kacperczyck, 2009).

Nonetheless, researchers also present evidence that such practices are not restricted to funds explicitly labeled as “socially responsible”, as non-SI funds consider non-pecuniary preferences of clients such as public pension funds (Hong and Kacperczyck, 2009), or when they are subjected to pro-social local political or religious environment (Borgers et al., 2015). Bello (2005) offers empirical evidence that compared to conventional funds, mutual funds that employ exclusionary screening do not hold substantially different assets, nor do they have a significantly lower degree of portfolio diversification and concentration. These portfolio similarities are also associated with no difference in investment performance (Bello, 2005). A study by Borgers et al. (2015) also illustrates how some US conventional funds, especially those with large institutional clientele, are similar to their SI counterparts and have zero or little exposure to “sin” stocks from 2004 to 2012.

With recent developments in SI markets, SI strategies other than exclusionary screening have emerged, for example, ESG integration and thematic investments (Amel-Zadeh and Serafeim, 2018; GSIA, 2018). Theoretical papers such as Landier and Lovo (2020) and Oehmeke and Opp (2020) discuss how socially-minded investors could make an impact and reduce negative externalities. Both papers argue that divestment from polluting firms, which corresponds to negative screening SI strategies, does not necessarily improve social welfare. Instead, pro-social investors such as SI funds should allocate their scarce socially responsible capital based on the amount of “avoided externalities.” (Oehmeke and Opp, 2020). In other words, if sustainable funds aim to reduce negative externalities, they may need to prioritize investments in sectors or firms that are more polluting or controversial to improve societal welfare (Landier & Lovo, 2020). As such, one may argue that both SI and conventional funds should have stakes in the same poor-ESG-performing organizations, making it possible for both SI and Conventional funds to react to ESG-related controversies by increasing their holdings in those firms involved in such controversies. On the other hand, Chowdhry et al. (2019) propose a model of contracting designs that organizations can use to align profit-motivated and socially-motivated investors. They argue that conventional, for-profit, and socially driven investors should hold opposite contracts, and socially driven investors are guaranteed payment when organizations or projects fail to deliver social values.

Most of the theoretical papers discussed above are formed upon the assumption that SI mutual funds and investors either have non-pecuniary interests, i.e., they hold different "taste", or preferences compared to conventional funds, or that they perceive CSR or ESG to be value-relevant. Such assumptions imply that SI and conventional funds may hold different portfolios. Using a sample of US mutual funds, I show in Table 4.1 that the number of firms held by both SI and conventional funds between 2009 and 2019 represents over 70% of all unique firms in my sample from 2009 to 2019.

Table 4.1 Fund portfolio firm overlap between SI and Conventional

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Firms held by both	379	336	329	357	362	365	465	452	484	491	475
Total number of unique firms	506	456	470	496	504	525	626	601	595	591	563
Common holdings %	75%	74%	70%	72%	72%	70%	74%	75%	81%	83%	84%

This table reports the number of unique US-listed firms held by both SI and conventional funds each year between 2009 to 2019. The last row of common holdings % is computed as the number of firms held by both (first row) divided by the total number of unique firms (third row).

In addition, Table 4.2 below illustrates that SI funds do seem to select stocks from a smaller universe (329 to 491 stocks compared to 456 to over 600 stocks from the conventionally held firm universe), and in most years, an average SI fund owns a smaller number of firms in their portfolio. Compared to conventional investors, SI investors also hold lower stakes in any given firm (measured by shares held by SI funds over total shares outstanding of the firm). Both types of funds hold firms from over 40 industries. In general, the observed differences are small in magnitude, which is consistent with prior literature, indicating that SI and conventional funds may not be easily distinguishable based solely on the stocks in their portfolio.

Table 4.2 SI and Conventional mutual fund holding characteristics by year

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
SI funds											
Stock selection universe	379	336	329	357	362	365	465	452	484	491	475
Average number of firms held	39	33	34	38	35	35	50	54	53	56	52
Portfolio industry coverage	41	40	41	40	41	40	43	43	42	43	42
Average stake in firms (%)	0.07	0.07	0.07	0.07	0.07	0.07	0.06	0.05	0.04	0.04	0.03
Conventional funds											
Stock selection universe	506	456	470	496	504	525	626	601	595	591	563
Average number of firms held	46	36	37	41	38	40	53	52	52	55	52
Portfolio industry coverage	42	41	42	41	43	42	43	43	42	43	42
Average stake in firms (%)	0.19	0.20	0.20	0.19	0.20	0.19	0.16	0.16	0.15	0.14	0.13

This table reports the characteristics of firms held by SI and conventional mutual funds in my sample. The row "Stock selection universe" counts the total number of unique firms held by any SI (conventional) fund. The row "Average number of firms held" displays the average number of holdings by a SI (conventional) fund. The row "Portfolio industry coverage" is the number of unique industries held by any SI (conventional) funds. The average stake in firms (%) is calculated as the average percentage of firms' shares held by any SI (conventional) fund out of the total shares outstanding of firms.

On the other hand, portfolio differences could also be manifested through ESG performance. For example, a recent paper by Gibson et al. (2022) use a global sample of mutual funds, ESG rating data from Refinitiv and MSCI combined to evaluate the ESG performance of United Nations Principal of Responsible Investment (PRI) signatories (institutional investors who publicly commit to incorporate ESG issues in analysis and selection of investments). They find that PRI signatories display better "ESG footprints" (measured by portfolio ESG ratings) than non-PRI investors. However, the authors do not observe this pattern among US institutional investors and highlight that US PRIs with worsened ESG scores after signing PRI are also those who perform poorly, cater to retail clients, and have low ESG profiles in their own fund operations. Gibson et al. (2022) thus point out that these US PRI investors may be involved in greenwashing instead of committing to ESG. Kim and Yoon (2022) use firm-level ESG scores from MSCI, Sustainalytics, and TruValue Labs, construct fund-level ESG performance measures for PRI and non-PRI funds, and document PRI signatories holding assets that lead to high fund-level ESG score than non-PRI funds.

In both Gibson et al. (2022) and Kim & Yoon (2022), the key dependent variable is fund-level ESG performance, constructed using firm-level ESG ratings. In the following subsection, I discuss potential concerns regarding ESG ratings.

ESG rating divergence and discrepancies

As described above, existing research that attempts to differentiate SI and Conventional funds in their ESG performance and ESG commitment use portfolio ESG scores as the dependent variable, where weighted portfolio ESG scores are constructed from firm-level ESG ratings offered by third-party databases. Some of the most prominent firm-level ESG data vendors include Refinitiv, previously known as Thomson Reuter's Asset4 ESG rating), Sustainalytics (now owned by Morningstar), MSCI ESG rating (formerly KLD ESG data), Vigeo Eiris, and Bloomberg ESG data. A survey by Amel-Zadeh and Serafeim (2017) shows that ESG ratings are used by investors managing assets of $30 trillion when making investment decisions.

Nonetheless, growing evidence from the literature on ESG ratings suggests that these ratings should be used with caution. One of the most well-documented facts on ESG ratings is how ratings from data providers do not converge. While credit ratings are highly correlated between credit rating companies (e.g., S&P and Moody's credit rating correlation is approximately 0.99 according to Berge et al., 2022), the top six rating agencies: Refinitiv, Moody's ESG, Sustainalytics, KLD, MSCI, and S&P Global have an overall correlation on their ESG scores is only 0.54 (Berg et al., 2022). When examining each "pillar" score individually, the authors find that a low level of correlation persists for the Environment pillar score; the correlation is 0.54. The correlation for social and governance pillars are 0.42 and 0.3, respectively. Chatterji et al. (2016) also find a lack of convergence between raters and show that raters disagree not only on their theorization of corporate social responsibility (i.e., their definition of CSR) but also on measuring methodologies applied to the same underlying metric. Consequently, even for the same company in question, the ESG ratings it receives from different providers cannot be meaningfully compared. Semenova and Hassel (2015) show that in terms of environmental performance ratings, although KLD (now MSCI EST STATS) and ASSET4 (Refinitiv) share common scope and dimensions of environmental performance (strengths in KLD) and risk (concerns in KLD) considerations, they do not converge on the aggregate level. The authors suggest that researchers should follow prior studies such as Mattingly and Berman (2006) and use non-convergent metrics separately instead of constructing combined metrics that confuse interpretations. In other words, combining different ratings will introduce noise due to measurement and scope divergence between raters (Berg et al., 2022), and the resulting combined rating will mask firms' actual ESG fundamentals and confuses interpretations.

In addition, there may also exist a "rater effect": under the same rating agency, companies receiving high ratings or scores in one ESG category, for example, the environment pillar, are also more likely to receive high ratings in other categories (Berg et al., 2022). A more recent study by Gibson et al. (2019) documents the importance of ESG data providers' geographical location on rating divergence. Raters based in civil law countries tilt their focus towards labor and social issues, and raters from common-law countries are more concerned with shareholder rights and general governance topics.

Based on data definition documents from Refinitiv and webinars by Morningstar's Sustainalytics, I find that both data providers' rating processes involve some subjectivity elements, for example, in their aggregation process of combining indicators into overall or pillar scores.[8] Furthermore, as evidenced by Amel-Zadeh and Serafeim (2017), non-SI investors may consider ESG issues for their risk-return relevance, creating a larger potential clientele for ESG data vendors. Consequently, data vendors may attempt to design their ESG ratings in a way that allows one to make a financial case on ESG issues (Bouten et al., 2017).

The difficulty in using ESG ratings is heightened by the possibility that raters may modify their ESG rating methodologies over time. For example, in September 2018, Morningstar Sustainalytics launched a new "ESG risk" rating. The ratings made changes to the original metrics in several ways, such as accounting for inherent differences in ESG risks between industries and inverting the measurement scale so that a low ESG rating represents low ESG risk exposure. Although Morningstar kept its original data for legacy users, there is no direct mapping available between the two versions. This implies that data users cannot compare Sustainalytics ESG ratings over time to assess whether there is any actual ESG performance change. Another ESG data provider, Refinitiv, also performed a methodology rewriting in April 2020. During this rewriting process, Refinitiv also retroactively modified historical ESG scores in

[8] See for example a webinar hosted by Sustainalytics for corporate issuers discussing their rating and review process here: https://www.sustainalytics.com/esg-research/resource/corporate-esg-blog/your-company's-esg-ratings-understanding-sustainalytics-research-process

their database, affecting around 6000 firms and 29,828 firm-year observations (Berg et al., 2021). Comparing Refinitiv ESG data downloaded before and after the rewriting, Berg et al. (2021) document an average decrease in overall ESG score of -20.6% across the database, which originates predominantly from downgrades on the Environment pillar score (-47.4%). They also argue that these changes might have been "data-mined" to reconcile firstly with the ESG performance of specific firms in specific historical years and secondly with firms' realized stock returns. This finding is consistent with Bouten et al. (2017), who argues raters produce scores to satisfy investors' need to make value-relevant predictions with ESG data.

Overall, despite ESG ratings being used widely by investors and researchers (Berg et al., 2021), substantial gaps remain in consolidating standards, processes, and measurement methods. Therefore, whether ESG ratings reflect firm-level ESG fundamentals such as controversies is unclear.

I provide preliminary evidence on this question by plotting ESG score changes after well-publicized corporate scandals. Figures 4.8 and 4.9 below show two well-known corporate environment-related scandals: the BP oil spill in 2010 and the Volkswagen emission scandal in 2015. As detailed in Chapter 2, the BP oil spill represents one of the most damaging man-caused disasters to the environment. However, Figure 4.8 shows that after the disaster, BP saw no downgrades in its overall ESG score and its Environment pillar score, but both scores increased in the year following the incident. Figure 4.9 illustrates the ESG score movements after the Volkswagen emission standard scandal in 2015. The scandal led to millions of cars worldwide being recalled, approximately $18 billion in fines, and a first-in-15-years quarterly loss of $2.5 billion (Hotten, 2015). In Figure 4.9, I observe a slight decrease in the overall ESG ratings at the event year and an increase in the Environment pillar score one year after the scandal. The two pieces of anecdotal cases appear to conform with existing literature in two ways, Refinitiv may have retrofitted the ESG scores to reflect business implications of ESG scandals, and overall ESG scores are constructed in an opaque manner that may fail to account for "actual ESG" performance indicators such as controversies and scandals.

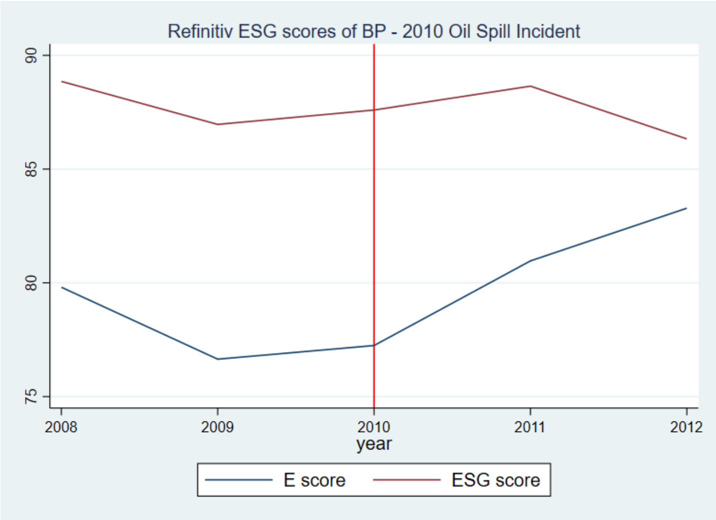


Figure 4.8 BP Oil Spill and Refinitiv ESG Score change

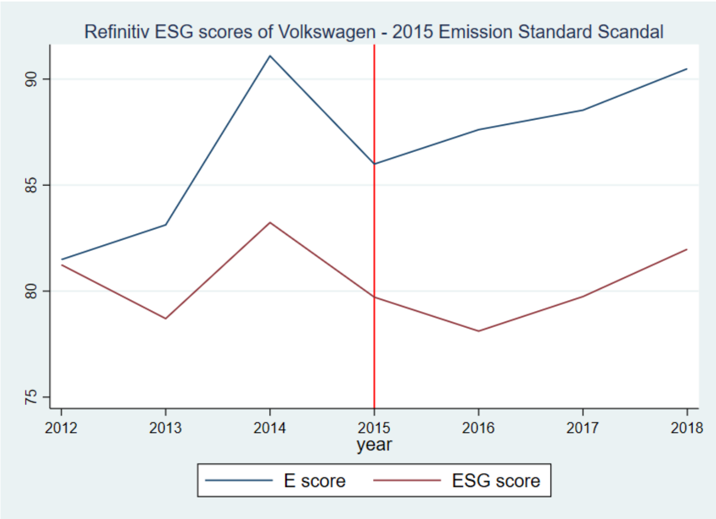


Figure 4.9 Volkswagen Emission Standard Scandal 2015 and Refinitiv ESG Score change

To avoid such rating ambiguity and ensure I measure the actual ESG performance of firms and how mutual funds react to such changes as objectively as possible, I abstain from using ESG rating scores at the firm and fund levels. Instead, I resort to firm ESG controversy events and portfolio composition of mutual funds and employ a difference-in-differences approach to test whether SI mutual funds and conventional funds react differently to firm-level controversies.

In this Chapter, I complement these studies and examine whether mutual fund portfolios are rebalanced in a way that would reflect their social preferences after corporate ESG controversies take place. I first give anecdotal evidence around two well-known ESG controversies and then move to a formal analysis.

In 2010, it was reported that 15 workers at Foxconn committed suicide, most of which took place in May. As the assembly factory for Apple's iPhone products, Foxconn's incidents raise questions about Apple's supply chain ESG management and assembly line employee welfare; the event represents a "social" controversy.

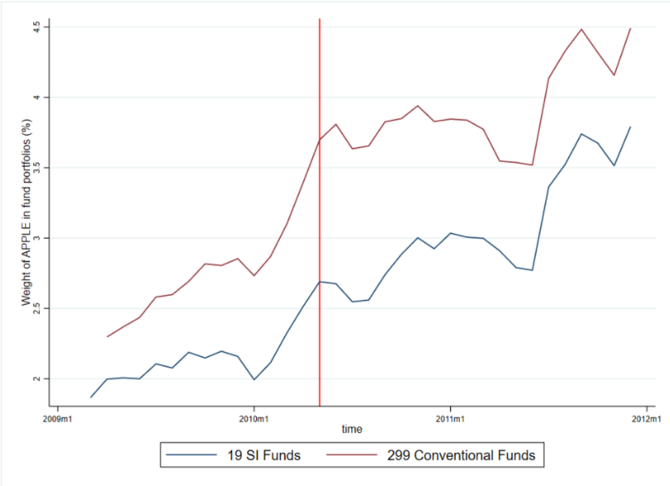


Figure 4.10 Monthly portfolio weight of AAPLE around Foxcoom suicide incidents (Mar 2010)

In Figure 4.10, I show that the average portfolio weight of Apple among 19 SI and 299 conventional funds follows similar patterns before and after the scandal. In other words, both funds react similarly to this social controversy. Another recent example of a well-known ESG scandal is related to Facebook. In March 2018, it was reported that Cambridge Analytics obtained Facebook users' data without consent for political advertising (Rosenberg et al., 2018). Facebook's CEO, Mark Zuckerberg, was subsequently called for hearings at the US Congress in April (Watson, 2018).

Figure 4.11 Monthly portfolio weight of Facebook around the Cambridge Analytica scandal (Mar 2018)

From Figure 4.11, I observe that the average holdings of Facebook for 41 US SI funds and 713 US conventional funds both increased and followed similar trends after the shock. Put simply, fund managers of SI and conventional funds reacted similarly to this governance controversy. Both examples indicate that there may be some similarity between SI and conventional fund managers' decisions to adjust the weights of their holdings. I formally analyze SI and conventional fund holdings reaction to controversies in the following sections of this Chapter.

4.3 Empirical Design

4.3.1 Methodology

Compared to ESG ratings which may be subject to various endogeneity concerns, one advantage of using controversy counts is that they are objective indicators of what is taking place in firms with respect to environmental, social, and governance issues. Although these ESG controversies are likely independent of mutual funds, they are mostly not exogenous to firms. At the same time, some incidents took place due to random human errors, such as an operational error in handling equipment or following safety protocols; other controversies, such as accounting scandals or shareholder rights violations, may stem from management oversight, weak governance, or interest in cost reduction for higher profitability relative to competitors. In this case, some ESG controversies may occur due to firms' bottom-line considerations and subsequently affect stock prices once the controversies become known to market participants. In addition, regardless of a "Sustainable Investment" label, mutual funds are investment products that warrant a financial reward. Therefore, if I do not find significant differences between SI and conventional portfolio rebalance decisions after controversies, one could argue that although SI funds may hold non-pecuniary considerations, the valuation implications of controversies may cancel out social preferences when deciding to rebalance portfolios. In this Chapter, I test whether SI and conventional funds trade stocks in their portfolio after controversies take place, whether they trade in different directions, and whether any difference in rebalancing decisions depends on the direction of market reactions to controversial ESG events.

One feature of Refinitiv's ESG controversy counts lies in the granularity of different controversies. There are 23 types of controversy metrics available for each firm-year observation, each type captures a different ESG aspect. For example: child labor controversy, employee health and safety controversy, intellectual property controversy, shareholder rights controversy etc. As such, it is natural to observe that some firms covered by Refinitiv have multiple types of controversies in a given year (for example, environment resource impact controversy, anti-competition controversy, and business ethics controversy at the same time). In addition, for the same type of controversy (e.g. business ethics controversy), one may observe multiple incidents of it in the same year. In other words, the Refinitiv ESG controversy dataset contains: types of controversies, and for each type, an annual count of that controversy.

When multiple types of controversies occur in the same year, it becomes impossible to assign any portfolio changes or market reactions to one particular controversy as impacts of different controversies contaminate each other. . Therefore, when defining my "shock" or "event", I ensure that for each firm-year observation, the "treated" firms in my sample are those with non-zero controversy count for only one type of controversy in any particular year, and therefore firm-year observations with multiple types of controversies are excluded from the sample.

To establish whether SI and conventional funds react differently to ESG controversies, I specify two difference-in-differences models which can be written in the following functional forms:

$$ShrsHeld_{Csho_{i,j,t}} = \alpha_i + \eta_t + \psi_j + \beta_1 SI_i + \beta_2 Controv_{,t} + \beta_3 Controv_{j,t} \times SI_i + \delta FundControl_{i,t} + \vartheta FirmControl_{j,t} + \varepsilon_{i,t} \quad (4.1)$$

$$ShrsHeld_{Csho_{i,j,t}} = \alpha_i + \eta_t + \psi_j + \beta_1 SI_i + \beta_2 Sign_{j,t} + \beta_3 Sign_{j,t} \times SI_i + \delta FundControl_{i,t} + \vartheta FirmControl_{j,t} + \varepsilon_{i,t} \quad (4.2)$$

In Equation (4.1), the variable $Controv_{j,t}$ is a dummy variable that takes the value of 1 if a firm j experienced non-zero counts of controversies in year *t* regardless of any stock market reaction to that controversy.

In Equation (4.2), I propose an alternative specification. This specification tests if SI funds make portfolio rebalancing decisions based not only on the fact that firms experience ESG controversies but also on the valuation implications of ESG controversies. To do so, I introduce the variable $Sign_{j,t}$ to indicate the potential for joint consideration of financial and social concerns in portfolio rebalancing.

Specifically, to construct the variable $Sign_{j,t}$, I follow standard event study procedures and compute cumulative abnormal return (CAR) for each firm-year-controversy observation. Firms' daily return data are obtained from Compustat via WRDS and US market return data from the Fama-French daily factor return dataset is used as my market benchmark. Once a firm is identified as a "treated firm" in year t (i.e., non-zero number of controversies of any one particular type of controversy, but zero number of controversies for all other types in year t), I use the first ESG controversy source date disclosed by Thompson Reuters in year t as the event date. Under a 250-day estimation window and a 20-day event window setting, I estimate CARs and standard deviations of CAR, and determine the significance of the CARs for all "treated firms" in my sample.

As a result, firms in my sample can now be categorized in to four groups. First, there are firms with zero controversies whatsoever in year *t* (i.e., zero count of incidents across all 23 types of controversies in year *t*), this group of firms is therefore defined to be the 'baseline' firms in the regression model with $Sign_{j,t}$ equal to 0. The second group of firms are those who are "treated" (i.e. has non-zero controversies of one particular type in year *t*), but the controversy does not lead to any significant market reactions as measured by CAR. These "controversy but has insignificant CAR" firms has $Sign_{j,t}$ equal to 1. In similar fashion, the other two groups of firms are: "controversy but has positive CAR" firms with $Sign_{j,t}$ equal to 2, and "controversy but has negative CAR" firms with $Sign_{j,t}$ equal to 3. To summarize, the variable $Sign_{j,t}$ captures whether firm *j* experienced controversies in year *t*, and if so, the corresponding market reaction.

In both DiD models, my dependent variable is $ShrsHeld\_Csho_{i,j,t}$: annual average proportion of shares of firm *j* held by fund *i* in year *t*. To obtain this variable, I first compute the monthly proportion of shares held for each fund-firm-year-month observation $ShrsHeld\_Csho_{i,j,t,m}$ as:

$$ShrsHeld\_Csho_{i,j,t,m} = 100 \times \frac{ShrsHeld_{i,j,t,m}}{Csho_{j,t,m}} \tag{4.3}$$

In Equation (4.3), $ShrsHeld_{i,j,t,m}$ is the number of firm *j*'s shares held by fund *i* at year *t* month *m*, and $Csho_{j,t,m}$ is the total number of common shares outstanding for company *j* at year *t* and month *m*. The dependent variable in my regression model is the annual average of this monthly measure of funds' stakes in companies. SI_i is a dummy variable equal to 1 if fund *i* is identified by Morningstar as "Sustainable Investment" and 0 otherwise. α_i are fund-group fixed effects, which is the Morningstar global category a fund belongs to, such as: US Equity Large Cap Growth, Aggressive Allocation. I also include year fixed effects η_t., industry fixed effects, fund size decile group fixed effects and fund family fixed effects. Fund family fixed effects controls for unobservable differences in investment styles and expertise across fund management companies. Fund size decile fixed effects rule out unobserved operational, reputational, and visibility differences between large and small funds.

Following prior literature studying SI funds, such as Bauer et al. (2005), Renneboog et al. (2008), and Borgers et al. (2015), the variable $FundControl_{i,t}$ includes $Concentration_{i,t}$ (portfolio concentration measured as aggregate assets, expressed as a percentage, of top 10 portfolio holdings); $FundSize_{i,t}$ (fund-level total net asset in millions of USD); $SharpeRatio_{i,t-1}$ (1-year lagged fund Sharpe ratio); $FundReturn_{i,t}$ (fund total return in year t); $FundRisk_{i,t}$ (standard deviation of monthly fund returns for the past 36 months from the end of year t.); $Expratio_{i,t}$ (annual report net expense ratio expressed as a percentage of fund assets), and a dummy $Institutional_i$ to indicate institutional funds. It is worth noting that the type and number of incidents experienced by firms may also differ by industry. For example, oil and gas companies naturally face higher risks of spills and leakages than financial institutions. Survey results from Amel-Zadeh and Serafeim (2017) suggest that portfolio managers consider firms in different industries to face systematically different ESG risks. To account for industry effects, I use industry fixed effects in my regression specifications.

Considering prior studies on mutual fund stock holdings, such as Falkenstein (1996), I define $FirmControl_{j,t}$ as a vector of firm-specific control variables including the following firm characteristics: $Volatility_{i,t}$ (variance of firms' monthly returns over the past 24 months from the end of year *t*), $Size_{i,t-1}$ (1-year lagged total net assets), $MktVal_{j,t-1}$ (1-year lagged market value in \$ millions), $SalesGrowth_{j,t}$ (growth in sales per share from the previous year in percentage), $ROA_{j,t-1}$ (1-year lagged return on assets), $DebtAsset_{j,t-1}$ (1-year lagged debt to asset ratio) and $FirmSharpeRatio_{j,t-1}$ (1-year lagged firm Sharpe ratio). When estimating my regression model, I cluster standard errors by firm-year to account for potential dependence across funds from the same firm-year combination.

4.3.2 Data and sample

In this Chapter, I obtained firm-level ESG controversy data from Thompson Reuters Refinitiv ESG database, previously known as ASSET4 ESG. Refinitiv ESG database provides more than 630 different ESG-related metrics. These metrics include the overall ESG ratings and ESG pillar scores (Environment, Social, and Government scores) and cover a total of 12,000 public companies globally that represent over 80% of the global market cap. Figure 4.12 summarizes the process of how Refinitiv uses its 630 data points to construct the ESG overall score. The 630 ESG metrics represent the most granular data points available in Refinitiv, while another 23 ESG controversy data points are used to construct Refinitiv's ESGC score, which is the weighted average of ESG scores and ESG controversy scores (Refinitiv, 2022).

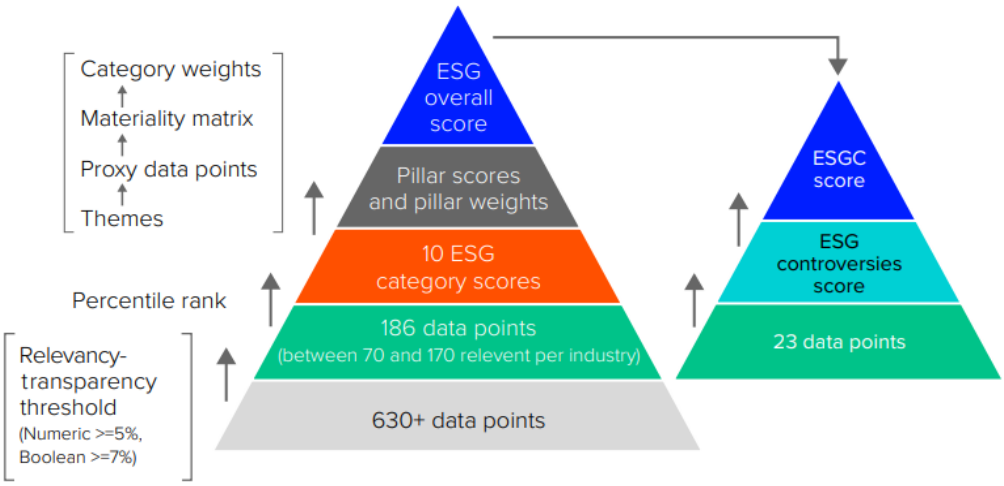


Figure 4.12 Refinitiv ESG scoring methodology process flow (source: Refinitiv)

In this chapter, I use all 23 ESG controversy data points, which cover the following aspects: Community, Human Rights, Product Responsibility, Resource use, Shareholders, and Workforce. These data points reflect controversies or scandals experienced by firms in the current year, as well as new developments such as lawsuits or fines related to past adverse events., For example, *TR.ControvEnv* counts the number of controversies published in the media relating to the impact of the company's operations on the environment and natural resources. For all controversy measures, the default value is 0. I download data on all 23 controversies from 2009 to 2019 for all companies covered by the Refinitiv database. These firm-level ESG controversies occur or do not occur at any point in time. Therefore, unlike ESG rating scores (typically from 0 to 100), they are not subject to ambiguous and potentially subjective criteria issued by rating agencies and analysts. Furthermore, investor demand for firms' shares will cause ESG controversies. Thus, I argue that ESG incident counts are exogenous to fund holdings. A summary tabulation of all controversies obtained from Refinitiv is available in my appendix.

Morningstar Direct is my primary data source for mutual fund financial and portfolio composition data. Following mutual fund and sustainable investing literature, such as Carhart (1997) and Benson and Humphrey (2008), I exclude real estate, money markets, commodities, precious metals, bond funds, ETFs, sector funds, international funds, real estate funds and funds with less than 60% in equity. In addition, I scrutinize my mutual fund sample for irregularities to detect potential data errors. Data errors discovered include funds with more than 100% total portfolio weighting; more than 100% portfolio weight on multiple stocks; negative portfolio weighting in the top 10 holdings; a fund that holds only three real-estate company stocks but is not identified as a real-estate sector fund by Morningstar.

For each fund *i*, I obtain the top 100 stock holdings and their corresponding monthly number of shares held by fund *i* between 2009 and 2019 from Morningstar Direct. From communications with Morningstar Direct, I understand that most mutual funds voluntarily disclose portfolio compositions to Morningstar monthly or quarterly; however, they may also hold off disclosing holdings information for a certain period of days and resume reporting to Morningstar at a later stage. Morningstar does not provide additional information as to why mutual funds might act in this manner. The heterogeneity in reporting frequency and consistency leads to an imbalanced monthly panel dataset of holdings. After deleting funds with no holdings records from 2009 to 2019 and dropping observations that appear in funds' holdings records but cannot be identified due to a lack of an identification code in Morningstar, my initial mutual fund holdings dataset consists of 3,012 US mutual funds, and 14,496 publicly traded firms.

I use the companies' International Securities Identification Number (ISIN) code and Ticker to merge the above Morningstar fund holdings data with the Refinitiv ESG controversy count dataset and Refinitiv ESG controversy source date dataset. Since the same firm can have multiple ISINs, and multiple firms may share one Ticker code, I verify that the firms in my sample are those listed on the US and North American stock exchanges, including New York Stock Exchange, Nasdaq Stock Exchange.

In the Refinitiv ESG dataset, I am able to match a sample of 798 unique US firms that also appear in the mutual fund holdings dataset. However, a few discrepancies exist between Refinitiv's ESG controversy date information and its controversy count data. For example, given that some firm-controversy type-year observations with a positive controversy count do not have a corresponding controversy source date entry, I exclude these observations from my sample. I further clean my data by dropping firms that do not have clear industry peer group designation given by Refinitiv and eliminating duplicate records of the same fund-firm-year-month observation. My final sample contains 751 US-listed companies from 43 Refinitiv ESG peer groups (such as Software & Computer Services; Oil, Gas, and Coal) that were held at different points from 2009 to 2019 across 118 US SI funds and 2200 US conventional funds.

For each firm, I obtain the number of shares outstanding from Morningstar Direct, Compustat via WRDS, and Thompson Reuters at a monthly frequency. I download the same variable from multiple databases, given that certain databases have limited data coverage on firms. This is especially the case for Morningstar Direct, which is predominantly a mutual fund database. Compustat and Thompson Reuters offer similar data coverage on the monthly number of shares outstanding. In this Chapter, I resort to Thompson Reuters as it is also the primary source of other firm-level financial data.

Table 4.3 below summarizes the financial characteristics of firms in my sample. I observe that in most years, there are fewer firms that SI funds hold compared to conventional funds in my sample. In an average year, there are 408 firms held across different SI funds, while 539 firms are held across different conventional funds. This is evidence consistent with Table 4.2, which illustrates that SI funds typically select stocks from a smaller pool of assets. In addition, SI funds also appear to hold larger firms in terms of the total net asset (*Size*) and market capitalization (*MktVal*). Firms held by both types of funds are similar in their risk-adjusted returns measured by Sharpe ratio (*FirmSharpeRatio*), but firms in SI funds have higher rates of return to asset (*ROA*).

Table 4.3 Firm summary stats

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
SI funds											
Size	19.3	14.7	16.8	18.5	16.3	20.9	28.5	27.9	26.2	28.0	32.5
MktVal	8.4	8.7	9.0	11.0	13.3	16.2	18.6	18.4	19.0	17.9	22.2
FirmSharpeRatio	0.87	0.91	0.27	0.91	1.60	0.76	-0.08	0.68	1.08	-0.44	1.08
ROA	6.54	6.22	7.44	8.20	7.83	7.19	7.14	6.85	6.08	6.59	7.41
N	379	336	329	357	362	365	465	452	484	491	475
Conventional funds											
Size	16.0	12.0	13.2	15.0	14.2	17.4	22.9	23.5	23.5	25.4	28.9
MktVal	6.9	7.1	7.1	8.9	10.9	12.7	14.8	15.0	16.5	15.6	19.3
FirmSharpeRatio	0.93	0.90	0.22	0.92	1.60	0.68	-0.13	0.71	1.02	-0.50	0.96
ROA	5.33	5.95	6.53	7.10	7.02	6.31	6.54	5.63	5.47	6.02	6.79
N	506	456	470	496	504	525	626	601	595	591	563

Fund summary statistics are provided in Table 4.4; I show that conventional funds are, on average, about four to five times the size of SI funds measured in total net assets. On all other risk-return aspects, such as total return and Sharpe ratio, SI and conventional funds do not show substantial differences on a yearly basis. This is consistent with the graphic evidence provided in Section 4.2 and the broader body of literature on the inconclusiveness of SI outperformance.

Table 4.4 Fund summary stats

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
SI funds											
Size	466.6	524.6	513.4	522.4	696.6	805.8	683.3	672.1	676.1	592.6	776.6
Return	31.83	16.28	-1.00	14.06	31.67	9.61	-1.31	10.57	19.43	-5.75	29.65
Sharpe	0.41	0.25	0.01	0.36	0.91	0.31	-0.02	0.26	1.05	-0.13	0.55
beta	0.89	0.98	1.07	1.01	0.95	1.05	0.90	1.07	0.89	0.98	1.01
alpha	6.55	1.38	-3.02	-1.69	0.93	-4.29	-2.60	-1.95	0.21	-1.61	-1.55
N	59	59	61	64	66	65	75	86	98	105	116
Conventional funds											
Size	2,045.7	2,299.0	2,149.7	2,341.5	3,101.3	3,328.5	3,128.1	3,292.3	3,725.6	3,390.3	4,138.2
Return	32.83	19.12	-1.18	15.19	33.81	8.54	-1.45	12.40	19.09	-7.23	27.38
Sharpe	0.40	0.28	0.01	0.37	0.92	0.26	-0.02	0.27	0.96	-0.15	0.47
beta	0.98	1.02	1.15	1.04	0.99	1.09	0.88	1.18	0.81	1.00	1.08
alpha	5.69	3.56	-3.35	-0.98	1.75	-5.77	-2.89	-1.34	1.22	-3.00	-5.12
N	1,417	1,462	1,543	1,605	1,660	1,729	1,823	1,901	2,017	2,075	2,164

Table 4.5 below summarizes the number of firms facing different market reactions after controversies. In an average year, I am able to identify 50 firms with controversies that are held by SI and 58 controversial firms held by conventional funds. This is consistent with my earlier findings in Section 3.2 that SI funds on average, hold fewer stocks in their portfolio, and a large proportion of their holdings overlap with firms in the conventional fund portfolio. It is also worth noting that compared to conventional funds, a larger proportion of the firms held by SI funds experienced controversies (12.2% compared to 10.7%). This finding implies that from an investor's perspective, investing in SI funds does not guarantee a "free from negative externalities" portfolio. From a

fund manager’s perspective, I also notice that due to my strict definition of “treatment firms”, I do not have a large sample of controversies for which I can obtain CARs. Nonetheless, I observe an interesting pattern: most uncontaminated ESG controversy events do not lead to strong and definitive market reactions. On average, over 75% of all controversies are Zero-CAR controversies. I hypothesize that this is due to disagreement among market participants and show evidence for this argument in the following section. Meanwhile, I acknowledge that the limited number of Positive-CAR and Negative-CAR firms may lead to a lack of power in my regression results.

Table 4.5 number of companies held by funds

	$Sign_{j,t}$	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Firms held by SI funds												
No Controversy	0	332	260	277	302	313	303	438	400	453	444	423
Zero CAR	1	45	59	40	46	39	48	15	42	24	33	42
Positive CAR	2	0	8	9	2	3	7	3	3	4	6	4
Negative CAR	3	2	9	3	7	7	7	9	7	3	8	6
N		379	336	329	357	362	365	465	452	484	491	475
Firms held by Conventional funds												
No Controversy	0	456	374	410	433	445	456	594	536	554	532	501
Zero CAR	1	48	63	47	54	47	52	18	51	31	42	50
Positive CAR	2	0	8	10	2	3	8	4	5	4	7	5
Negative CAR	3	2	11	3	7	9	9	10	9	6	10	7
N		506	456	470	496	504	525	626	601	595	591	563

4.4 Results and Discussion

In Table 4.6, I show results from my baseline model. The first column shows results from estimating equation (1). I find that when controversies take place, conventional funds reduce their stakes in that firm, while SI funds increase their holdings. The reduction in conventional funds holding amounts to 0.015% of the firm’s total shares outstanding. Increases in holdings by SI funds are larger, representing 0.0182% of the firm’s existing shares. Both the increase and the reduction in holdings are significant at the 1% level.

When I distinguish between different types of controversies, i.e., by their market reactions measured with CARs, my findings are similar to those of the first model. Specifically, conventional mutual funds sell controversial firms when the event induces no market reaction (Zero CAR). Compared to the first specification, the size of the reduction is larger: 0.019% of the firm's total shares outstanding. SI funds act in the opposite direction and almost offset the selling pressure from conventional investors with a significant increase of 0.019%. For positive and negative CAR controversies, I do not find significant differences in portfolio rebalancing decisions between SI and conventional funds.

Table 4.6 The effect of ESG controversies on SI fund holdings relative to conventional funds.

	All controversies	Controversy by CAR sign
SI	0.0393***	0.0393***
	(0.00245)	(0.00245)
Controversy	-0.0150***	
	(0.00360)	
SI*Controversy	0.0182***	
	(0.00484)	
ZeroCAR		-0.01930***
		(0.00429)
PositiveCAR		-0.00474
		(0.0192)
NegativeCAR		0.00316
		(0.0111)
SI*ZeroCAR		0.01932***
		(0.00492)
SI*PositiveCAR		-0.00499
		(0.0142)
SI*NegativeCAR		0.0258
		(0.0164)
adj. R^2	0.330	0.330
No. of firms held by SI funds	280	280
No. of SI funds	91	91
No. of firms held by Conventional funds	333	333
No. of Conventional funds	1846	1846
N	476988	476988

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table reports the baseline regression results under the DiD model specified in Equation (4.1) and Equation (4.2). The dependent variable is $ShrsHeld\_Csho_{i,j,t}$, which represents the stake (in %) of a firm *i* held by fund *j* at year *t*. The first column shows the result from using $Controv_{j,t}$ as the independent variable. The second column shows the result using $Sign_{j,t}$ as the independent variable. The coefficient for ZeroCAR is equivalent to $Sign_{j,t} = 1$, while PositiveCAR and NegativeCAR correspond to cases where $Sign_{j,t}$ takes the value of 2 and 3, respectively.

One may argue that the lack of significance in my baseline regression is due to lack of power. As the definition of my independent variable $Sign_{j,t}$ depends on the event studies CAR estimation, it is impossible to increase the number of observations at the firm level. Therefore, I use an alternative definition of SI to increase the number of SI funds in my sample. Specifically, I define US PRI funds as SI and non-PRI funds as conventional. Results obtained in this alternative specification are in line with my baseline results, although the increase in holding is larger (0.02%) when the market reaction to the controversy is not statistically significant.

Table 4.7 SI defined as PRI= 1

	All controversies	Controversy by CAR sign
SI	-0.0225***	-0.0224***
	(0.00498)	(0.00498)
Controv	-0.0149***	
	(0.00360)	
SI*Controv	0.0180***	
	(0.00553)	
ZeroCAR		-0.0192***
		(0.00429)
PositiveCAR		-0.00468
		(0.0188)
NegativeCAR		0.00334
		(0.0114)
SI*ZeroCAR		0.0206***
		(0.00648)
SI*PositiveCAR		-0.0113
		(0.0109)
SI*NegativeCAR		0.0196
		(0.0128)
adj. R^2	0.329	0.329
No. of firms held by SI funds	292	292
No. of SI funds	117	117
No. of firms held by Conventional funds	333	333
No. of Conventional funds	1906	1906
N	476988	476988

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table reports the regression results under the DiD model specified in Equation (4.1) and Equation (4.2), but defines SI as funds that are PRI signatories. The dependent variable is $ShrsHeld\_Csho_{i,j,t}$, which represents the stake (in %) of a firm *i* held by fund *j* at year *t*. The first column shows the result from using $Controv_{j,t}$ as the independent variable. The second column shows the result using $Sign_{j,t}$ as the independent variable. The coefficient for ZeroCAR is equivalent to $Sign_{j,t} = 1$, while PositiveCAR and NegativeCAR correspond to cases where $Sign_{j,t}$ takes the value of 2 and 3, respectively.

Under the theoretical framework by Friedman and Heinle (2016), in cases where all market participants agree upon the information content of CSR-related disclosures, variation in investor tastes and preferences for sustainability will lead to differences in portfolio rebalancing decisions post-disclosure. However, when the market reaction to the controversy is unanimous (Positive CAR and Negative CAR), SI and conventional funds do not rebalance their portfolio, which may be interpreted as if SI funds did not hold a distinct and stronger taste for sustainability or CSR. This finding is also consistent with the equilibrium model constructed by Pederson,

Table 4.7 above summarizes the volatility of CARs depending on the direction of market reaction. I observe that Zero-CAR controversy events have a larger average standard deviation of 6.68 and are more sparsely distributed than Positive-CAR and Negative-CAR events. Further, the variance ratio test in Table 4.8 shows that volatilities between Zero-CAR and Negative-CAR and Zero-Car and Positive-CAR are not the same. This suggests that compared to events with a clear positive or negative

CAR outcome, there is a lack of clear consensus among market participants on the stock valuation implications of some of the controversies I analyze. Moreover, the larger CAR volatility indicates opposing and potentially strong opinions, which illicit upward and downward pressure on stock prices simultaneously, resulting in a net-zero CAR effect.

Table 4.8 Variance ratio test – CAR volatility

	Obs	Standard Error	Standard Deviation	p-value for ratio not equal to 1
Zero CAR	520	0.0159	0.363	0.0000***
Positive CAR	66	0.0963	0.783	
Zero CAR	520	0.0159	0.363	0.0000***
Negative CAR	66	0.0670	0.636	

The table reports the variance ratio test of CAR. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I compute the average of CAR variances for ZeroCAR events, PositiveCAR events, and NegativeCAR events, then take the square root and obtain standard deviation to use in the STATA command: *sdtesti*, which tests the null hypothesis that sd(ZeroCAR)/sd(PositiveCAR) = 1. The last column shows the p-value for the alternative hypothesis that the ratio is not equal to 1.

When there is disagreement about the financial implication of controversial ESG events (Zero-CAR events), conventional funds sell shares of these companies. However, SI funds increase the holdings of such firms. I hypothesize that this is because instead of valuing "pure ESG" performance, SI funds pursue opportunities to improve their fund-level ESG scores. Firstly, during my sample period, there is a lack of stringent and consistent disclosure regulations that funds need to adhere to to be marketed as sustainable investment funds. For example, the SEC proposed a new, coherent policy on fund-level ESG integration or screening disclosure in May 2022, which sets out tiered standards for disclosure of ESG incorporation strategies when mutual funds want to attach various SI labels to themselves. Prior to this implementation, it is likely that SI funds did not have other apparent or visible means of demonstrating ESG commitment except through enhancing their fund-level ESG scores, such as by obtaining a higher Morningstar Globes rating to attract flows (Hartzmark & Sussman, 2019).

I provide evidence on the relationship between ESG ratings and controversies by estimating the following model:

$$ESGScore_{j,t} = \alpha_i + \eta_t + \beta_2 Sign_{j,t} + \vartheta FirmControl_{j,t} + \varepsilon_{j,t} \quad (4.4)$$

In Equation (4.4), $ESGScore_{j,t}$ is the ASSET4 overall ESG rating of company j at year

t, $Sign_{j,t}$ is the categorical variable for whether firm j had a controversy in year t and the corresponding market reaction measured by CAR.

Table 4.9 illustrates the predictability of controversies on ESG ratings contemporaneously and in a forward-looking manner. Combining all controversies, a firm's current period controversy predicts an increase in its firm-level overall ESG scores. When I analyze the market reaction to these controversies, Zero CAR and Positive CAR events also positively and significantly impact firms' ESG scores. The predictability between higher ESG scores, controversies in general, and controversies with Zero CAR events may help explain why SI funds increase holdings after ESG controversies. Previous research demonstrates that investors' capital favors high Morningstar Globed funds over low Morningstar globe funds (Hartzmark & Sussman, 2019). Fund managers thus have incentives to improve their portfolio or fund-level ESG ratings to "signal" their social preference to SI investors and attract flows. As a result, even though controversies are bad news to companies' actual ESG performance, as long as they do not directly induce ESG downgrades, controversial firms may still be held by SI funds. In addition, the increase in the ESG score during and after the year of the ESG controversy can be understood if one references the literature review presented in Section 3.2. In short, academic literature established that ESG ratings are often opaquely defined and assigned, composed of many different ESG "opportunities" and "risks" metrics. After controversies, companies may enforce more stringent policies to ensure further controversies do not occur again. And because ESG rating agencies such as Refinitiv consider information from a wide range of ESG metrics, including policy-related metrics, when assigning ESG scores (Refinitiv, 2022), any improvement in policies, processes, and operational systems would have been incorporated by Refinitiv when producing ESG rating, creating a potential increase in ESG scores after controversies.

Table 4.9 Predictability of contemporaneous and forward-looking ESG scores using ESG controversies

	Controversy in general		Controversy by type	
	Contemporaneous ESG score	1-year forward ESG score	Contemporaneous ESG score	1-year forward ESG score
Controversy	4.506^{***}	3.935^{***}		
	(0.731)	(0.766)		
ZeroCAR			4.698^{***}	4.653^{***}
			(0.809)	(0.842)
PositiveCAR			8.427^{***}	6.540^{***}
			(2.352)	(2.100)
NegativeCAR			0.784	-1.580

			(2.201)	(2.582)
Volatility	-0.00428***	-0.00374***	-0.00428***	-0.00374***
	(0.000613)	(0.000724)	(0.000613)	(0.000724)
Size	0.0198***	0.0296***	0.0198***	0.0297***
	(0.00305)	(0.00399)	(0.00305)	(0.00399)
MarketCap	0.318***	0.297***	0.318***	0.297***
	(0.0114)	(0.0155)	(0.0114)	(0.0155)
SalesGrowth	-0.0665***	-0.0204***	-0.0666***	-0.0204***
	(0.00582)	(0.00435)	(0.00580)	(0.00435)
PB	-0.000751***	-0.000639**	-0.000751***	-0.000640**
	(0.0000952)	(0.000252)	(0.0000952)	(0.000252)
ROA	-0.0367**	-0.0330**	-0.0368**	-0.0330**
	(0.0151)	(0.0165)	(0.0151)	(0.0165)
SharpeRatio	-0.638***	-0.477***	-0.636***	-0.480***
	(0.0806)	(0.0977)	(0.0806)	(0.0977)
DebtAsset	-0.0159***	-0.0112**	-0.0158***	-0.0111**
	(0.00540)	(0.00530)	(0.00540)	(0.00530)
constant	31.56***	32.72***	31.57***	32.73***
	(0.928)	(1.013)	(0.928)	(1.013)
adj. R^2	0.338	0.325	0.338	0.325
Number of Firms	337	321	337	321
N	79748	74693	79748	74693

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table reports the baseline regression results under the model specified in Equation (4.4). The dependent variable is $ESGScore_{j,t}$ and represents the Refinitiv ESG score of company *j*. The first two columns show results from using $Controv_{j,t}$ as the independent variable. The third and fourth column show results using $Sign_{j,t}$ as the independent variable. The coefficient for ZeroCAR is equivalent to $Sign_{j,t} = 1$, while PositiveCAR and NegativeCAR correspond to cases where $Sign_{j,t}$ takes the value of 2 and 3, respectively. The second and fourth columns use 1-year forward ESG ratings as the dependent variable. Control variables include: Volatility (variance of firms' monthly returns over the past 24 months from the end of year t), $Size_{i,t-1}$ (1-year lagged total net assets), MktCap (1-year lagged market value in $ millions), SalesGrowth (growth in sales per share from the previous year in percentage), ROA (1-year lagged return on assets), DebtAsset (1-year lagged debt to asset ratio) and SharpeRatio (1-year lagged firm Sharpe ratio).

Thus far, I have established that SI mutual funds increase holdings for zero-CAR controversies because there is a positive association between zero-CAR ESG controversies and ESG ratings. However, it is unclear why they do not increase holdings for positive CAR controversies if it also leads to an increase in ESG ratings (the overall ESG score is scaled from 0 to 100, and the size of the increase is 6.5 to 8.4). I argue that this is because a significant positive CAR indicates that the market has already consistently incorporated the ESG controversy news; any potential monetary reward that could be earned due to this information has already been exhausted. The same logic applies to negative CAR controversies. In these cases, both SI and conventional funds act as mean-variance investors and do not rebalance their portfolios. However, when controversies induce disagreement between market participants on the price implication (zero CAR), mutual funds face a "gamble" of whether to enter the market as a buyer. I find that conventional funds are not constrained by any commitment to ESG to become sellers of these stocks. On the other hand, SI mutual funds become buyers because they are incentivized by receiving higher ESG ratings.

4.5 Conclusion

In this Chapter, I find consistent evidence with existing SI literature that compares SI and conventional fund portfolio performance (such as those reviewed in Renneboog et al., 2008). There is little difference between SI and conventional mutual funds: they share similar risk-return profiles and have a high level of overlap in terms of assets held in their portfolios. I also document anecdotal evidence that neither SI mutual fund portfolio weightings nor ESG ratings react to well-publicized corporate scandals such as the BP oil spill and Facebook's data privacy scandal. My empirical design has the following distinct features: (1) the use of objective ESG controversy counts instead of ESG ratings to identify actual ESG deterioration; (2) identifying price implications of negative externalities created by firms using event studies; (3) precise measurement of SI mutual funds portfolio choices using funds' proportional stake in firms; (4) provided more granularity in understanding the interaction between SI fund rebalancing decisions, ESG controversies, and stock market reactions. I also examine possible mechanisms for SI and conventional funds to hold different portfolios by examining the impact of controversies on ESG ratings. Baseline results from this chapter find differences in SI and conventional rebalancing decisions, but only when controversies yield zero CAR and not when CAR is strongly positive or negative.

My empirical findings contribute to the existing literature on ESG preferences. Theoretically, my finding of no rebalancing after positive and negative CARs confirms the following theoretical prediction from Pederson, Fitzgibbons, and Pomorski (2021). When the price implications of ESG controversies are clearly understood by market participants, both SI and conventional funds behave as the "ESG signal" mean-variance type of investors, and SI funds do not demonstrate preferences for high ESG scores. Empirically, this is akin to evidence provided by Bebchuk et al. (2013), where all market participants gradually learn about the value-relevance of a governance indicator and begin to incorporate it into stock prices so that high ESG scores do not have return predictability. In other words, there is insufficient dispersion in market participants' social or ESG preferences to create more than one optimal market portfolio (Pastor et al., 2021). When firms' ESG information leads to disagreement on short-term profitability (I show this with the variance of zero-CARs), SI funds in my sample increase holdings in expectation of receiving higher ESG ratings. This finding is in line with the theoretical framework by Pederson, Fitzgibbons, and Pomorski (2021), where SI funds act as investors with preferences for high ESG scores and hold different portfolios compared to conventional funds.

I acknowledge that there may be alternative explanations for my findings. Firstly, I exclude companies with more than one type of controversy over the course of a year to avoid contamination for my event studies. However, Refinitiv counts controversies based on information gathered via media reports. Finding multiple types of controversies could therefore indicate higher levels of media and public attention. Thus, my baseline results may be tilted toward smaller companies. In addition, the number of positive and negative CAR controversies is smaller than zero CAR controversies, which could result in low statistical power.

Nonetheless, my robustness test using PRI as the definition for SI increases the estimation sample, and my baseline result remains unchanged. In my third chapter, I do not completely rule out the possibility for some SI funds to hold social preferences for actual ESG performance. Since the identification strategy is based on ESG controversies, I am examining SI fund reaction to negative externalities, a more comprehensive conclusion could be reached if there exists an objective measure for positive externalities. Meanwhile, future work could also consider incorporating ESG engagement data to examine whether SI funds differ from conventional funds in making proxy voting and shareholder proposals after ESG controversies.

Appendix 4.1 – All controversies from Refinitiv

Table 4.1.1 Total Number of Controversies by Type & year – US firms covered by Refinitiv

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Total Controversies	**616**	**647**	**1,102**	**835**	**687**	**938**	**917**	**388**	**753**	**733**	**930**
Anti-Competition	56	65	91	49	50	81	83	47	178	134	177
Business Ethics	86	135	206	183	147	223	173	64	138	169	124
Intellectual Property	114	121	184	81	49	42	28	4	23	41	73
Critical Country	10	8	7	0	0	0	2	0	0	0	2
Public Health	15	8	23	19	25	40	30	7	11	11	21
Tax Fraud	5	11	13	13	7	13	6	16	27	17	28
Child Labour	1	3	6	6	7	2	5	0	0	0	0
Human Right	15	11	10	5	3	2	3	0	2	0	1
Magement Compensation	3	8	4	2	1	3	3	1	1	2	2
Consumer Controversy	76	30	89	49	26	46	80	33	34	46	69
Customer Health and Safety	46	44	90	64	44	62	99	24	90	68	40
Privacy	9	12	23	23	19	38	60	28	61	89	171
Product Access	0	0	4	5	8	16	11	4	6	1	4
Responsible Marketing	26	30	73	73	75	90	86	12	15	15	30
Responsible R&D	5	2	3	3	4	3	0	1	1	1	0
Environmental Resource Impact	0	2	11	6	4	11	17	14	62	29	37
Accounting	22	10	4	2	1	6	5	0	0	3	10
Insider Dealing	2	4	5	6	6	7	6	3	2	4	7
Shareholder Rights	0	4	40	49	51	54	38	28	6	7	34
Diversity & Opprty	48	41	35	43	37	41	37	30	17	36	46
Employee Health & Safety	30	41	79	72	57	64	62	18	18	7	7
Wage Work Condition	33	50	87	60	48	78	68	40	40	25	25
Strike	14	7	15	22	18	16	15	14	21	28	22

Chapter 5 : Conclusion

This thesis investigates whether investors and managers make sustainable investment decisions based on pro-social preferences. I find that both investors and managers are not behaving in line with pro-social behavior. In Chapter 2, I focus on the disentangling of social preferences from philanthropy flow data and examine the impact of different exogenous shocks on individuals' willingness to give. I exploit granular and purpose-based charity information and financial data from the US and the UK from 2008 to 2018. I find that during this period, donation flows reacted significantly and positively to environmental shocks (except for the BP oil spill) but did not change after two humanitarian crises: the Haiti earthquake and the Syrian refugee crisis. I show that in line with previous theoretical and empirical evidence, reduced monetary incentives due to a lower tax shield from donation leads to a lower total donation to charities. Meanwhile, scandals that disrepute charities also create a sector-wide loss of confidence, resulting in fewer donation flows. I also rule out the alternative hypothesis that one of my test results is driven by the reputation effect rather than pro-social preferences.

In Chapter 3, I examine whether investors direct flows into SI mutual funds to demonstrate their social preferences. I first use the shocks that induced significant social preference changes, measured in donation flows in Chapter 2, to conduct my quasi-experiment on SI mutual fund flows. Results from this Chapter suggest that SI mutual fund investments are unlikely to be driven by pro-social preferences. After environmental disasters or policies that could potentially increase environmental externalities, investors do not direct more resources to SI mutual funds after explicitly controlling for risk-return profiles of SI mutual funds. My empirical design also addresses alternative channels to demonstrate social preferences using SI investment. I use charity scandals to rule out the possibility that SI funds may be perceived as less irrelevant or effective options for pro-social individuals. Considering the tax reform shock examined in this Chapter, I illustrate that individuals retain their current positions in SI mutual funds and do not redirect resource from philanthropy when monetary incentives to donate is lacking. Results in this Chapter are robust to the reduced search cost and short-term effect, and differences between institutional and retail investors do not alter my findings. Taken together, the results in Chapter 3 point to the conclusion that SI investment cannot be driven by individuals' social preferences alone. One possibility not explored in Chapter 3 is that investors do not see SI as a vehicle for satisfying their social preferences because SI funds do not hold pro-social portfolios that are different from conventional funds that only pursue financial gains.

In Chapter 4, I test the hypothesis that SI mutual fund managers do not exhibit significantly different social preferences by comparing holdings of SI and conventional funds after corporate ESG controversies. I show that SI fund managers do not reduce their stakes in firms that incur actual ESG controversies. This suggests that the primary motivation for SI fund portfolio rebalancing decisions is not concern over externalities corporations generate. My empirical evidence also indicates that SI fund managers do not make different rebalancing decisions compared to their conventional counterparts. As long as controversies cause clear short-term stock price reactions, both types of funds keep their existing stakes in these controversial firms. I then analyze the potential mechanism for the lack of portfolio differences by revealing a positive relationship between the two. In light of these findings, I argue that SI fund managers reflect concerns over portfolio ESG ratings and financial motivations in their investment decisions but not their preferences over the actual ESG performance of companies.

Overall, I find that both SI investors and fund managers do not demonstrate strong social preferences in their investment decisions. Individuals contribute to charities to support pro-social causes but do not consider SI an appropriate vehicle for providing social impact. Interestingly, investment flows towards SI remain unchanged even after charities become less attractive to pro-social agents. There is also little dispersion among fund managers' revealed preferences for ESG, as they hold similar portfolios before and after corporate ESG controversy events. The effect I find in donation, SI flows, and SI portfolio holdings are robust to pecuniary motivations of investors and fund managers. I thus argue that market participants hold similar rather than diverse social preferences. However, I do acknowledge that SI fund managers can demonstrate social preferences in the form of engagement rather than divestment. Thus, future work could consider examining shareholder engagement around corporate ESG incidents.

Indeed, my findings appear to contradict the recent growth of the SI industry and surveys on investor ESG attitudes. However, empirical findings from my stringent test designs use revealed preferences and show consistency with existing empirical evidence of ambiguity in the performance of "green" assets and the similarity in SI and conventional mutual fund performances (such as reviewed by Liang & Renneboog, 2020). I also provide empirical evidence that complements theoretical predictions of recently developed equilibrium models with investors of different tastes (Pastor et al., 2020; Pedersen et al., 2021), and shed light on the future development of theoretical models that might incorporate assumptions on taste for ESG.

The policy implications of this thesis are also important; instead of implicitly assuming social preferences for investors and fund managers, pro-social policymakers should consider strategies to promote awareness of ESG and aim to amplify the social preferences of economic agents to achieve social objectives while enhancing the credibility of SI mutual funds as vehicles for providing social impact. My findings also suggest that existing definitions of SI and ESG ratings are insufficient indicators for actual portfolio ESG performance. Thus, regulators could consider establishing more stringent rules and standards to discern SI and conventional mutual funds. Finally, it would be interesting to revisit this research once new disclosure requirements for SI mutual fund ESG practices, such as the one proposed in May 2022 by the US SEC become widely adopted.

References

Adams, R. B. (2016). Women on boards: The superheroes of tomorrow? *The Leadership Quarterly*, *27*(3), 371–386. https://doi.org/10.1016/j.leaqua.2015.11.001

Agnew, H., Klasa, A., & Mundy, S. (2022, June 6). How ESG investing came to a reckoning. *Financial Times*. https://www.ft.com/content/5ec1dfcf-eea3-42af-aea2-19d739ef8a55

Albuquerque, R., Koskinen, Y., & Zhang, C. (2019). Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence. *Management Science*, *65*(10), 4451–4469. https://doi.org/10.1287/mnsc.2018.3043

Amel-Zadeh, A., & Serafeim, G. (2018). Why and How Investors Use ESG Information: Evidence from a Global Survey. *Financial Analysts Journal*, *74*(3), 87–103. https://doi.org/10.2469/faj.v74.n3.2

Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy*, *97*(6), 1447–1458.

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, *100*(401), 464–477. https://doi.org/10.2307/2234133

Andreoni, J. (2006). Leadership Giving in Charitable Fund-Raising. *Journal of Public Economic Theory*, *8*(1), 1–22. https://doi.org/10.1111/j.1467-9779.2006.00250.x

Andreoni, J., & Bernheim, B. D. (2009). Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica*, *77*(5), 1607–1636.

Andreoni, J., & Miller, J. (2002). Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica*, *70*(2), 737–753. https://doi.org/10.1111/1468-0262.00302

Andreoni, J., & Payne, A. A. (2003). Do Government Grants to Private Charities Crowd out Giving or Fund-Raising? *The American Economic Review*, *93*(3), 792–812.

Andreoni, J., & Payne, A. A. (2011). Is crowding out due entirely to fundraising? Evidence from a panel of charities. *Journal of Public Economics*, *95*(5), 334–343. https://doi.org/10.1016/j.jpubeco.2010.11.011

Andreoni, J., Rao, J. M., & Trachtman, H. (2017). Avoiding the Ask: A Field Experiment on Altruism, Empathy, and Charitable Giving. *Journal of Political Economy*. https://doi.org/10.1086/691703

Ariely, D., Bracha, A., & Meier, S. (2009). Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *The American Economic Review*, *99*(1), 544–555.

Auten, G. E., Sieg, H., & Clotfelter, C. T. (2002). Charitable Giving, Income, and Taxes: An Analysis of Panel Data. *The American Economic Review*, *92*(1), 371–382.

Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, *211*(4489), 1390–1396. https://doi.org/10.1126/science.7466396

Barber, B. M., Morse, A., & Yasuda, A. (2021). Impact investing. *Journal of Financial Economics*, *139*(1), 162–185. https://doi.org/10.1016/j.jfineco.2020.07.008

Baer, J., & Lim, D. (2021, June 12). The Hedge-Fund Manager Who Did Battle With Exxon—And Won. *Wall Street Journal*. https://www.wsj.com/articles/the-hedge-fund-manager-who-did-battle-with-exxonand-won-11623470420

Baker, M., Bergstresser, D., Serafeim, G., & Wurgler, J. (2022). The Pricing and Ownership of US Green Bonds. *Annual Review of Financial Economics*, *14*(1), 415–437. https://doi.org/10.1146/annurev-financial-111620-014802

Basil, D. Z., Ridgway, N. M., & Basil, M. D. (2006). Guilt appeals: The mediating effect of responsibility. *Psychology & Marketing*, *23*(12), 1035–1054. https://doi.org/10.1002/mar.20145

Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, *29*(7), 1751–1767. https://doi.org/10.1016/j.jbankfin.2004.06.035

Bauer, R., Ruof, T., & Smeets, P. (2021). Get Real! Individuals Prefer More Sustainable Investments. *The Review of Financial Studies*, *34*(8), 3976–4043. https://doi.org/10.1093/rfs/hhab037

Bebchuk, L. A., Cohen, A., & Wang, C. C. Y. (2013). Learning and the disappearing association between governance and returns. *Journal of Financial Economics*, *108*(2), 323–348. https://doi.org/10.1016/j.jfineco.2012.10.004

Becker, G. S. (1974). A Theory of Social Interactions. *Journal of Political Economy*, *82*(6), 1063–1093. https://doi.org/10.1086/260265

Bello, Z. Y. (2005). Socially Responsible Investing and Portfolio Diversification. *Journal of Financial Research*, *28*(1), 41–57. https://doi.org/10.1111/j.1475-6803.2005.00113.x

Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, *96*(5), 1652–1678. https://doi.org/10.1257/aer.96.5.1652

Bendapudi, N., Singh, S. N., & Bendapudi, V. (1996). Enhancing Helping Behavior: An Integrative Framework for Promotion Planning. *Journal of Marketing*, *60*(3), 33–49. https://doi.org/10.1177/002224299606000303

Benson, K. L., & Humphrey, J. E. (2008). Socially responsible investment funds: Investor reaction to current and past returns. *Journal of Banking & Finance*, *32*(9), 1850–1859. https://doi.org/10.1016/j.jbankfin.2007.12.013

Berg, F., Fabisik, K., & Sautner, Z. (2021). *Is History Repeating Itself? The (Un)Predictable Past of ESG Ratings* (SSRN Scholarly Paper No. 3722087). https://doi.org/10.2139/ssrn.3722087

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. *Review of Finance*, *26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to Effective Altruism: The Role of Subjective Preferences in Charitable Giving. *Psychological Science*, *29*(5), 834–844. https://doi.org/10.1177/0956797617747648

Bernheim, B. D. (1986). On the Voluntary and Involuntary Provision of Public Goods. *The American Economic Review*, *76*(4), 789–793.

Białkowski, J., & Starks, L. T. (2016). SRI Funds: Investor Demand, Exogenous Shocks and ESG Profiles. *Working Papers in Economics*, Article 16/11. https://ideas.repec.org//p/cbt/econwp/16-11.html

Bohnet, I., & Frey, B. S. (1999). Social Distance and Other-Regarding Behavior in Dictator Games: Comment. *American Economic Review*, *89*(1), 335–339. https://doi.org/10.1257/aer.89.1.335

Bollen, N. P. B. (2007). Mutual Fund Attributes and Investor Behavior. *The Journal of Financial and Quantitative Analysis*, *42*(3), 683–708.

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, *142*(2), 517–549. https://doi.org/10.1016/j.jfineco.2021.05.008

Boffo, R., & Patalano, R. (2020). *ESG Investing: Practices, Progress and Challenges*. https://www.oecd.org/finance/ESG-Investing-Practices-Progress-Challenges.pdf

Borgers, A., Derwall, J., Koedijk, K., & ter Horst, J. (2015). Do social factors influence investment behavior and performance? Evidence from mutual fund holdings. *Journal of Banking & Finance*, *60*, 112–126. https://doi.org/10.1016/j.jbankfin.2015.07.001

Bousso, R. (2021, May 18). Shell shareholders increase pressure for further climate action. *Reuters*. https://www.reuters.com/business/energy/shell-shareholders-overwhelmingly-support-energy-transition-plan-2021-05-18/

Bouten, L., Cho, C. H., Michelon, G., & Roberts, R. W. (2017). *CSR Performance Proxies in Large-Sample Studies: 'Umbrella Advocates', Construct Clarity and the 'Validity Police'* (SSRN Scholarly Paper No. 3107182). https://doi.org/10.2139/ssrn.3107182

Bruni, L., & Zamagni, S. (2013). *Handbook on the Economics of Reciprocity and Social Enterprise*. Edward Elgar Publishing. https://ideas.repec.org//b/elg/eebook/14159.html

Burgoyne, C. B., Young, B., & Walker, C. M. (2005). Deciding to give to charity: A focus group study in the context of the household economy. *Journal of Community & Applied Social Psychology*, *15*(5), 383–405. https://doi.org/10.1002/casp.832

CAF. (2016). *CAF World Giving Index 2016 | Giving and volunteering around the world.* Retrieved 6 January 2023, from https://www.cafonline.org/about-us/publications/2016-publications/caf-world-giving-index-2016

CAF. (2019). *CAF World Giving Index 2019| Ten years of giving trends.* Retrieved 6 January 2023, from https://www.cafonline.org/about-us/publications/2019-publications/caf-world-giving-index-10th-edition

Cao, J., Titman, S., Zhan, X., & Zhang, W. E. (2021). *ESG Preference, Institutional Trading, and Stock Return Patterns* (SSRN Scholarly Paper No. 3353623). https://doi.org/10.2139/ssrn.3353623

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Cavale, S. (2021, May 5). Unilever says majority of shareholders voted in favour of climate action plan. *Reuters*. https://www.reuters.com/business/sustainable-business/unilever-says-majority-shareholders-voted-favour-climate-action-plan-2021-05-05/

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics*, *117*(3), 817–869. https://doi.org/10.1162/003355302760193904

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, *37*(8), 1597–1614. https://doi.org/10.1002/smj.2407

Chava, S. (2014). Environmental Externalities and Cost of Capital. *Management Science*, *60*(9), 2223–2247.

Chen, G. (2009). Does meeting standards affect charitable giving? An empirical study of New York metropolitan area charities. *Nonprofit Management and Leadership*, *19*(3), 349–365. https://doi.org/10.1002/nml.223

Cheung, C.-K., & Chan, C.-M. (2000). Social-cognitive factors of donating money to charity, with special attention to an international relief organization. *Evaluation and Program Planning*, *23*(2), 241–253. https://doi.org/10.1016/S0149-7189(00)00003-3

Chevalier, J., & Ellison, G. (1997). Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy* 105 (6), 1167–1200. https://doi.org/10.1086/516389

Chowdhry, B., Davies, S. W., & Waters, B. (2019). Investing for Impact. *The Review of Financial Studies*, *32*(3), 864–904. https://doi.org/10.1093/rfs/hhy068

Clotfelter, C. T. (1980). Tax incentives and charitable giving: Evidence from a panel of taxpayers. *Journal of Public Economics*, *13*(3), 319–340. https://doi.org/10.1016/0047-2727(86)90009-5

Cochran, P. L., & Wood, R. A. (1984). Corporate Social Responsibility and Financial Performance. *The Academy of Management Journal*, *27*(1), Article 1. JSTOR. https://doi.org/10.2307/255956

De Angelis, T., Tankov, P., & Zerbib, O. D. (2022). Climate Impact Investing. *Management Science*. https://doi.org/10.1287/mnsc.2022.4472

De Pelsmacker, P., Driesen, L., & Rayp, G. (2005). Do Consumers Care about Ethics? Willingness to Pay for Fair-Trade Coffee. *Journal of Consumer Affairs*, *39*(2), 363–385. https://doi.org/10.1111/j.1745-6606.2005.00019.x

Dimson, E., Karakaş, O., & Li, X. (2015). Active Ownership. *The Review of Financial Studies*, *28*(12), 3225–3268. https://doi.org/10.1093/rfs/hhv044

Dragusanu, R., Giovannucci, D., & Nunn, N. (2014). The Economics of Fair Trade. *Journal of Economic Perspectives*, *28*(3), 217–236. https://doi.org/10.1257/jep.28.3.217

Duncan, B. (1999). Modeling charitable contributions of time and money. *Journal of Public Economics*, *72*(2), 213–242. https://doi.org/10.1016/S0047-2727(98)00097-8

Duquette, N. J. (2016). Do tax incentives affect charitable contributions? Evidence from public charities' reported revenues. *Journal of Public Economics*, *137*, 51–69. https://doi.org/10.1016/j.jpubeco.2016.02.002

Eccles, R. G., & Klimenko, S. (2019, May 1). The Investor Revolution. *Harvard Business Review*. https://hbr.org/2019/05/the-investor-revolution

Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics*, *101*(3), 621–640. https://doi.org/10.1016/j.jfineco.2011.03.021

Eckel, C. C., & Grossman, P. J. (1996). Altruism in Anonymous Dictator Games. *Games and Economic Behavior*, *16*(2), 181–191. https://doi.org/10.1006/game.1996.0081

Elliott, D. (2013, December 21). For BP Cleanup, 2013 Meant 4.6 Million Pounds Of Oily Gunk. *NPR*. https://www.npr.org/2013/12/21/255843362/for-bp-cleanup-2013-meant-4-6-million-pounds-of-gulf-coast-oil

Erikson, R. S., & Wlezien, C. (2012). *The Timeline of Presidential Elections: How Campaigns Do (and Do Not) Matter*. University of Chicago Press.

Eurosif. (2016). *SRI Study 2016*. https://www.eurosif.org/sri-study-2016/

Evans, R. B. (2010). Mutual Fund Incubation. *The Journal of Finance*, *65*(4), 1581–1611. https://doi.org/10.1111/j.1540-6261.2010.01579.x

Falkenstein, E. G. (1996). Preferences for Stock Characteristics As Revealed by Mutual Fund Portfolio Holdings. *The Journal of Finance*, *51*(1), 111–135. https://doi.org/10.1111/j.1540-6261.1996.tb05204.x

Fama, E. F., & French, K. R. (2007). Disagreement, tastes, and asset prices. *Journal of Financial Economics*, *83*(3), 667–689. https://doi.org/10.1016/j.jfineco.2006.01.003

Fehr, E., & Schmidt, K. M. (2006). Chapter 8 The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories. In S.-C. Kolm & J. M. Ythier (Eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity* (Vol. 1, pp. 615–691). Elsevier. https://doi.org/10.1016/S1574-0714(06)01008-6

Feldman, N. E. (2010). Time Is Money: Choosing between Charitable Activities. *American Economic Journal: Economic Policy*, *2*(1), 103–130. https://doi.org/10.1257/pol.2.1.103

Feldstein, M., & Taylor, A. (1976). The Income Tax and Charitable Contributions. *Econometrica*, *44*(6), 1201–1222. https://doi.org/10.2307/1914255

Fink. (2021). *Larry Fink CEO Letter*. BlackRock. https://www.blackrock.com/corporate/investor-relations/2021-larry-fink-ceo-letter

Franzoni, F., & Schmalz, M. C. (2017). Fund Flows and Market States. *The Review of Financial Studies*, *30*(8), Article 8. https://doi.org/10.1093/rfs/hhx015

Fombrun, C., & Shanley, M. (1990). What's in a Name? Reputation Building and Corporate Strategy. *The Academy of Management Journal*, *33*(2), 233–258. https://doi.org/10.2307/256324

Friedman, H. L., & Heinle, M. S. (2016). Taste, information, and asset prices: Implications for the valuation of CSR. *Review of Accounting Studies*, *21*(3), 740–767. https://doi.org/10.1007/s11142-016-9359-x

Friedman, M., (1970). 'The Social Responsibility Of Business Is to Increase Its Profits'. *The New York Times*, September 1970. https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html.

Geczy, C. C., Stambaugh, R. F., & Levin, D. (2021). Investing in Socially Responsible Mutual Funds. *The Review of Asset Pricing Studies*, *11*(2), 309–351. https://doi.org/10.1093/rapstu/raab004

Gibson, R., Glossner, S., Krueger, P., Matos, P., & Steffen, T. (2022). Do Responsible Investors Invest Responsibly?. *Review of Finance*, *26*(6), 1389–1432. https://doi.org/10.1093/rof/rfac064

Gibson, R., Krueger, P., & Schmidt, P. S. (2019). *ESG Rating Disagreement and Stock Returns* (SSRN Scholarly Paper No. 3433728). https://doi.org/10.2139/ssrn.3433728

Gibson, R., Krueger, P., & Mitali, S. F. (2021). *The Sustainability Footprint of Institutional Investors: ESG Driven Price Pressure and Performance* (SSRN Scholarly Paper No. 2918926). https://doi.org/10.2139/ssrn.2918926

Gintis, H. (2000). Strong Reciprocity and Human Sociality. *Journal of Theoretical Biology*, *206*(2), 169–179. https://doi.org/10.1006/jtbi.2000.2111

Giving USA. (2019, June 18). *Giving USA 2019: Americans gave $427.71 billion to charity in 2018 amid complex year for charitable giving | Giving USA*.

https://givingusa.org/giving-usa-2019-americans-gave-427-71-billion-to-charity-in-2018-amid-complex-year-for-charitable-giving/

Glazer, A., & Konrad, K. A. (1996). A Signaling Explanation for Charity. *The American Economic Review*, *86*(4), 1019–1028.

Gordon, T. P., Knock, C. L., & Neely, D. G. (2009). The role of rating agencies in the market for charitable contributions: An empirical test. *Journal of Accounting and Public Policy*, *28*(6), 469–484. https://doi.org/10.1016/j.jaccpubpol.2009.08.001

Goss, A., & Roberts, G. S. (2011). The impact of corporate social responsibility on the cost of bank loans. *Journal of Banking & Finance*, *35*(7), 1794–1810. https://doi.org/10.1016/j.jbankfin.2010.12.002

Green, M. (2018, February 16). Oxfam and the dark side of the aid industry. *Financial Times*. https://www.ft.com/content/8799725c-123c-11e8-8cb6-b9ccc4c4dbbb

Gruber, J., & Hungerman, D. M. (2007). Faith-based charity and crowd-out during the great depression. *Journal of Public Economics*, *91*(5), 1043–1069. https://doi.org/10.1016/j.jpubeco.2006.11.004

GSIA. (2018) *Trends Report 2018.* Retrieved May 12, 2021, from http://www.gsi-alliance.org/trends-report-2018/

Guenster, N., Bauer, R., Derwall, J., & Koedijk, K. (2011). The Economic Value of Corporate Eco-Efficiency. *European Financial Management*, *17*(4), Article 4. https://doi.org/10.1111/j.1468-036X.2009.00532.x

Hale, J. (2021, January 28). *A Broken Record: Flows for U.S. Sustainable Funds Again Reach New Heights*. Morningstar, Inc. https://www.morningstar.com/articles/1019195/a-broken-record-flows-for-us-sustainable-funds-again-reach-new-heights

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, *26*(3), 245–256. https://doi.org/10.1016/j.evolhumbehav.2005.01.002

Hart, O., & Zingales, L. (2017). Companies Should Maximize Shareholder Welfare Not Market Value. *Journal of Law, Finance, and Accounting*, *2*(2), 247–275. https://doi.org/10.1561/108.00000022

Hartzmark, S. M., & Sussman, A. B. (2019). Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows. *The Journal of Finance*, *74*(6), 2789–2837. https://doi.org/10.1111/jofi.12841

Harbaugh, W. T. (1998). The Prestige Motive for Making Charitable Transfers. *The American Economic Review*, *88*(2), 277–282.

Harrington, R. (2016, November 9). President-elect Donald Trump doesn't believe in climate change. Here's his platform on the environment. *Business Insider*. https://www.businessinsider.com/donald-trump-climate-change-global-warming-environment-policies-plans-platforms-2016-10

Heinkel, R., Kraus, A., & Zechner, J. (2001). The Effect of Green Investment on Corporate Behavior. *The Journal of Financial and Quantitative Analysis*, *36*(4), 431–449. https://doi.org/10.2307/2676219

Henry, D. (2017, February 6). Former EPA employees urge Senate to oppose Pruitt. *The Hill*. https://thehill.com/policy/energy-environment/318081-former-epa-employees-urge-senate-to-oppose-pruitt/

Hibbert, S., Smith, A., Davies, A., & Ireland, F. (2007). Guilt appeals: Persuasion knowledge and charitable giving. *Psychology & Marketing*, *24*(8), 723–742. https://doi.org/10.1002/mar.20181

Hichri, W., & Kirman, A. (2007). The emergence of coordination in public good games. *The European Physical Journal B*, *55*(2), 149–159. https://doi.org/10.1140/epjb/e2007-00032-8

Hickey, R., Minaker, B., & Payne, A. A. (2019). The sensitivity of charitable giving to the timing and salience of tax credits. *National Tax Journal*, *72*(1), 79–110. https://doi.org/10.17310/ntj.2019.1.03

Hoepner, A., Oikonomou, I., Scholtens, B., & Schröder, M. (2016). The Effects of Corporate and Country Sustainability Characteristics on the Cost of Debt: An International Investigation. *Journal of Business Finance & Accounting*, *43*(1–2), 158–190. https://doi.org/10.1111/jbfa.12183

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, Property Rights, and Anonymity in Bargaining Games. *Games and Economic Behavior*, *7*(3), 346–380. https://doi.org/10.1006/game.1994.1056

Hong, H., & Liskovich, I. (2015). *Crime, Punishment and the Halo Effect of Corporate Social Responsibility* (NBER Working Paper No. 21215). National Bureau of Economic Research, Inc. https://econpapers.repec.org/paper/nbrnberwo/21215.htm

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, *93*(1), 15–36. https://doi.org/10.1016/j.jfineco.2008.09.001

Hong, H., Wang, N., & Yang, J. (2020). *Mitigating Disaster Risks in the Age of Climate Change* (No. w27066). National Bureau of Economic Research. https://doi.org/10.3386/w27066

Hortaçsu, A., & Syverson, C. (2004). Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds. *The Quarterly Journal of Economics*, *119*(2), 403–456. https://doi.org/10.1162/0033553041382184

Hotten, R. (2015, December 10). Volkswagen: The scandal explained. *BBC News*. https://www.bbc.com/news/business-34324772

Huhmann, B. A., & Brotherton, T. P. (1997). A Content Analysis of Guilt Appeals in Popular Magazine Advertisements. *Journal of Advertising*, *26*(2), 35–45. https://doi.org/10.1080/00913367.1997.10673521

Jenni, K., & Loewenstein, G. (1997). Explaining the Identifiable Victim Effect. *Journal of Risk and Uncertainty*, *14*(3), 235–257. https://doi.org/10.1023/A:1007740225484

Katz, J. (2016, July 19). 2016 Election Forecast: Who Will Be President? *The New York Times*. https://www.nytimes.com/interactive/2016/upshot/presidential-polls-forecast.html

Kennedy, C., Blumenthal, M., Clement, S., Clinton, J. D., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., Saad, L., Witt, G. E., & Wlezien, C. (2018). An Evaluation of the 2016 Election Polls in the United States. *Public Opinion Quarterly*, *82*(1), 1–33. https://doi.org/10.1093/poq/nfx047

Kim, S., & Yoon, A. (2022). Analyzing Active Fund Managers' Commitment to ESG: Evidence from the United Nations Principles for Responsible Investment. *Management Science*. https://doi.org/10.1287/mnsc.2022.4394

Kirman, A., & Teschl, M. (2010). Selfish or selfless? The role of empathy in economics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1538), 303–317. https://doi.org/10.1098/rstb.2009.0192

Korschun, D., Bhattacharya, C. B., & Swain, S. D. (2014). Corporate Social Responsibility, Customer Orientation, and the Job Performance of Frontline Employees. *Journal of Marketing*, *78*(3), 20–37. https://doi.org/10.1509/jm.11.0245

Kotchen, M. J. (2006). 'Green Markets and Private Provision of Public Goods'. *Journal of Political Economy* 114 (4): 816–34. https://doi.org/10.1086/506337.

Krasteva, S., & Yildirim, H. (2016). Information, competition, and the quality of charities. *Journal of Public Economics*, *144*, 64–77. https://doi.org/10.1016/j.jpubeco.2016.10.008

Krueger, P. (2015). Corporate goodness and shareholder wealth. *Journal of Financial Economics*, *115*(2), 304–329. https://doi.org/10.1016/j.jfineco.2014.09.008

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The Importance of Climate Risks for Institutional Investors. *The Review of Financial Studies*, *33*(3), 1067–1111. https://doi.org/10.1093/rfs/hhz137

Landier, A., & Lovo, S. (2020). *ESG Investing: How to Optimize Impact?* (SSRN Scholarly Paper No. 3508938). https://doi.org/10.2139/ssrn.3508938

LeClair, M. S. (2019). Reported Instances of Nonprofit Corruption: Do Donors Respond to Scandals in the Charitable Sector? *Corporate Reputation Review*, *22*(2), 39–47. https://doi.org/10.1057/s41299-018-0056-5

Levine, D. K. (1998). Modeling Altruism and Spitefulness in Experiments. *Review of Economic Dynamics*, *1*(3), 593–622. https://doi.org/10.1006/redy.1998.0023

Levitt, S. D., & List, J. A. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives*, *21*(2), 153–174. https://doi.org/10.1257/jep.21.2.153

Liang, H., & Renneboog, L. (2020). Corporate Social Responsibility and Sustainable

Finance: A Review of the Literature. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3698631

Liebowitz, S. (2009, July 4). New Evidence on the Foreclosure Crisis. *Wall Street Journal*. https://www.wsj.com/articles/SB124657539489189043

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility during the Financial Crisis. *The Journal of Finance*, *72*(4), 1785–1824. https://doi.org/10.1111/jofi.12505

List, J. A. (2011). The Market for Charitable Giving. *Journal of Economic Perspectives*, *25*(2), 157–180. https://doi.org/10.1257/jep.25.2.157

Loewenstein, G., Rick, S., & Cohen, J. D. (2008). Neuroeconomics. *Annual Review of Psychology*, *59*(1), 647–672. https://doi.org/10.1146/annurev.psych.59.103006.093710

Luo, X., & Bhattacharya, C. B. (2006). Corporate Social Responsibility, Customer Satisfaction, and Market Value. *Journal of Marketing*, *70*(4), Article 4. https://doi.org/10.1509/jmkg.70.4.001

Majumdar, S. K., & Marcus, A. A. (2001). Rules versus Discretion: The Productivity Consequences of Flexible Regulation. *The Academy of Management Journal*, *44*(1), 170–179. https://doi.org/10.2307/3069344

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.2307/2975974

Marron, D. B., & Toder, E. J. (2014). Tax Policy Issues in Designing a Carbon Tax. *American Economic Review*, *104*(5), 563–568. https://doi.org/10.1257/aer.104.5.563

Mason, J. (2010, June 16). BP agrees to $20 billion spill fund. *Reuters*. https://www.reuters.com/article/uk-oil-spill-idUKTRE64U0OW20100616

Mattingly, J. E., & Berman, S. L. (2006). Measurement of Corporate Social Action: Discovering Taxonomy in the Kinder Lydenburg Domini Ratings Data. *Business & Society*, *45*(1), 20–46. https://doi.org/10.1177/0007650305281939

Mellström, C., & Johannesson, M. (2008). Crowding Out in Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, *6*(4), 845–863. https://doi.org/10.1162/JEEA.2008.6.4.845

Michel, G., & Rieunier, S. (2012). Nonprofit brand image and typicality influences on charitable giving. *Journal of Business Research*, *65*(5), 701–707. https://doi.org/10.1016/j.jbusres.2011.04.002

Mooney, A., & Mathurin, P. (2021, February 6). ESG funds defy havoc to ratchet huge inflows. *Financial Times*. https://www.ft.com/content/8e9f8204-83bf-4217-bc9e-d89396279c5b

Mueller, K., Hattrup, K., Spiess, S.-O., & Lin-Hi, N. (2012). The effects of corporate social responsibility on employees' affective commitment: A cross-cultural investigation. *Journal of Applied Psychology*, *97*, 1186–1200.

https://doi.org/10.1037/a0030204

Oehmke, M., & Opp, M. M. (2022). A Theory of Socially Responsible Investment. *Swedish House of Finance Research Paper Series*. https://doi.org/10.2139/ssrn.3467644

Okten, C., & Weisbrod, B. A. (2000). Determinants of donations in private nonprofit markets. *Journal of Public Economics*, *75*(2), 255–272. https://doi.org/10.1016/S0047-2727(99)00066-3

Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate Social and Financial Performance: A Meta-Analysis. *Organization Studies*, *24*(3), Article 3. https://doi.org/10.1177/0170840603024003910

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, *142*(2), 550–571. https://doi.org/10.1016/j.jfineco.2020.12.011

Payne, A. A. (2001). Measuring the Effect of Federal Research Funding on Private Donations at Research Universities: Is Federal Research Funding More than a Substitute for Private Donations? *International Tax and Public Finance*, *8*(5), 731–751. https://doi.org/10.1023/A:1012843227003

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, *142*(2), 572–597. https://doi.org/10.1016/j.jfineco.2020.11.001

Peloza, J., & Steel, P. (2005). The Price Elasticities of Charitable Contributions: A Meta-Analysis. *Journal of Public Policy & Marketing*, *24*(2), 260–272. https://doi.org/10.1509/jppm.2005.24.2.260

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20. https://doi.org/10.1017/S0140525X02000018

PRI. (2020). *Annual Report 2020*. PRI. Retrieved December 27, 2022, from https://www.unpri.org/about-the-pri/annual-report-2020/6811.article

PwC. (2020, April 11). 2022 — The growth opportunity of the century. *PwC*. Retrieved January, 6, 2023, from https://www.pwc.lu/en/sustainable-finance/esg-report-the-growth-opportunity-of-the-century.html

Quinson, T. (2020, September 5). Is Trump's Bid to Curb ESG to Protect Retirees or Oil? *Bloomberg*. https://www.bloomberg.com/news/articles/2020-09-05/is-trump-bid-to-curb-esg-to-protect-retirees-or-oil-quicktake

Ramelli, S., Ossola, E., & Rancan, M. (2021). Stock price effects of climate activism: Evidence from the first Global Climate Strike. *Journal of Corporate Finance*, *69*, 102018. https://doi.org/10.1016/j.jcorpfin.2021.102018

Ramelli, S., Wagner, A. F., Zeckhauser, R. J., & Ziegler, A. (2021). Investor Rewards to Climate Responsibility: Stock-Price Responses to the Opposite Shocks of the 2016 and 2020 U.S. Elections. *The Review of Corporate Finance Studies*, *10*(4), 748–787. https://doi.org/10.1093/rcfs/cfab010

Reece, W. S., & Zieschang, K. D. (1985). Consistent Estimation of the Impact of Tax Deductibility on the Level of Charitable Contributions. *Econometrica*, *53*(2), 271–293. https://doi.org/10.2307/1911236

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, *14*(3), 302–322. https://doi.org/10.1016/j.jcorpfin.2008.03.009

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, *32*(9), 1723–1742. https://doi.org/10.1016/j.jbankfin.2007.12.039

Renneboog, L., Ter Horst, J., & Zhang, C. (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, *20*(4), 562–588. https://doi.org/10.1016/j.jfi.2010.12.003

Riedl, A., & Smeets, P. (2017). Why Do Investors Hold Socially Responsible Mutual Funds? *The Journal of Finance*, *72*(6), 2505–2550. https://doi.org/10.1111/jofi.12547

Rosenberg, M., Confessore, N., & Cadwalladr, C. (2018, March 17). How Trump Consultants Exploited the Facebook Data of Millions. *The New York Times*. https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html

Roussanov, N., Ruan, H., & Wei, Y. (2021). Marketing Mutual Funds. *The Review of Financial Studies*, *34*(6), 3045–3094. https://doi.org/10.1093/rfs/hhaa095

Semenova, N., & Hassel, L. G. (2015). On the Validity of Environmental Performance Metrics. *Journal of Business Ethics*, *132*(2), 249–258. https://doi.org/10.1007/s10551-014-2323-4

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., & Frith, C. D. (2004). Brain Responses to the Acquired Moral Status of Faces. *Neuron*, *41*(4), 653–662. https://doi.org/10.1016/S0896-6273(04)00014-5

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but Not Sensory Components of Pain. *Science*, *303*(5661), 1157–1162.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, *439*(7075), Article 7075. https://doi.org/10.1038/nature04271

Sirri, E. R., & Tufano, P. (1998). Costly Search and Mutual Fund Flows. *The Journal of Finance* 53 (5): 1589–1622. https://doi.org/10.1111/0022-1082.00066.

Smith, A. (1759). *The Theory of Moral Sentiments*.

Statman, M. (2005). The Religions of Social Responsibility. *The Journal of Investing*, *14*(3), 14–21. https://doi.org/10.3905/joi.2005.580542

Steinberg, W. (1989). Depression: A discussion of Jung's ideas. *The Journal of Analytical Psychology*, *34*(4), 339–352. https://doi.org/10.1111/j.1465-5922.1989.00339.x

Taraporevala, C. (2021, January 11). *CEO's Letter on Our 2021 Proxy Voting Agenda*. https://www.ssga.com/international/en/institutional/ic/insights/ceo-letter-2021-proxy-voting-agenda

The Economist. (2008, October 23). *A helping hand to homeowners*. https://www.economist.com/finance-and-economics/2008/10/23/a-helping-hand-to-homeowners

Tocqueville, A. de. (1835). *Democracy in America: Volume I*. Saunders and Otley.

Tocqueville, A. de. (1840). *Democracy in America*. A. A. Knopf.

UBS. (2020). *Setting a new course*. UBS Investor Watch. https://www.ubs.com/us/en/investor-watch/setting-a-new-course.html

Ullmann, A. A. (1985). Data in Search of a Theory: A Critical Examination of the Relationships among Social Performance, Social Disclosure, and Economic Performance of U. S. Firms. *The Academy of Management Review*, *10*(3), Article 3. JSTOR. https://doi.org/10.2307/258135

Vogel, D. (2006). *The Market for Virtue*. 2nd ed. Brookings Institution Press. http://www.jstor.org/stable/10.7864/j.ctt6wpg2c.

Volcovici, V., & Shepardson, D. (2016, December 7). Trump to pick foe of Obama climate agenda to run EPA -source. *Reuters*. https://www.reuters.com/article/us-usa-trump-epa-idUSKBN13W2HE

Waddock, S. A., & Graves, S. B. (1997). The Corporate Social Performance-Financial Performance Link. *Strategic Management Journal*, *18*(4), 303–319.

Waite, S., Massa, A., & Cannon, C. (2019, August 8). Asset Managers With $74 Trillion on Brink of Historic Shakeout. *Bloomberg.* https://www.bloomberg.com/graphics/2019-asset-management-in-decline/

Watson, C. (2018, April 11). The key moments from Mark Zuckerberg's testimony to Congress. *The Guardian*. https://www.theguardian.com/technology/2018/apr/11/mark-zuckerbergs-testimony-to-congress-the-key-moments

White, H. K., Hsing, P.-Y., Cho, W., Shank, T. M., Cordes, E. E., Quattrini, A. M., Nelson, R. K., Camilli, R., Demopoulos, A. W. J., German, C. R., Brooks, J. M., Roberts, H. H., Shedd, W., Reddy, C. M., & Fisher, C. R. (2012). Impact of the Deepwater Horizon oil spill on a deep-water coral community in the Gulf of Mexico. *Proceedings of the National Academy of Sciences*, *109*(50), 20303–20308.

https://doi.org/10.1073/pnas.1118029109

Wood, D. J., & Jones, R. E. (1995). Stakeholder Mismatching: A Theoretical Problem in Empirical Research on Corporate Social Performance. *The International Journal of Organizational Analysis*, *3*(3), Article 3. https://doi.org/10.1108/eb028831

Wright, F. A., & Wright, A. A. (2018). How surprising was Trump's victory? Evaluations of the 2016 U.S. presidential election and a new poll aggregation model. *Electoral Studies*, *54*, 81–89. https://doi.org/10.1016/j.electstud.2018.05.001

Zivin, J. G., & Small, A. (2005). A Modigliani-Miller Theory of Altruistic Corporate Social Responsibility. *The B.E. Journal of Economic Analysis & Policy*, *5*(1). https://doi.org/10.1515/1538-0653.1369

Appendices

Appendix 1. Charity flow changes in response to shocks to altruism

Table 1. BP oil spill 2010 - UK charity voluntary income flow by size groups

	Baseline definition of $Treat_i$				Alternative definition of $Treat_i$			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A. Size Group 2 and 3								
$Post_t$	82.46	23.13	30.73	37.53	82.33	21.00	27.79	34.22
	(59.85)	(53.55)	(53.20)	(52.76)	(68.61)	(60.95)	(60.60)	(60.12)
$Treat_i$	-37.35	14.69	2.34	10.38	12.61	114.43	107.16	117.54^{*}
	(102.76)	(91.76)	(91.15)	(90.37)	(80.26)	(71.44)	(71.03)	(70.48)
$Post_t \times Treat_i$	-23.24	-67.21	-53.43	-60.61	-17.86	-39.77	-32.51	-36.06
	(127.73)	(114.00)	(113.23)	(112.25)	(99.44)	(88.17)	(87.65)	(86.93)
$Total\_income_{i,t}$		0.37^{***}	0.37^{***}	0.38^{***}		0.39^{***}	0.39^{***}	0.40^{***}
		(0.03)	(0.03)	(0.03)		(0.02)	(0.02)	(0.02)
$Volunteers_{i,t}$			0.68^{***}	0.67^{***}			0.63^{***}	0.62^{***}
			(0.19)	(0.18)			(0.18)	(0.17)
$Fixed\_assets_{i,t}$				-0.01^{***}				-0.01^{***}
				(0.00)				(0.00)
Intercept	542.82^{***}	-75.67	-108.45^{*}	-91.23	564.13^{***}	-108.01^{*}	-137.40^{**}	-126.07^{*}
	(48.53)	(60.59)	(60.83)	(60.44)	(55.71)	(64.73)	(64.85)	(64.37)
N	831	831	831	831	949	949	949	949
Adjusted R^2	-0.000	0.204	0.215	0.229	-0.001	0.213	0.223	0.236
B. All Sizes								
$Post_t$	93.68	11.04	26.01	26.29	90.44	11.83	28.62	28.82
	(118.80)	(101.73)	(101.72)	(101.76)	(136.23)	(115.94)	(115.82)	(115.86)
$Treat_i$	-100.91	-14.71	-7.84	-7.34	-46.57	-66.04	-62.32	-62.75
	(200.97)	(172.02)	(171.72)	(171.78)	(155.73)	(132.48)	(132.19)	(132.24)
$Post_t \times Treat_i$	357.58	346.38	353.79^{*}	353.45^{*}	182.61	150.31	152.39	152.26
	(250.21)	(214.11)	(213.74)	(213.81)	(193.78)	(164.86)	(164.49)	(164.54)
Total $income_{i,t}$		0.22^{***}	0.22^{***}	0.22^{***}		0.23^{***}	0.23^{***}	0.23^{***}
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
$Volunteers_{i,t}$			0.88^{**}	0.88^{**}			0.95^{***}	0.95^{***}
			(0.36)	(0.36)			(0.33)	(0.33)
Fixed $assets_{i,t}$				-0.00				-0.00
				(0.00)				(0.00)
Intercept	818.74^{***}	241.00^{***}	190.11^{**}	191.21^{**}	849.80^{***}	288.10^{***}	232.42^{**}	233.38^{**}
	(95.88)	(85.93)	(88.20)	(88.28)	(110.12)	(96.37)	(98.07)	(98.17)
N	1404	1404	1404	1404	1616	1616	1616	1616
Adjusted R^2	0.002	0.269	0.272	0.271	0.001	0.277	0.280	0.280

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $Post_t = 1$ if year = 2010, 2011 and $Post_t = 0$ if year = 2009. Baseline definition: $Treat_i = 1$ if a charity is classified by UK Charity Commission with classification number 112(Environment) and is not simultaneously classified with other charitable purposes codes: 103(Saving of Lives) 106(Overseas Aid) 107(Accommodation/Housing) 110(Amateur Sport) 113(Economic/Community Development/Employment) 114(Armed Forces/Emergency Service Efficiency) 115(Human Rights/Religious Or Racial Harmony/Equality or Diversity) 116(Recreation) 309(Acts As An Umbrella Or Resource Body). Alternative definition of $Treat_i = 1$ includes charities with code 111(Animals) additional to charities in the baseline definition, and is not simultaneously identified with the aforementioned other charitable purposes. $Treat_i = 0$ if a charity is classified with 109(Arts/Culture) and none of the other charitable purposes mentioned above. We employ logit propensity score matching using 1-year lagged covaraites: Total $assets_{i,t}$, Total $revenue_{i,t}$, Total functional $expense_{i,t}$, Total government $grants_{i,t}$, Fundraising $expense_{i,t}$. Payroll $tax_{i,t}$, as of shock year 2010. Matched sample is obtained by truncating the full sample to keep observations with logit propensity score between 0.3 and 0.8 (inclusive). Logit regression results, covariate balance test results and parallel trends are available upon request.

Table 2. US Paris Agreement withdrawal- UK charity voluntary income flow by size groups

	Baseline definition of $Treat_i$				Alternative definition of $Treat_i$			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A. Size Group 2 and 3								
$Post_t$	0.59	-149.64^{**}	-158.02^{**}	-150.32^{**}	-30.83	-181.11^{**}	-181.45^{**}	-176.23^{**}
	(91.03)	(63.07)	(61.32)	(60.29)	(101.17)	(70.81)	(70.81)	(69.89)
$Treat_i$	-199.51	51.01	80.89	105.14	2.58	266.68^{***}	267.16^{***}	286.32^{***}
	(141.50)	(98.09)	(95.46)	(93.95)	(103.78)	(72.98)	(72.98)	(72.13)
$Post_t \times Treat_i$	378.12^{*}	273.05^{**}	300.93^{**}	289.75^{**}	194.19	174.96^{*}	179.60^{*}	179.38^{*}
	(195.17)	(134.80)	(131.10)	(128.88)	(143.53)	(100.22)	(100.32)	(99.00)
$Total\_income_{i,t}$		0.58^{***}	0.62^{***}	0.64^{***}		0.59^{***}	0.59^{***}	0.62^{***}
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)

	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
$Volunteers_{i,t}$			-0.93***	-1.05***			-0.04	-0.05
			(0.14)	(0.14)			(0.04)	(0.04)
$Fixed\_assets_{i,t}$				-0.01***				-0.01***
				(0.00)				(0.00)
Intercept	579.54***	-526.18***	-484.49***	-468.64***	607.71***	-581.77***	-580.21***	-580.19***
	(65.47)	(59.63)	(58.31)	(57.40)	(72.42)	(63.94)	(63.96)	(63.11)
N	739	739	739	739	882	882	882	882
Adjusted R^2	0.002	0.525	0.551	0.566	0.002	0.513	0.514	0.526
B. All Sizes								
$Post_t$	-91.37	-236.92	-230.71	-226.60	-114.48	-211.55	-210.47	-206.79
	(262.72)	(213.85)	(213.28)	(213.24)	(282.63)	(233.73)	(233.67)	(233.65)
$Treat_i$	-529.48	-209.38	-237.41	-225.51	-323.48	-254.63	-254.72	-251.33
	(408.00)	(332.22)	(331.45)	(331.48)	(284.91)	(235.60)	(235.55)	(235.52)
$Post_t \times Treat_i$	1144.45**	869.11*	809.76*	809.47*	537.19	391.51	374.54	374.27
	(563.90)	(458.97)	(458.16)	(458.04)	(396.31)	(327.74)	(327.91)	(327.86)
$Total\_income_{i,t}$		0.30***	0.30***	0.30***		0.28***	0.28***	0.28***
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
$Volunteers_{i,t}$			1.44***	1.39***			0.21	0.20
			(0.49)	(0.49)			(0.16)	(0.16)
$Fixed\_assets_{i,t}$				-0.00				-0.00
				(0.00)				(0.00)
Intercept	1275.09***	101.18	-56.08	-33.89	1336.13***	312.44*	287.24*	302.42*
	(187.77)	(159.16)	(167.65)	(168.46)	(201.47)	(170.71)	(171.72)	(172.11)
N	1361	1361	1361	1361	1632	1632	1632	1632
Adjusted R^2	0.001	0.339	0.342	0.343	-0.000	0.316	0.316	0.317

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2016. Baseline definition: $Treat_i = 1$ if a charity is classified by UK Charity Commission with classification number 112(Environment) and is not simultaneously classified with other charitable purposes codes: 103(Saving of Lives) 106(Overseas Aid) 107(Accommodation/Housing) 110(Amateur Sport) 113(Economic/Community Development/Employment) 114(Armed Forces/Emergency Service Efficiency) 115(Human Rights/Religious Or Racial Harmony/Equality or Diversity) 116(Recreation) 309(Acts As An Umbrella Or Resource Body). Alternative definition of $Treat_i = 1$ includes charities with code 111(Animals) additional to charities in the baseline definition, and is not simultaneously identified with the aforementioned other charitable purposes. $Treat_i = 0$ if a charity is classified with 109(Arts/Culture) and none of the other charitable purposes mentioned above. We employ logit propensity score matching using 1-year lagged covaraites: Total $assets_{i,t}$, Total $revenue_{i,t}$, Total functional $expense_{i,t}$, Total government $grants_{i,t}$, Fundraising $expense_{i,t}$. Payroll $tax_{i,t}$ as of shock year 2016. Matched sample is obtained by truncating the full sample to keep observations with logit propensity score between 0.3 and 0.9 (inclusive). Logit regression results, covariate balance test results and parallel trends are available upon request.

Table 3. 2016 US presidential election - US charity voluntary income flow by size groups

	Baseline definition of $Treat_i$				Alternative definition of $Treat_i$			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A. Size Group 2 and 3								
$Post_t$	-12.33	-23.54***	-23.52***	-22.87***	-12.52	-24.17***	-24.00***	-23.51***
	(8.92)	(6.93)	(6.93)	(6.89)	(8.78)	(6.71)	(6.65)	(6.61)
$Treat_i$	10.67	18.70	18.88	18.82	0.25	11.89	12.47*	11.22
	(15.73)	(12.14)	(12.13)	(12.07)	(9.74)	(7.39)	(7.34)	(7.29)
$Post_t \times Treat_i$	16.54	27.78*	28.05*	28.24*	14.52	15.15	14.65	15.37*
	(19.77)	(15.41)	(15.41)	(15.32)	(12.15)	(9.29)	(9.22)	(9.16)
Total $assets_{i,t}$	0.03***	-0.02***	-0.02***	-0.02***	0.04***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Total $revenue_{i,t}$	0.46***	0.49***	0.49***	0.48***	0.46***	0.56***	0.55***	0.55***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Total functional $expense_{i,t}$		-0.25***	-0.25***	-0.22***		-0.33***	-0.34***	-0.30***
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Total government $grants_{i,t}$		0.14***	0.14***	0.14***		0.15***	0.15***	0.15***
		(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Fundraising $expense_{i,t}$			0.13	0.15			0.55***	0.56***
			(0.11)	(0.11)			(0.08)	(0.08)
Payroll $tax_{i,t}$				-1.72***				-1.64***
				(0.36)				(0.26)
Intercept	-12.33	-23.54***	-23.52***	-22.87***	38.25***	-4.51	-2.37	-2.04
	(8.92)	(6.93)	(6.93)	(6.89)	(8.23)	(6.45)	(6.40)	(6.36)
N	1973	1932	1932	1932	3305	3214	3214	3214
Adjusted R^2	0.351	0.618	0.618	0.622	0.353	0.633	0.638	0.643
B. All Sizes								
$Post_t$	-1.08	-10.27	-10.33	-11.13*	-1.11	-10.68*	-10.59*	-10.86*
	(8.76)	(6.52)	(6.52)	(6.48)	(8.08)	(5.91)	(5.90)	(5.90)
$Treat_i$	39.83***	35.47***	35.36***	35.75***	16.49*	18.64***	18.76***	18.96***

	(15.44)	(11.48)	(11.48)	(11.40)	(8.98)	(6.55)	(6.55)	(6.54)
$Post_t \times Treat_i$	17.81	16.36	16.35	14.87	7.03	5.02	4.74	4.29
	(19.76)	(14.68)	(14.68)	(14.59)	(11.38)	(8.31)	(8.31)	(8.30)
Total $assets_{i,t}$	0.05***	-0.02***	-0.02***	-0.03***	0.05***	-0.03***	-0.03***	-0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Total $revenue_{i,t}$	0.43***	0.58***	0.58***	0.59***	0.42***	0.62***	0.62***	0.62***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Total functional $expense_{i,t}$		-0.40***	-0.40***	-0.43***		-0.44***	-0.44***	-0.46***
		(0.02)	(0.02)	(0.02)		(0.01)	(0.01)	(0.01)
Total government $grants_{i,t}$		0.17***	0.17***	0.17***		0.17***	0.17***	0.17***
		(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Fundraising $expense_{i,t}$			-0.18	-0.21*			0.23***	0.22***
			(0.11)	(0.11)			(0.08)	(0.08)
Payroll $tax_{i,t}$				1.66***				0.74***
				(0.26)				(0.19)
Intercept	24.85***	-8.58	-8.22	-9.79*	24.29***	-12.21**	-12.55**	-12.80***
	(7.37)	(5.60)	(5.61)	(5.58)	(6.66)	(4.94)	(4.94)	(4.94)
N	3304	3304	3304	3304	5375	5375	5375	5375
Adjusted R^2	0.502	0.725	0.726	0.729	0.506	0.737	0.737	0.738

The table shows change in donation for the matched sample of treated and control charities in different size groups. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $Post_t = 1$ if year = 2016, 2017 and $Post_t = 0$ if year = 2015. In baseline definition, $Treat_i = 1$ if a charity has NTEE code C20 (Pollution Abatement & Control), C27 (Recycling), C30 (Natural Resources Conservation & Protection), C32 (Water Resources, Wetlands Conservation & Management). In alternative definition of $Treat_i$ = 1 include charities classified with an NTEE code D20 (Wildlife Preservation & Protection), D31 (Protection of Endangered Species), additional to charities defined under the baseline. For both baseline and alternative definitions, $Treat_i = 0$ if a charity is classified by NTEE with either code: A20 (Arts & Culture), A23 (Cultural & Ethnic Awareness). We employ logit propensity score matching using 1-year lagged covaraites: Total $assets_{i,t}$, Total $revenue_{i,t}$, Total functional $expense_{i,t}$, Total government $grants_{i,t}$, Fundraising $expense_{i,t}$. Payroll $tax_{i,t}$ as of shock year 2016. Matched sample is obtained by truncating the full sample to keep observations with logit propensity score between 0.5 and 0.79 (inclusive). Logit regression results, covariate balance test results and parallel trends are available upon request.

Table 4. US Paris Agreement withdrawal- US charity voluntary income flow by size groups

	Baseline definition of $Treat_i$				Alternative definition of $Treat_i$			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
A. Size Group 2 and 3								
$Post_t$	-5.51	-14.45***	-14.43***	-14.24***	-6.21	-14.53***	-14.45***	-14.43***
	(5.91)	(4.69)	(4.69)	(4.62)	(5.46)	(4.34)	(4.34)	(4.29)
$Treat_i$	22.32**	5.94	6.01	10.71	11.12*	-4.20	-4.01	-3.48
	(10.86)	(8.55)	(8.57)	(8.46)	(6.09)	(4.81)	(4.82)	(4.76)
$Post_t \times Treat_i$	13.07	21.07*	21.07*	21.37**	14.63*	18.52***	18.46***	18.82***
	(13.60)	(10.81)	(10.82)	(10.66)	(7.64)	(6.07)	(6.07)	(6.00)
Total $assets_{i,t}$	0.01	-0.01*	-0.01*	-0.01**	0.01**	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Total $revenue_{i,t}$	0.62***	0.72***	0.72***	0.72***	0.64***	0.74***	0.74***	0.74***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)
Total functional $expense_{i,t}$		-0.32***	-0.32***	-0.27***		-0.34***	-0.34***	-0.30***
		(0.03)	(0.03)	(0.03)		(0.02)	(0.02)	(0.02)
Total government $grants_{i,t}$		0.11***	0.11***	0.11***		0.12***	0.12***	0.11***
		(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Fundraising $expense_{i,t}$			-0.02	0.00			-0.09	-0.05
			(0.11)	(0.10)			(0.09)	(0.08)
Payroll $tax_{i,t}$				-2.26***				-2.01***
				(0.30)				(0.22)
Intercept	-3.95	-20.18***	-20.12***	-20.66***	-9.15*	-20.17***	-19.99***	-21.54***
	(6.04)	(4.83)	(4.85)	(4.78)	(5.24)	(4.21)	(4.22)	(4.17)
N	1915	1870	1870	1870	3183	3090	3090	3090
Adjusted R^2	0.485	0.685	0.685	0.694	0.497	0.691	0.691	0.699
B. All Sizes								
$Post_t$	-11.12*	-20.16***	-20.10***	-20.07***	-11.82**	-20.05***	-19.98***	-20.05***
	(6.19)	(5.14)	(5.15)	(5.04)	(5.64)	(4.64)	(4.64)	(4.57)
$Treat_i$	21.60*	6.80	7.01	10.06	9.70	-3.72	-3.49	-3.84
	(11.22)	(9.29)	(9.30)	(9.11)	(6.33)	(5.21)	(5.22)	(5.14)
$Post_t \times Treat_i$	22.37	26.21**	26.20**	27.10**	20.66***	22.58***	22.53***	23.08***
	(14.16)	(11.72)	(11.72)	(11.47)	(7.98)	(6.56)	(6.56)	(6.46)
Total $assets_{i,t}$	0.01*	-0.01***	-0.01***	-0.01***	0.01***	-0.01***	-0.01***	-0.01***

	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Total revenue$_{i,t}$	0.58***	0.53***	0.53***	0.54***	0.61***	0.59***	0.59***	0.59***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Total functional expense$_{i,t}$		-0.16***	-0.16***	-0.10***		-0.21***	-0.21***	-0.16***
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Total government grants$_{i,t}$		0.11***	0.11***	0.11***		0.11***	0.11***	0.11***
		(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Fundraising expense$_{i,t}$			-0.06	-0.01			-0.09	-0.05
			(0.12)	(0.12)			(0.09)	(0.09)
Payroll tax$_{i,t}$				-3.14***				-2.50***
				(0.32)				(0.23)
Intercept	11.88*	0.16	0.40	-1.51	4.26	-2.78	-2.54	-5.01
	(6.06)	(5.20)	(5.22)	(5.11)	(5.31)	(4.48)	(4.49)	(4.42)
N	2221	2221	2221	2221	3587	3587	3587	3587
Adjusted R^2	0.499	0.657	0.657	0.671	0.507	0.667	0.667	0.677

The table shows change in donation for the matched sample of treated and control charities in different size groups. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Post$_t$ = 1 if year = 2017, 2018 and Post$_t$ = 0 if year = 2016. . In baseline definition, Treat$_i$ = 1 if a charity has NTEE code C20 (Pollution Abatement & Control), C27 (Recycling), C30 (Natural Resources Conservation & Protection), C32 (Water Resources, Wetlands Conservation & Management). In alternative definition of Treat$_i$ = 1 include charities classified with an NTEE code D20 (Wildlife Preservation & Protection), D31 (Protection of Endangered Species), additional to charities defined under the baseline. For both baseline and alternative definitions, Treat$_i$ = 0 if a charity is classified by NTEE with either code: A20 (Arts & Culture), A23 (Cultural & Ethnic Awareness). We employ logit propensity score matching using 1-year lagged covaraites: Total assets$_{i,t}$, Total revenue$_{i,t}$, Total functional expense$_{i,t}$, Total government grants$_{i,t}$, Fundraising expense$_{i,t}$. Payroll tax$_{i,t,}$ as of shock year 2017. Matched sample is obtained by truncating the full sample to keep observations with logit propensity score between 0.4 and 0.9 (inclusive). Logit regression results, covariate balance test results and parallel trends are available upon request.

Appendix 2. UK fund flow changes in response to shocks to altruism

Table 1. BP oil spill 2010 – matched sample results

Panel A. Baseline $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	254.69***	237.75***	232.24***	43.42	51.98	54.37
	(60.05)	(57.61)	(57.94)	(37.19)	(36.78)	(37.08)
$Treat_i$	66.08***	45.58***	46.81***	22.54	23.18	29.87*
	(16.28)	(16.19)	(16.25)	(14.72)	(14.78)	(15.79)
$Post_t \times Treat_i$	-61.76***	-30.13*	-33.18*	-33.92*	-39.63**	-43.63**
	(17.70)	(16.61)	(16.95)	(17.45)	(17.47)	(18.62)
Lagged $RTN_{i,t}$	-1.33***	-1.60***	-1.57***	-0.04	-0.15	-0.15
	(0.50)	(0.45)	(0.46)	(0.35)	(0.34)	(0.34)
$RTN_{i,t}$	3.10***	2.79***	2.74***	0.50	0.43	0.44
	(0.60)	(0.57)	(0.57)	(0.33)	(0.33)	(0.33)
$Age_{i,t}$	0.40	-0.11	-0.13	-1.04*	-1.06*	-1.03*
	(0.61)	(0.55)	(0.55)	(0.55)	(0.55)	(0.55)
Lagged $Fundsize_{i,t}$	-0.08***	-0.10***	-0.10***	-0.09***	-0.09***	-0.09***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$Institutional_i$	-13.89	8.53	-6.34	-11.45	14.62	41.38
	(14.54)	(12.85)	(20.73)	(12.12)	(12.75)	(32.26)
Lagged Expense $Ratio_{i,t}$		43.31***	43.68***		30.12***	30.38***
		(5.26)	(5.28)		(5.33)	(5.34)
$Post_t \times Institutional_i$			22.10			-14.24
			(24.16)			(37.17)
$Treat_i \times Institutional_i$						-55.11
						(45.08)
$Post_t \times Treat_i \times Institutional_i$						30.63
						(52.84)
Lagged $\alpha_{i,t}$	35.55**	37.13**	36.89**	14.81	6.50	8.79
	(18.05)	(15.88)	(15.89)	(11.08)	(11.17)	(11.31)
Lagged $\beta_{mkt,i,t}$	69.13	178.64***	177.31***	-8.25	-45.24	-52.21
	(57.43)	(57.80)	(57.84)	(32.57)	(33.76)	(34.15)
Lagged $\beta_{smb,i,t}$	50.30**	105.90***	107.00***	21.51*	26.44**	28.26**
	(20.24)	(18.37)	(18.42)	(11.24)	(11.18)	(11.27)
Lagged $\beta_{hml,i,t}$	-29.11	46.03*	44.07*	-54.00***	-50.67***	-49.97***
	(26.34)	(24.17)	(24.28)	(17.09)	(16.90)	(16.93)
Lagged $\beta_{mom,i,t}$	70.99	-125.94***	-130.96***	-4.51	-31.12	-31.22
	(50.74)	(48.13)	(48.45)	(28.36)	(29.44)	(29.49)
Intercept	-296.10***	-474.25***	-469.44***	-3.76	-24.26	-23.04
	(72.20)	(77.32)	(77.52)	(41.23)	(41.24)	(41.37)
N	315	307	307	448	424	424
Adjusted R^2	0.307	0.493	0.493	0.358	0.411	0.410

Panel B. Alternative $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	-237.46***	-205.82***	-231.22***	-53.30	-45.77	-45.17
	(78.09)	(73.19)	(73.43)	(49.60)	(45.54)	(45.66)
$Treat_i$	98.47***	67.85***	70.89***	22.02	20.87	25.74*
	(14.59)	(14.40)	(14.33)	(14.56)	(13.74)	(14.74)
$Post_t \times Treat_i$	-118.00***	-58.50***	-69.75***	-20.06	-27.18	-25.79
	(20.05)	(19.24)	(19.76)	(19.49)	(18.22)	(19.65)
Lagged $RTN_{i,t}$	5.14***	3.39***	3.70***	1.22**	1.08**	1.10**
	(0.89)	(0.83)	(0.83)	(0.54)	(0.50)	(0.50)
$RTN_{i,t}$	5.19***	2.66***	2.73***	1.37**	1.15**	1.26**
	(0.91)	(0.88)	(0.87)	(0.56)	(0.52)	(0.53)
$Age_{i,t}$	0.16	-0.89	-0.89	-1.28*	-1.11*	-1.04
	(0.67)	(0.60)	(0.60)	(0.69)	(0.65)	(0.65)
Lagged $Fundsize_{i,t}$	0.05***	-0.00	0.00	-0.08***	-0.09***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
$Institutional_i$	-5.37	14.58	-10.43	-13.64	20.89	43.57
	(16.05)	(14.18)	(18.27)	(14.56)	(14.20)	(29.44)
Lagged Expense $Ratio_{i,t}$		39.61***	39.56***		38.34***	38.74***
		(5.28)	(5.23)		(5.41)	(5.41)
$Post_t \times Institutional_i$			49.72**			2.95
			(23.26)			(38.04)
$Treat_i \times Institutional_i$						-43.49
						(41.25)

$Post_t$ × $Treat_i$ × $Institutional_i$						-11.11
						(54.82)
Lagged $\alpha_{i,t}$	-29.29	-25.48	-28.73*	8.74	-5.13	-2.38
	(18.96)	(16.83)	(16.73)	(13.99)	(13.08)	(13.21)
Lagged $\beta_{mkt,i,t}$	-3.90	106.92	101.25	2.93	-31.67	-42.76
	(69.87)	(66.00)	(65.40)	(39.47)	(37.59)	(38.11)
Lagged $\beta_{smb,i,t}$	-58.43**	39.15	35.29	-1.37	4.48	6.03
	(26.62)	(26.25)	(26.06)	(15.95)	(14.78)	(14.87)
Lagged $\beta_{hml,i,t}$	-88.67***	-12.90	-20.41	-75.98***	-67.56***	-66.51***
	(29.55)	(27.65)	(27.60)	(19.12)	(17.65)	(17.65)
Lagged $\beta_{mom,i,t}$	124.38**	-30.94	-37.54	22.51	-16.81	-16.61
	(50.76)	(48.99)	(48.60)	(31.76)	(30.60)	(30.60)
Intercept	-3.29	-155.08*	-135.64	23.21	-11.30	-8.79
	(83.81)	(85.62)	(85.26)	(47.25)	(43.89)	(43.91)
N	196	192	192	287	273	273
Adjusted R^2	0.593	0.698	0.704	0.372	0.484	0.485

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Panel A: $Post_t$ = 1 if year = 2010, 2011 and $Post_t$ = 0 if year = 2009. In Panel B: $Post_t$ = 1 if year = 2010 and $Post_t$ = 0 if year = 2009. Additional test results using $Post_t$ = 1 if year = 2010, 2011 and $Post_t$ = 0 if year = 2009, 2010 is available upon request. In Column (iii), interaction terms: $Post_t$ × $Treat_i$ × $Institutional_i$ and $Treat_i$ × $Institutional_i$ is omitted from the regression due to lack of observations. In Columns (i) to (iii), we employ 1(treated)-to-2(control) matching by pre-shock fund characteristics for environmental SI funds (treated) and non-environmental SI funds (control); in columns (iv) to (vi), we employ 1(treated)-to-2(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i).

Table 2. US Paris Agreement withdrawal– matched sample results

Panel A. Baseline $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	-3.42	5.51	5.14	-3.06	-12.92	-22.30
	(3.15)	(4.80)	(5.03)	(8.90)	(17.57)	(18.25)
$Treat_i$	3.49	6.51	6.20	-4.92	-14.45	-10.42
	(2.90)	(4.62)	(4.88)	(8.27)	(14.62)	(15.89)
$Post_t$ × $Treat_i$	2.25	-3.65	-5.26	-3.11	10.60	22.07
	(3.50)	(6.42)	(6.89)	(9.63)	(18.37)	(19.88)
Lagged $RTN_{i,t}$	0.14	0.24	0.22	0.42	0.54	0.49
	(0.13)	(0.28)	(0.28)	(0.33)	(0.72)	(0.72)
$RTN_{i,t}$	0.12	0.07	0.06	0.22	0.18	0.09
	(0.09)	(0.18)	(0.18)	(0.24)	(0.50)	(0.50)
$Age_{i,t}$	-0.60***	-0.95***	-0.96***	-0.41	-1.50**	-1.63**
	(0.11)	(0.26)	(0.26)	(0.26)	(0.70)	(0.69)
Lagged $Fundsize_{i,t}$	-0.01***	-0.00	-0.00	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Institutional_i$	4.83**	0.58	-2.13	23.63***	30.33**	24.70
	(2.11)	(4.03)	(6.86)	(6.06)	(14.63)	(28.37)
Lagged Expense $Ratio_{i,t}$		-4.25*	-4.02		-0.09	1.54
		(2.56)	(2.60)		(9.35)	(9.28)
$Post_t$ × $Institutional_i$			2.97			79.32**
			(8.57)			(35.80)
$Treat_i$ × $Institutional_i$			-3.47			-24.50
			(10.95)			(37.03)
$Post_t$ × $Treat_i$ × $Institutional_i$			15.33			-81.14*
			(15.24)			(48.99)
Lagged $\alpha_{i,t}$	7.58**	3.65	2.61	17.53**	15.45	16.08
	(3.35)	(5.73)	(5.83)	(8.14)	(15.55)	(15.44)
Lagged $\beta_{mkt,i,t}$	-10.65	4.65	2.38	3.54	14.39	6.02
	(6.86)	(14.04)	(14.13)	(17.76)	(36.43)	(36.15)
Lagged $\beta_{smb,i,t}$	5.48*	-1.47	-0.75	11.26	2.35	4.49
	(3.00)	(5.75)	(5.78)	(7.77)	(17.62)	(17.45)
Lagged $\beta_{hml,i,t}$	-10.22**	12.30	12.66	8.92	32.19	28.10
	(3.98)	(7.66)	(7.68)	(9.66)	(22.79)	(22.62)
Lagged $\beta_{mom,i,t}$	5.53	-21.77**	-22.20**	-13.07	-38.55	-36.07
	(5.38)	(9.53)	(9.55)	(12.54)	(29.08)	(28.81)
Intercept	12.36*	4.16	6.63	5.65	17.23	24.38
	(6.80)	(14.24)	(14.36)	(17.76)	(39.55)	(39.44)
N	964	308	308	1351	524	524
Adjusted R^2	0.065	0.115	0.114	0.088	0.099	0.117

Panel B. Alternative $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	-1.65	-18.79	-19.88*	-23.45	-30.54	-53.86
	(4.61)	(12.08)	(11.93)	(15.48)	(37.10)	(37.13)
$Treat_i$	0.70	3.83	4.08	-13.76	-21.61	-17.99
	(2.68)	(5.45)	(5.55)	(9.33)	(17.42)	(18.72)
$Post_t$ × $Treat_i$	5.69	8.07	5.22	7.40	23.66	44.50
	(3.60)	(8.85)	(9.05)	(12.21)	(25.48)	(27.10)
Lagged $RTN_{i,t}$	-0.19	0.02	0.07	-0.19	0.05	-0.01
	(0.17)	(0.37)	(0.36)	(0.49)	(1.13)	(1.10)
$RTN_{i,t}$	-0.02	0.84**	0.81**	0.62	0.50	0.50
	(0.14)	(0.40)	(0.39)	(0.47)	(1.11)	(1.09)
$Age_{i,t}$	-0.45***	-0.95***	-0.92***	-0.63*	-1.74*	-1.92**
	(0.12)	(0.34)	(0.34)	(0.36)	(0.97)	(0.95)
Lagged $Fundsize_{i,t}$	-0.00*	-0.01*	-0.01**	-0.03***	-0.03***	-0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Institutional_i$	6.77***	4.58	-1.36	31.69***	45.86**	24.36
	(2.32)	(5.30)	(7.37)	(8.36)	(20.40)	(33.22)
Lagged Expense $Ratio_{i,t}$		-2.15	-1.85		1.06	4.01
		(3.26)	(3.23)		(12.39)	(12.18)
$Post_t$ × $Institutional_i$			14.62			158.39***
			(12.32)			(47.71)
$Treat_i$ × $Institutional_i$			-2.61			-22.55
			(11.76)			(42.79)
$Post_t$ × $Treat_i$ × $Institutional_i$			45.44*			-149.78**
			(23.33)			(69.66)
Lagged $\alpha_{i,t}$	9.14**	11.10	9.75	33.27***	28.39	29.71
	(4.37)	(11.66)	(11.61)	(12.32)	(26.01)	(25.54)
Lagged $\beta_{mkt,i,t}$	-12.54	24.67	25.22	5.89	13.33	2.57
	(7.99)	(19.63)	(19.41)	(26.26)	(55.75)	(54.71)
Lagged $\beta_{smb,i,t}$	3.69	-8.63	-9.52	18.73*	9.01	11.87
	(3.36)	(7.58)	(7.51)	(11.07)	(26.23)	(25.71)
Lagged $\beta_{hml,i,t}$	4.83	13.29	14.50	11.62	32.31	27.11
	(4.42)	(9.25)	(9.08)	(14.09)	(31.26)	(30.67)
Lagged $\beta_{mom,i,t}$	-4.42	-10.89	-14.17	1.16	-26.90	-19.30
	(6.25)	(13.44)	(13.23)	(18.31)	(43.92)	(43.10)
Intercept	13.99*	-14.47	-14.19	17.91	25.99	36.64
	(7.91)	(19.88)	(19.54)	(25.88)	(59.25)	(58.28)
N	634	224	224	883	370	370
Adjusted R^2	0.058	0.113	0.146	0.136	0.106	0.142

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Panel A: $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2016. In Panel B: $Post_t = 1$ if year = 2017 and $Post_t = 0$ if year = 2016. Additional test results using $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2015, 2016 is available upon request. In Columns (i) to (iii), we employ 1(treated)-to-2(control) matching by pre-shock fund characteristics for environmental SI funds (treated) and non-environmental SI funds (control); in columns (iv) to (vi), we employ 1(treated)-to-1(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i).

Table 3. Oxfam scandal 2011 – matched sample results

	Baseline $Post_t$ definition			Alternative $Post_t$ definition		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	31.53*	33.59	35.35	31.11	29.80	29.51
	(19.02)	(22.60)	(23.57)	(27.75)	(29.84)	(31.18)
$Treat_i$	-13.90	-14.16	-11.95	-23.08	-21.51	-19.51
	(12.31)	(13.92)	(14.85)	(14.54)	(15.99)	(17.02)
$Post_t$ × $Treat_i$	1.23	-1.97	-5.89	4.57	3.63	0.48
	(14.48)	(16.52)	(17.71)	(19.14)	(20.92)	(22.35)
Lagged $RTN_{i,t}$	0.93**	0.82*	0.84*	-0.36	-0.65	-0.67
	(0.39)	(0.49)	(0.49)	(0.77)	(0.85)	(0.86)
$RTN_{i,t}$	0.92***	0.95***	0.97***	2.51***	2.61***	2.66***
	(0.31)	(0.36)	(0.37)	(0.73)	(0.80)	(0.80)
$Age_{i,t}$	-0.53	-0.51	-0.52	-0.43	-0.47	-0.48
	(0.40)	(0.50)	(0.50)	(0.59)	(0.63)	(0.63)
Lagged $Fundsize_{i,t}$	-0.05***	-0.04***	-0.04***	-0.05***	-0.04***	-0.04***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$Institutional_i$	-12.60	0.38	3.01	-16.80	2.01	0.24
	(9.58)	(12.14)	(30.96)	(13.90)	(16.58)	(35.03)
Lagged Expense $Ratio_{i,t}$		14.79**	14.76**		18.41**	18.52**
		(5.79)	(5.82)		(7.73)	(7.79)
$Post_t$ × $Institutional_i$			-5.72			4.99
			(35.47)			(43.82)

$Treat_i \times Institutional_i$			-18.80			-18.88
			(42.06)			(47.37)
$Post_t \times Treat_i \times Institutional_i$			31.63			31.41
			(49.38)			(62.07)
Lagged $\alpha_{i,t}$	-1.38	-1.33	-2.02	26.03	27.61	27.36
	(8.78)	(12.96)	(13.04)	(18.18)	(20.91)	(21.04)
Lagged $\beta_{mkt,i,t}$	-21.06	-35.46	-36.07	9.45	-9.82	-9.06
	(26.31)	(31.02)	(31.16)	(39.32)	(43.76)	(43.99)
Lagged $\beta_{smb,i,t}$	-6.80	-2.25	-2.37	0.42	7.03	7.97
	(11.37)	(13.90)	(14.02)	(19.23)	(20.44)	(20.71)
Lagged $\beta_{hml,i,t}$	-2.27	7.71	7.99	-1.59	7.50	8.23
	(15.52)	(18.27)	(18.38)	(21.39)	(23.14)	(23.33)
Lagged $\beta_{mom,i,t}$	1.39	-12.36	-12.10	-33.50	-53.87	-55.65
	(28.01)	(37.75)	(38.00)	(44.06)	(50.33)	(50.75)
Intercept	3.07	-10.10	-10.71	11.13	0.10	-0.75
	(33.28)	(38.77)	(39.41)	(46.46)	(50.72)	(51.54)
N	500	405	405	322	292	292
Adjusted R^2	0.210	0.183	0.178	0.180	0.194	0.187

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Baseline definition: $Post_t = 1$ if year = 2011, 2012 and $Post_t = 0$ if year = 2010. Alternative definition: $Post_t = 1$ if year = 2011 and $Post_t = 0$ if year = 2010. Additional test results using $Post_t = 1$ if year = 2011, 2012 and $Post_t = 0$ if year = 2009, 2010 is available upon request. We employ 1(treated)-to-2(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i).

Appendix 3. US fund flow changes in response to shocks to altruism

Table 1. BP oil spill 2010 – matched sample results

Panel A. Baseline $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	74.50	72.97	60.40	104.15*	106.38**	97.18*
	(50.41)	(50.44)	(51.83)	(54.12)	(53.91)	(54.51)
$Treat_i$	47.92**	47.24**	37.27*	58.62***	57.56***	72.31***
	(19.05)	(19.06)	(21.42)	(20.50)	(20.42)	(22.24)
$Post_t \times Treat_i$	-44.14*	-42.90*	-30.01	-20.80	-20.60	-19.27
	(22.70)	(22.74)	(25.75)	(24.70)	(24.60)	(27.03)
Lagged $RTN_{i,t}$	-1.18*	-1.17*	-1.16*	-1.17	-1.19	-1.13
	(0.69)	(0.69)	(0.70)	(0.75)	(0.75)	(0.75)
$RTN_{i,t}$	-0.49	-0.49	-0.50	0.33	0.35	0.36
	(0.58)	(0.58)	(0.58)	(0.60)	(0.60)	(0.60)
$Age_{i,t}$	-1.64*	-1.75*	-1.76*	-7.89***	-7.59***	-7.69***
	(0.90)	(0.90)	(0.90)	(0.92)	(0.92)	(0.92)
Lagged $Fundsize_{i,t}$	-0.02***	-0.02***	-0.02***	-0.00**	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
$Institutional_i$	-23.02*	-29.84**	-69.48**	-13.14	3.07	24.19
	(13.22)	(15.03)	(33.79)	(15.23)	(16.40)	(43.20)
Lagged Expense $Ratio_{i,t}$		-11.60	-11.49		29.61***	28.38**
		(12.18)	(12.21)		(11.38)	(11.34)
$Post_t \times Institutional_i$			54.36			37.70
			(40.25)			(50.21)
$Treat_i \times Institutional_i$			46.85			-76.23
			(45.00)			(54.95)
$Post_t \times Treat_i \times Institutional_i$			-60.72			-12.98
			(54.63)			(65.51)
Lagged $\alpha_{i,t}$	116.87***	112.41***	112.82***	156.81***	163.82***	161.27***
	(25.21)	(25.64)	(25.73)	(22.46)	(22.53)	(22.46)
Lagged $\beta_{mkt,i,t}$	-63.50*	-69.18**	-69.57**	-194.30***	-187.13***	-196.55***
	(33.63)	(34.16)	(34.24)	(48.78)	(48.66)	(48.62)
Lagged $\beta_{smb,i,t}$	31.81	36.67	35.94	57.88***	50.57**	51.82**
	(25.31)	(25.83)	(25.99)	(21.09)	(21.20)	(21.13)
Lagged $\beta_{hml,i,t}$	-45.24	-51.56	-53.36	-31.12	-20.18	-14.27
	(32.24)	(32.92)	(33.06)	(30.15)	(30.33)	(30.33)
Lagged $\beta_{mom,i,t}$	-12.70	1.78	-1.26	155.97***	132.53***	149.58***
	(50.39)	(52.64)	(53.60)	(48.62)	(49.26)	(49.41)
Intercept	48.94	73.57	82.95	163.86**	109.82	121.00*
	(50.96)	(57.15)	(57.96)	(64.90)	(67.89)	(68.24)
N	363	363	363	732	732	732
Adjusted R^2	0.134	0.134	0.131	0.198	0.204	0.211

Panel B. Alternative $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	23.73	24.17	27.33	-107.90	-114.16*	-109.72
	(65.24)	(65.45)	(67.27)	(67.39)	(66.90)	(67.65)
$Treat_i$	42.89**	42.82**	32.70	58.74***	57.34***	71.08***
	(19.95)	(20.00)	(22.35)	(20.22)	(20.06)	(21.82)
$Post_t \times Treat_i$	-51.55*	-51.34*	-32.38	-13.56	-13.33	-13.11
	(26.91)	(27.00)	(30.47)	(27.85)	(27.63)	(30.44)
Lagged $RTN_{i,t}$	0.07	0.05	-0.07	2.69***	2.83***	2.84***
	(0.98)	(0.99)	(0.99)	(1.01)	(1.00)	(1.00)
$RTN_{i,t}$	1.72	1.69	1.81	4.90***	5.07***	5.21***
	(1.14)	(1.16)	(1.17)	(1.04)	(1.03)	(1.03)
$Age_{i,t}$	-1.38	-1.40	-1.40	-8.53***	-8.15***	-8.28***
	(1.16)	(1.17)	(1.17)	(1.10)	(1.10)	(1.10)
Lagged $Fundsize_{i,t}$	-0.03***	-0.03***	-0.03***	-0.00**	-0.00**	-0.00*
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
$Institutional_i$	-50.18***	-51.52***	-68.24*	-50.88***	-29.16	33.25
	(16.75)	(19.00)	(35.21)	(18.47)	(19.77)	(42.54)
Lagged Expense $Ratio_{i,t}$		-2.35	-2.06		39.10***	38.41***
		(15.59)	(15.65)		(13.39)	(13.37)
$Post_t \times Institutional_i$			31.36			-30.29
			(47.73)			(55.55)
$Treat_i \times Institutional_i$			45.47			-92.21*
			(46.47)			(53.85)

$Post_t$ × $Treat_i$ × $Institutional_i$			-87.04			21.03
			(64.83)			(73.07)
Lagged $\alpha_{i,t}$	166.45***	165.37***	168.73***	177.58***	187.58***	185.51***
	(34.44)	(35.25)	(35.48)	(28.63)	(28.61)	(28.59)
Lagged $\beta_{mkt,i,t}$	-99.30**	-100.06**	-102.30**	-282.89***	-276.09***	-281.43***
	(42.94)	(43.33)	(43.46)	(60.37)	(59.94)	(60.04)
Lagged $\beta_{smb,i,t}$	45.00	46.17	46.29	16.07	4.38	3.05
	(32.46)	(33.44)	(33.75)	(26.40)	(26.49)	(26.48)
Lagged $\beta_{hml,i,t}$	-2.45	-3.79	-6.24	35.59	50.98	58.30
	(38.36)	(39.46)	(39.64)	(35.41)	(35.53)	(35.62)
Lagged $\beta_{mom,i,t}$	107.66	109.21	107.61	326.07***	303.90***	318.63***
	(68.75)	(69.66)	(70.74)	(57.72)	(57.77)	(58.09)
Intercept	70.29	75.02	72.30	276.78***	209.47***	206.10***
	(57.97)	(66.07)	(66.95)	(71.58)	(74.66)	(75.12)
N	244	244	244	487	487	487
Adjusted R^2	0.208	0.204	0.201	0.283	0.294	0.297

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Panel A: $Post_t$ = 1 if year = 2010, 2011 and $Post_t$ = 0 if year = 2009. In Panel B: $Post_t$ = 1 if year = 2010 and $Post_t$ = 0 if year = 2009. Additional test results using $Post_t$ = 1 if year = 2010, 2011 and $Post_t$ = 0 if year = 2009, 2010 is available upon request. In Columns (i) to (iii), we match environmental SI funds (treated) and non-environmental SI funds (control) by their pre-shock fund characteristics, by excluding funds without data on Lagged $\alpha_{i,t}$ we obtain a matched sample; in columns (iv) to (vi), we employ 1(treated)-to-2(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i).

Table 2. US Paris Agreement withdrawal– matched sample results

Panel A. Baseline $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	-38.10	-38.59	-40.48	26.11	26.31	14.89
	(23.70)	(23.74)	(28.18)	(47.20)	(47.25)	(50.64)
$Treat_i$	33.12*	32.31	35.74	28.26	25.38	19.50
	(19.74)	(19.76)	(25.33)	(40.82)	(41.21)	(48.10)
$Post_t$ × $Treat_i$	-15.70	-14.29	-37.88	-27.91	-26.63	-24.47
	(24.61)	(24.67)	(31.48)	(49.16)	(49.25)	(57.77)
Lagged $RTN_{i,t}$	2.91***	3.07***	3.22***	0.52	0.54	0.52
	(1.09)	(1.10)	(1.10)	(2.12)	(2.12)	(2.13)
$RTN_{i,t}$	1.37**	1.45**	1.55**	0.70	0.70	0.72
	(0.61)	(0.61)	(0.62)	(1.16)	(1.16)	(1.17)
$Age_{i,t}$	-4.39***	-4.50***	-4.32***	-16.82***	-16.96***	-16.92***
	(0.94)	(0.95)	(0.96)	(1.64)	(1.65)	(1.66)
Lagged $Fundsize_{i,t}$	-0.01***	-0.01***	-0.01***	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Institutional_i$	28.81**	37.41**	19.19	-36.77	-32.31	-76.96
	(13.31)	(14.69)	(31.47)	(27.12)	(28.51)	(70.19)
Lagged Expense $Ratio_{i,t}$		19.23	16.51		12.40	12.22
		(13.98)	(14.18)		(22.78)	(22.81)
$Post_t$ × $Institutional_i$			1.88			52.63
			(35.51)			(83.56)
$Treat_i$ × $Institutional_i$			-13.83			31.84
			(42.30)			(93.03)
$Post_t$ × $Treat_i$ × $Institutional_i$			64.59			-22.52
			(50.68)			(112.55)
Lagged $\alpha_{i,t}$	44.86	61.00*	50.68	215.07***	221.44***	221.68***
	(30.91)	(33.06)	(34.22)	(68.60)	(69.54)	(69.82)
Lagged $\beta_{mkt,i,t}$	-16.93	-1.39	-21.59	148.94	150.41	150.76
	(50.13)	(51.41)	(53.30)	(98.79)	(98.91)	(99.52)
Lagged $\beta_{smb,i,t}$	-39.77*	-47.81*	-39.13	-93.97*	-97.66*	-96.27*
	(23.70)	(24.44)	(25.54)	(51.07)	(51.53)	(51.71)
Lagged $\beta_{hml,i,t}$	15.25	24.47	23.17	86.83	86.27	84.98
	(25.83)	(26.69)	(26.75)	(54.90)	(54.96)	(55.07)
Lagged $\beta_{mom,i,t}$	-55.68	-69.66	-60.91	-224.12**	-226.82**	-229.47**
	(66.58)	(67.39)	(68.31)	(101.30)	(101.48)	(101.88)
Intercept	49.79	10.90	36.08	122.85	110.03	118.92
	(49.58)	(57.00)	(61.42)	(90.67)	(93.76)	(94.89)
N	675	673	673	920	919	919
Adjusted R^2	0.078	0.079	0.081	0.384	0.384	0.382

Panel B. Alternative $Post_t$ definition

	Environmental vs Non-Environmental SI			SI vs Conventional funds		

	funds					
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	-85.06**	-115.13***	-119.32***	-14.54	-17.70	-20.10
	(38.89)	(40.11)	(43.13)	(73.42)	(73.99)	(76.85)
$Treat_i$	30.76	31.15	40.50	26.14	23.98	12.89
	(19.77)	(19.63)	(25.38)	(41.29)	(41.75)	(48.70)
$Post_t$ × $Treat_i$	-9.11	-7.80	-26.34	-21.06	-20.71	-24.49
	(28.18)	(27.97)	(35.97)	(56.49)	(56.54)	(66.35)
Lagged $RTN_{i,t}$	3.99**	5.62***	5.64***	1.00	1.16	1.05
	(1.96)	(2.03)	(2.05)	(3.38)	(3.41)	(3.43)
$RTN_{i,t}$	4.18***	5.86***	6.00***	2.87	3.04	3.02
	(1.52)	(1.63)	(1.63)	(2.85)	(2.89)	(2.90)
$Age_{i,t}$	-4.21***	-4.51***	-4.49***	-17.58***	-17.68***	-17.60***
	(1.15)	(1.15)	(1.17)	(2.01)	(2.03)	(2.04)
Lagged $Fundsize_{i,t}$	-0.00	0.00	0.00	-0.03***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
$Institutional_i$	15.85	35.17**	32.03	-65.71**	-62.03*	-97.91
	(16.36)	(17.68)	(32.36)	(33.31)	(34.86)	(70.85)
Lagged Expense $Ratio_{i,t}$		46.37***	46.03***		9.83	9.56
		(16.75)	(16.96)		(27.25)	(27.31)
$Post_t$ × $Institutional_i$			5.69			18.62
			(40.71)			(96.11)
$Treat_i$ × $Institutional_i$			-27.95			44.78
			(42.61)			(93.49)
$Post_t$ × $Treat_i$ × $Institutional_i$			55.92			6.25
			(57.81)			(129.46)
Lagged $\alpha_{i,t}$	76.59**	106.60***	103.23**	300.31***	302.79***	305.71***
	(37.86)	(39.11)	(40.16)	(85.24)	(85.58)	(86.07)
Lagged $\beta_{mkt,i,t}$	-135.37**	-105.22*	-109.95*	103.74	103.71	110.76
	(61.95)	(62.46)	(64.30)	(123.39)	(123.47)	(124.40)
Lagged $\beta_{smb,i,t}$	-58.87**	-84.27***	-84.00***	-95.83	-99.41	-96.00
	(29.56)	(30.74)	(32.25)	(60.87)	(61.72)	(62.04)
Lagged $\beta_{hml,i,t}$	1.71	9.65	6.87	80.24	78.95	77.70
	(34.28)	(34.15)	(34.32)	(67.83)	(67.97)	(68.13)
Lagged $\beta_{mom,i,t}$	-112.18	-147.34*	-152.85*	-327.07**	-328.07**	-334.84**
	(83.66)	(84.01)	(85.86)	(129.49)	(129.61)	(130.29)
Intercept	136.96**	35.80	39.87	183.53*	172.46	173.49
	(61.49)	(71.15)	(74.68)	(111.17)	(115.41)	(116.28)
N	462	462	462	614	614	614
Adjusted R^2	0.067	0.080	0.079	0.411	0.411	0.408

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Panel A: $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2016. In Panel B: $Post_t = 1$ if year = 2017 and $Post_t = 0$ if year = 2016. Additional test results using $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2015, 2016 is available upon request. In Columns (i) to (iii), we employ 1(treated)-to-1(control) matching by pre-shock fund characteristics for environmental SI funds (treated) and non-environmental SI funds (control); in columns (iv) to (vi), we employ 1(treated)-to-1(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i).

Table 3.2016 US presidential election – matched sample results

Panel A. Baseline $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	-13.55	-13.30	-46.62	-49.34	-49.20	-65.79
	(25.92)	(25.92)	(32.28)	(43.68)	(43.59)	(48.49)
$Treat_i$	6.07	7.40	-32.21	-2.04	-7.50	-53.12
	(22.62)	(22.65)	(29.18)	(41.88)	(41.86)	(49.39)
$Post_t$ × $Treat_i$	2.83	1.70	21.74	68.49	68.13	74.28
	(27.63)	(27.64)	(36.29)	(50.41)	(50.30)	(59.76)
Lagged $RTN_{i,t}$	-0.20	-0.21	-0.52	0.98	1.13	1.08
	(0.99)	(0.99)	(1.00)	(1.66)	(1.66)	(1.65)
$RTN_{i,t}$	1.30	1.39	1.45	1.05	1.38	1.30
	(1.01)	(1.02)	(1.01)	(1.48)	(1.48)	(1.48)
$Age_{i,t}$	-4.21***	-4.24***	-4.00***	-5.89***	-6.56***	-6.59***
	(0.92)	(0.92)	(0.92)	(1.46)	(1.48)	(1.48)
Lagged $Fundsize_{i,t}$	0.00	0.00	0.00	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Institutional_i$	-22.30	-14.23	-89.51***	-83.71***	-63.59**	-179.22***
	(14.05)	(16.00)	(33.53)	(26.62)	(27.97)	(64.71)
Lagged Expense $Ratio_{i,t}$		16.81	13.64		52.46**	51.19**
		(15.94)	(15.89)		(22.82)	(22.77)

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$ × $Institutional_i$			62.14			70.44
			(39.43)			(76.70)
$Treat_i$ × $Institutional_i$			90.97*			158.63*
			(47.24)			(91.34)
$Post_t$ × $Treat_i$ × $Institutional_i$			-35.64			-32.17
			(56.80)			(110.06)
Lagged $\alpha_{i,t}$	167.55***	171.96***	172.79***	189.32***	210.82***	211.85***
	(37.81)	(38.04)	(37.86)	(52.64)	(53.35)	(53.27)
Lagged $\beta_{mkt,i,t}$	-151.80**	-151.85**	-210.72***	-160.33	-126.16	-102.47
	(72.61)	(72.60)	(74.88)	(97.88)	(98.80)	(98.98)
Lagged $\beta_{smb,i,t}$	22.76	19.61	30.32	-5.70	-24.06	-28.37
	(33.88)	(34.01)	(34.02)	(39.78)	(40.49)	(40.43)
Lagged $\beta_{hml,i,t}$	39.29	44.42	32.41	188.38***	203.11***	210.32***
	(36.09)	(36.41)	(36.46)	(44.52)	(44.89)	(44.90)
Lagged $\beta_{mom,i,t}$	47.66	39.84	22.06	-134.67	-140.21	-146.64
	(60.91)	(61.35)	(62.21)	(114.50)	(114.29)	(114.09)
Intercept	222.71***	198.99**	295.97***	273.11***	182.35*	194.95*
	(76.77)	(79.99)	(86.45)	(101.83)	(109.02)	(109.57)
N	529	529	529	1050	1050	1050
Adjusted R^2	0.078	0.079	0.089	0.209	0.213	0.216

Panel B. Alternative $Post_t$ definition

	Environmental vs Non-Environmental SI funds			SI vs Conventional funds		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
$Post_t$	-52.12	-50.27	-93.88**	-91.53	-85.30	-86.48
	(31.85)	(31.75)	(39.13)	(58.25)	(58.15)	(61.16)
$Treat_i$	1.49	3.97	-33.08	13.01	6.60	-37.41
	(18.38)	(18.36)	(24.02)	(37.95)	(37.95)	(44.75)
$Post_t$ × $Treat_i$	22.23	21.07	57.40*	73.21	71.51	68.00
	(25.58)	(25.50)	(34.06)	(52.19)	(52.04)	(61.96)
Lagged $RTN_{i,t}$	-2.24	-2.19	-3.34	-0.25	0.37	0.28
	(2.58)	(2.57)	(2.64)	(4.22)	(4.21)	(4.20)
$RTN_{i,t}$	2.04	2.08*	1.70	3.38	3.73*	3.63*
	(1.25)	(1.24)	(1.26)	(2.14)	(2.14)	(2.15)
$Age_{i,t}$	-3.78***	-3.83***	-3.76***	-6.25***	-6.88***	-6.96***
	(0.91)	(0.91)	(0.91)	(1.59)	(1.61)	(1.61)
Lagged $Fundsize_{i,t}$	0.02***	0.02***	0.02***	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Institutional_i$	-38.05***	-23.11	-84.66***	-111.21***	-89.74***	-172.89***
	(13.84)	(15.94)	(28.96)	(29.43)	(30.85)	(58.93)
Lagged Expense $Ratio_{i,t}$		29.34*	26.59*		56.26**	55.76**
		(15.71)	(15.70)		(24.99)	(24.92)
$Post_t$ × $Institutional_i$			70.99*			15.22
			(37.93)			(79.56)
$Treat_i$ × $Institutional_i$			86.23**			152.00*
			(39.08)			(82.55)
$Post_t$ × $Treat_i$ × $Institutional_i$			-73.35			0.17
			(53.28)			(113.53)
Lagged $\alpha_{i,t}$	141.07***	149.59***	164.01***	152.67**	168.93**	174.21**
	(53.42)	(53.42)	(54.18)	(73.67)	(73.81)	(73.65)
Lagged $\beta_{mkt,i,t}$	-304.78***	-304.29***	-376.98***	-155.78	-97.81	-63.00
	(82.56)	(82.27)	(86.75)	(115.18)	(117.69)	(118.11)
Lagged $\beta_{smb,i,t}$	43.18	39.05	48.57	3.03	-12.36	-17.90
	(36.04)	(35.98)	(36.04)	(52.62)	(52.91)	(52.79)
Lagged $\beta_{hml,i,t}$	2.90	14.95	-5.60	171.02***	192.29***	203.32***
	(37.04)	(37.47)	(38.36)	(55.04)	(55.68)	(55.69)
Lagged $\beta_{mom,i,t}$	45.50	35.98	11.86	-32.53	-26.76	-29.95
	(64.49)	(64.46)	(64.95)	(124.41)	(124.07)	(123.72)
Intercept	374.56***	331.50***	446.12***	276.51**	153.11	146.29
	(89.09)	(91.72)	(101.85)	(120.38)	(131.94)	(131.82)
N	362	362	362	702	702	702
Adjusted R^2	0.166	0.172	0.181	0.207	0.212	0.216

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Panel A: $Post_t = 1$ if year = 2016, 2017 and $Post_t = 0$ if year = 2015. In Panel B: $Post_t = 1$ if year = 2016 and $Post_t = 0$ if year = 2015. Additional test results using $Post_t = 1$ if year = 2016, 2017 and $Post_t = 0$ if year = 2014, 2015 is available upon request. In Columns (i) to (iii), we employ 1(treated)-to-1(control) matching by pre-shock fund characteristics for environmental SI funds (treated) and non-environmental SI funds (control); in columns (iv) to (vi), we employ 1(treated)-to-1(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i).

Table 4. Tax Cut and Jobs Act 2017 – matched sample results

	Baseline $Post_t$ definition			Alternative $Post_t$ definition		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	26.11	26.31	14.89	-14.54	-17.70	-20.10
	(47.20)	(47.25)	(50.64)	(73.42)	(73.99)	(76.85)
$Treat_i$	28.26	25.38	19.50	26.14	23.98	12.89
	(40.82)	(41.21)	(48.10)	(41.29)	(41.75)	(48.70)
$Post_t \times Treat_i$	-27.91	-26.63	-24.47	-21.06	-20.71	-24.49
	(49.16)	(49.25)	(57.77)	(56.49)	(56.54)	(66.35)
Lagged $RTN_{i,t}$	0.52	0.54	0.52	1.00	1.16	1.05
	(2.12)	(2.12)	(2.13)	(3.38)	(3.41)	(3.43)
$RTN_{i,t}$	0.70	0.70	0.72	2.87	3.04	3.02
	(1.16)	(1.16)	(1.17)	(2.85)	(2.89)	(2.90)
$Age_{i,t}$	-16.82***	-16.96***	-16.92***	-17.58***	-17.68***	-17.60***
	(1.64)	(1.65)	(1.66)	(2.01)	(2.03)	(2.04)
Lagged $Fundsize_{i,t}$	-0.02***	-0.02***	-0.02***	-0.03***	-0.03***	-0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Institutional_i$	-36.77	-32.31	-76.96	-65.71**	-62.03*	-97.91
	(27.12)	(28.51)	(70.19)	(33.31)	(34.86)	(70.85)
Lagged Expense $Ratio_{i,t}$		12.40	12.22		9.83	9.56
		(22.78)	(22.81)		(27.25)	(27.31)
$Post_t \times Institutional_i$			52.63			18.62
			(83.56)			(96.11)
$Treat_i \times Institutional_i$			31.84			44.78
			(93.03)			(93.49)
$Post_t \times Treat_i \times Institutional_i$			-22.52			6.25
			(112.55)			(129.46)
Lagged $\alpha_{i,t}$	215.07***	221.44***	221.68***	300.31***	302.79***	305.71***
	(68.60)	(69.54)	(69.82)	(85.24)	(85.58)	(86.07)
Lagged $\beta_{mkt,i,t}$	148.94	150.41	150.76	103.74	103.71	110.76
	(98.79)	(98.91)	(99.52)	(123.39)	(123.47)	(124.40)
Lagged $\beta_{smb,i,t}$	-93.97*	-97.66*	-96.27*	-95.83	-99.41	-96.00
	(51.07)	(51.53)	(51.71)	(60.87)	(61.72)	(62.04)
Lagged $\beta_{hml,i,t}$	86.83	86.27	84.98	80.24	78.95	77.70
	(54.90)	(54.96)	(55.07)	(67.83)	(67.97)	(68.13)
Lagged $\beta_{mom,i,t}$	-224.12**	-226.82**	-229.47**	-327.07**	-328.07**	-334.84**
	(101.30)	(101.48)	(101.88)	(129.49)	(129.61)	(130.29)
Intercept	122.85	110.03	118.92	183.53*	172.46	173.49
	(90.67)	(93.76)	(94.89)	(111.17)	(115.41)	(116.28)
N	920	919	919	614	614	614
Adjusted R^2	0.384	0.384	0.382	0.411	0.411	0.408

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Baseline definition: $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2016. Alternative definition: $Post_t = 1$ if year = 2017 and $Post_t = 0$ if year = 2016. Additional test results using $Post_t = 1$ if year = 2017, 2018 and $Post_t = 0$ if year = 2015, 2016 is available upon request. We employ 1(treated)-to-1(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged $RTN_{i,t}$), contemporaneous return ($RTN_{i,t}$), 1-year lagged fund size (Lagged $Fundsize_{i,t}$, in millions), lagged expense ratio (Expense $Ratio_{i,t}$) and fund age (Age_i).

Table 5. Tax Cut and Jobs Act 2017 – matched sample using high Morningstar Globes Rating as SI definition

	Baseline $Post_t$ definition			Alternative $Post_t$ definition		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
$Post_t$	34.49	47.56	62.38	-39.59	-47.85*	-59.24**
	(36.82)	(37.07)	(38.11)	(26.08)	(26.09)	(27.33)
$Treat_i$	20.88	15.04	28.06	44.15***	36.15**	50.18***
	(16.73)	(16.82)	(19.31)	(15.76)	(15.81)	(18.07)
$Post_t \times Treat_i$	-37.16	-34.81	-34.65	-21.81	-20.64	-18.89
	(24.24)	(24.32)	(28.13)	(22.58)	(22.57)	(25.85)
Lagged $RTN_{i,t}$	1.18	1.28	1.26	2.18*	2.54**	2.62**
	(0.87)	(0.87)	(0.87)	(1.18)	(1.18)	(1.18)
$RTN_{i,t}$	0.41	0.94	1.03	1.26	1.70*	1.83*
	(1.20)	(1.21)	(1.22)	(0.93)	(0.94)	(0.94)
$Age_{i,t}$	-12.27***	-12.59***	-12.64***	-11.64***	-12.07***	-12.15***
	(0.72)	(0.73)	(0.73)	(0.68)	(0.68)	(0.69)
Lagged $Fundsize_{i,t}$	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Institutional_i$	6.40	21.32	68.37**	13.97	34.67**	43.85*
	(14.06)	(14.59)	(26.59)	(13.46)	(14.04)	(26.13)
Lagged Expense $Ratio_{i,t}$		38.00***	37.80***		45.10***	44.01***
		(9.28)	(9.28)		(9.15)	(9.16)
$Post_t \times Institutional_i$			-50.89			38.45

			(37.01)			(36.12)
Treat$_i$ × Institutional$_i$			-53.90			-59.50
			(39.14)			(37.22)
Post$_t$ × Treat$_i$ × Institutional$_i$			7.30			-5.35
			(55.83)			(52.86)
Lagged $\alpha_{i,t}$	50.65*	61.00**	62.17**	12.79	19.31	20.51
	(29.76)	(29.90)	(29.91)	(27.22)	(27.23)	(27.24)
Lagged $\beta_{mkt,i,t}$	59.97	34.54	34.66	-49.16	-86.78**	-87.93**
	(47.97)	(48.50)	(48.49)	(42.83)	(43.54)	(43.53)
Lagged $\beta_{smb,i,t}$	20.86	11.43	9.49	-11.29	-26.77	-29.43
	(20.72)	(20.88)	(20.90)	(18.16)	(18.44)	(18.46)
Lagged $\beta_{hml,i,t}$	2.70	8.32	10.62	38.45*	33.19	33.38
	(29.34)	(29.43)	(29.46)	(22.71)	(22.70)	(22.69)
Lagged $\beta_{mom,i,t}$	-15.45	-2.82	-6.41	-3.56	16.53	14.27
	(44.34)	(44.51)	(44.54)	(41.76)	(41.96)	(41.94)
Intercept	54.84	30.62	18.99	152.41***	138.68***	139.00***
	(41.22)	(41.66)	(41.98)	(39.44)	(39.47)	(39.72)
N	3622	3604	3604	3833	3821	3821
Adjusted R^2	0.170	0.173	0.174	0.149	0.154	0.155

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Baseline definition: Post$_t$ = 1 if year = 2018 and Post$_t$ = 0 if year = 2017. Alternative definition: Post$_t$ = 1 if year = 2017 and Post$_t$ = 0 if year = 2016. To define treatment group, we obtain monthly Morningstar Sustainability Globes rating (earliest available from August 2018), which scales from 1 to 5 globes (indicating Low, Below-average, Average, Above-average, and High sustainability grades). Treat$_i$ is a dummy variable equal to 1 if fund i's annual average Morningstar Sustainability Globes rating in year 2018 is equal to 5; and 0 if fund i's annual average Morningstar Sustainability Globes rating in year 2018 is below or equal 2.5. Additional test results using Post$_t$ = 1 if year = 2017, 2018 and Post$_t$ = 0 if year = 2015, 2016 is available upon request. We employ 1(treated)-to-1(control) matching by pre-shock fund characteristics for SI funds (treated) and conventional funds (control). Control variables include 1-year lagged holding period return (Lagged RTN$_{i,t}$), contemporaneous return (RTN$_{i,t}$), 1-year lagged fund size (Lagged Fundsize$_{i,t}$, in millions), lagged expense ratio (Expense Ratio$_{i,t}$) and fund age (Age$_i$). Logit propensity score estimation, covariate balancing results and parallel trend graphs are available upon request.

Appendix 4. Covariate Balancing and Parallel Trends- UK funds

Table 1. BP oil spill 2010 – covariate balancing test

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	38.81	8.22	-0.08	0.94	0.35	0.00	-0.02
Treated	38.53	10.16	-0.29	0.87	0.65	-0.14	-0.07
t-stat	0.15	-1.65	2.67	2.30	-5.55	3.39	2.30
Median Test							
Control	35.88	5.00	-0.04	0.94	0.26	-0.04	0.04
Treated	37.74	9.00	-0.38	0.81	0.71	-0.08	-0.05
z-score	-1.47	-2.58	3.50	3.76	-5.13	2.11	3.03
N_control	94	115	70	70	70	70	70
N_treated	68	74	59	59	59	59	59
Panel B. After Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	37.73	10.08	-0.32	0.87	0.58	-0.14	-0.08
Treated	38.72	12.43	-0.34	0.84	0.64	-0.10	-0.09
t-stat	-0.97	-2.11	0.38	1.69	-1.67	-1.86	0.43
Median Test							
Control	36.49	7.00	-0.36	0.82	0.65	-0.12	-0.05
Treated	37.76	10.00	-0.38	0.81	0.71	-0.06	-0.06
z-score	-1.95	-4.24	0.92	2.32	-2.80	-3.54	1.96
N_control	102	102	102	102	102	102	102
N_treated	51	51	51	51	51	51	51
Panel C. Before Matching (SI vs Conventional Funds)							
Mean Test							
Control	42.38	8.62	-0.04	0.93	0.44	-0.13	-0.06
Treated	38.69	8.98	-0.18	0.90	0.49	-0.06	-0.05
t-stat	3.03	-0.55	3.82	1.50	-1.39	-2.88	-1.56
Median Test							
Control	39.19	6.00	-0.07	0.92	0.42	-0.14	-0.06
Treated	36.57	7.00	-0.22	0.89	0.47	-0.06	-0.03
z-score	3.36	-1.68	3.69	2.03	-2.37	-3.41	-1.81
N_control	1,829	2,204	1,458	1,458	1,458	1,458	1,458
N_treated	162	189	129	129	129	129	129
Panel D. After Matching (SI vs Conventional Funds)							
Mean Test							
Control	38.24	11.13	-0.13	0.93	0.51	-0.04	-0.03
Treated	37.97	11.17	-0.18	0.90	0.50	-0.06	-0.05
t-stat	0.25	-0.04	1.31	2.18	0.25	1.09	0.88
Median Test							
Control	36.70	8.00	-0.18	0.96	0.49	-0.04	-0.02
Treated	36.54	9.00	-0.25	0.86	0.49	-0.06	-0.03
z-score	0.27	-0.94	1.27	2.79	-0.35	0.57	0.92
N_control	254	254	254	254	254	254	254
N_treated	127	127	127	127	127	127	127

This table presents balance test results using 1-year lagged covariates as of the shock year 2010. Sample in Panel B is obtained by truncating the before matching sample (Panel A) to keep observations with logit propensity score between 0.15 and 0.2 (inclusive) or between 0.3 and 0.9 (inclusive). Sorting by propensity score, each environmental SI fund is matched by the closest two non-environmental funds. Sample in Panel D is obtained by truncating the before matching sample (Panel C) to keep observations with logit propensity score between 0 and 0.5 (inclusive). Sorting by propensity score, each SI fund is matched by the closest two conventional funds. Logit propensity score estimation results are shown in Appendix 6.

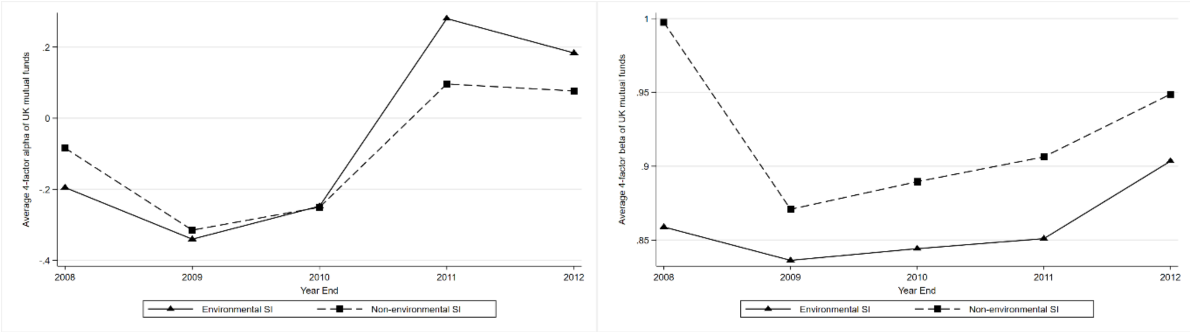


Figure 1. BP oil spill – Parallel trends on 4-Factor alpha (left) and 4-Factor market beta (right)

Table 2. US Paris Agreement withdrawal– covariate balancing test

	Lagged RTN$_{i,t}$	Age$_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	-2.15	9.39	-0.03	0.96	0.20	-0.21	-0.15
Treated	-4.86	8.87	-0.14	0.92	0.34	-0.33	-0.25
t-stat	3.27	0.63	3.42	1.85	-2.96	3.09	2.93
Median Test							
Control	-2.51	8.00	-0.05	0.99	0.14	-0.22	-0.08
Treated	-6.13	4.00	-0.11	0.97	0.35	-0.36	-0.28
z-score	3.89	1.19	2.76	2.26	-3.23	3.57	3.14
N_control	188	228	154	154	154	154	154
N_treated	139	184	124	124	124	124	124
Panel B. After Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	-4.38	9.35	-0.09	0.94	0.33	-0.32	-0.26
Treated	-3.39	10.67	-0.09	0.92	0.27	-0.28	-0.23
t-stat	-1.34	-1.70	-0.27	1.00	1.46	-1.24	-0.94
Median Test							
Control	-4.71	9.00	-0.13	0.99	0.26	-0.29	-0.13
Treated	-3.56	8.00	-0.06	0.97	0.26	-0.26	-0.21
z-score	-0.26	-0.41	-2.05	1.82	1.37	-0.97	-0.95
N_control	186	186	186	186	186	186	186
N_treated	93	93	93	93	93	93	93
Panel C. Before Matching (SI vs Conventional Funds)							
Mean Test							
Control	-3.47	8.41	-0.05	0.94	0.31	-0.23	-0.21
Treated	-3.30	9.16	-0.08	0.94	0.26	-0.26	-0.19
t-stat	-0.35	-1.76	2.03	-0.61	1.93	1.35	-0.70
Median Test							
Control	-4.09	5.00	-0.05	0.95	0.24	-0.26	-0.14
Treated	-3.32	6.00	-0.07	0.97	0.16	-0.23	-0.13
z-score	-1.47	-2.01	1.48	-1.64	1.40	0.82	-0.28
N_control	4,498	5,377	3,671	3,671	3,671	3,671	3,671
N_treated	327	412	278	278	278	278	278
Panel D. After Matching (SI vs Conventional Funds)							
Mean Test							
Control	-2.39	9.77	-0.06	0.93	0.21	-0.22	-0.17
Treated	-3.57	10.94	-0.07	0.94	0.27	-0.26	-0.20
t-stat	1.49	-1.57	0.44	-1.28	-1.66	1.35	1.05
Median Test							
Control	-3.52	6.50	-0.06	0.94	0.12	-0.22	-0.11
Treated	-2.99	8.50	-0.06	0.97	0.18	-0.24	-0.13
z-score	-0.15	-2.00	-0.03	-1.87	-1.95	1.33	1.13
N_control	264	264	264	264	264	264	264
N_treated	264	264	264	264	264	264	264

This table presents balance test results using 1-year lagged covariates as of the shock year 2017. Sample in Panel B is obtained by truncating the before matching sample (Panel A) to keep observations with logit propensity score between 0 and 0.6 (inclusive) or between 0.68 and 1 (inclusive). Sorting by propensity score, each environmental SI fund is matched by the closest two non-environmental funds. Sample in Panel D is obtained by truncating the before matching sample (Panel C) to keep observations with logit propensity score between 0 and 0.15 (inclusive). Sorting by propensity score, each SI fund is matched by the closest one conventional fund. Logit propensity score estimation results are shown in Appendix 6.

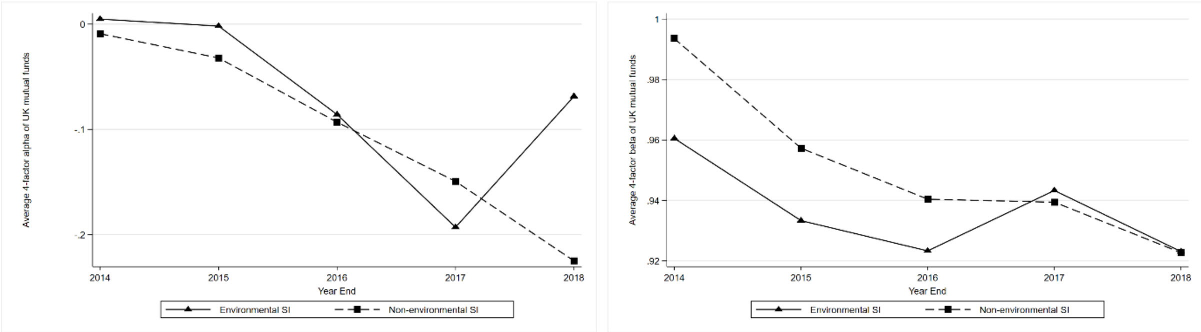


Figure 2.2016 US presidential election shock– Parallel trends on 4-Factor alpha (left) and 4-Factor market beta (right)

Table 3. Oxfam scandal 2011 – covariate balancing test

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (SI vs Conventional Funds)							
Mean Test							
Control	11.79	8.95	0.03	0.95	0.46	-0.04	-0.01
Treated	10.34	9.26	-0.18	0.92	0.49	-0.02	-0.02
t-stat	1.78	-0.50	5.80	2.04	-1.01	-1.17	0.61
Median Test							
Control	10.38	6.00	-0.02	0.94	0.42	-0.04	-0.01
Treated	10.69	7.00	-0.21	0.90	0.41	0.04	-0.00
z-score	0.93	-1.43	6.48	2.19	-1.75	-3.09	0.33
N_control	1,993	2,388	1,680	1,680	1,680	1,680	1,680
N_treated	167	205	150	150	150	150	150
Panel B. After Matching (SI vs Conventional Funds)							
Mean Test							
Control	11.87	12.65	-0.13	0.93	0.55	-0.04	-0.03
Treated	11.41	10.83	-0.13	0.92	0.50	-0.02	-0.01
t-stat	0.40	1.30	-0.06	0.65	0.88	-0.94	-1.42
Median Test							
Control	10.47	8.00	-0.14	0.91	0.49	-0.06	-0.02
Treated	11.22	10.00	-0.17	0.91	0.41	0.03	0.01
z-score	-0.11	-0.92	0.06	0.68	0.61	-2.07	-1.65
N_control	132	132	132	132	132	132	132
N_treated	132	132	132	132	132	132	132

This table presents balance test results using 1-year lagged covariates as of the shock year 2011. Sample in Panel B is obtained by truncating the before matching sample (Panel A) to keep observations with logit propensity score between 0 and 0.3 (inclusive). Sorting by propensity score, each SI fund is matched by the closest one conventional fund. Logit propensity score estimation results are shown in Appendix 6.

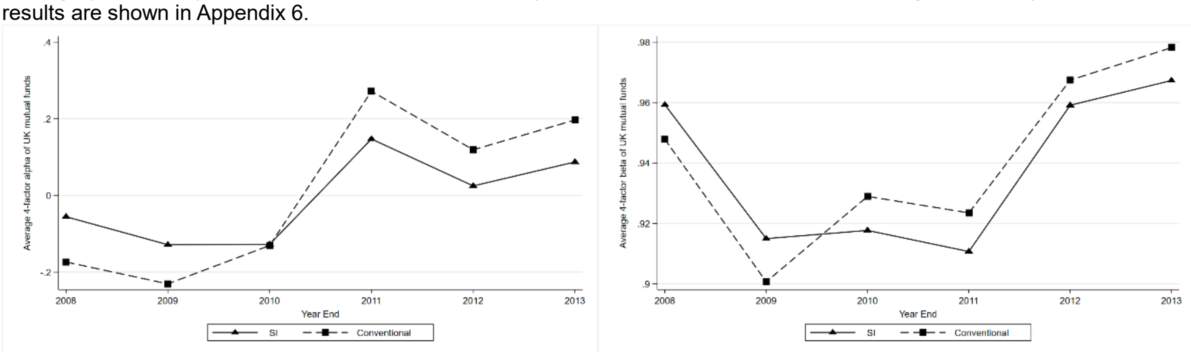


Figure 3. Oxfam scandal – Parallel trends on 4-Factor alpha (left) and 4-Factor market beta (right)

Appendix 5. Covariate Balancing and Parallel Trends– US funds

Table 1. BP oil spill 2010 – covariate balancing test

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	29.84	8.12	-0.04	1.01	0.15	-0.12	-0.00
Treated	32.23	8.12	-0.02	0.99	0.24	-0.15	-0.00
t-stat	-1.46	0.00	-0.46	0.70	-2.07	1.12	-0.11
Median Test							
Control	30.67	7.00	-0.10	1.02	0.07	-0.13	0.01
Treated	31.53	8.00	0.01	1.04	0.16	-0.15	0.01
z-score	-0.58	0.12	0.20	-1.10	-1.79	1.09	0.32
N_control	81	91	69	69	69	69	69
N_treated	96	109	79	79	79	79	79
Panel B. After Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	30.06	9.25	-0.04	1.01	0.15	-0.12	-0.00
Treated	29.78	9.95	-0.02	0.99	0.24	-0.15	-0.00
t-stat	0.17	-0.72	-0.46	0.70	-2.07	1.12	-0.11
Median Test							
Control	30.86	9.00	-0.10	1.02	0.07	-0.13	0.01
Treated	30.03	9.00	0.01	1.04	0.16	-0.15	0.01
z-score	0.73	-0.93	0.20	-1.10	-1.79	1.09	0.32
N_control	68	69	69	69	69	69	69
N_treated	76	79	79	79	79	79	79
Panel C. Before Matching (SI vs Conventional Funds)							
Mean Test							
Control	33.69	9.65	0.04	1.03	0.19	-0.14	-0.02
Treated	31.14	8.12	-0.03	1.00	0.20	-0.14	-0.00
t-stat	2.71	2.42	2.53	2.80	-0.39	-0.07	-1.83
Median Test							
Control	31.59	8.00	0.02	1.03	0.09	-0.12	-0.02
Treated	30.69	7.00	-0.03	1.02	0.15	-0.13	0.01
z-score	2.40	2.30	3.02	2.03	-1.46	0.22	-2.65
N_control	6,159	6,869	5,720	5,720	5,720	5,720	5,720
N_treated	177	200	148	148	148	148	148
Panel D. After Matching (SI vs Conventional Funds)							
Mean Test							
Control	30.67	9.44	0.02	1.00	0.16	-0.15	-0.00
Treated	30.72	9.70	-0.03	1.02	0.20	-0.13	-0.01
t-stat	-0.04	-0.35	1.41	-1.31	-1.19	-0.68	0.73
Median Test							
Control	30.38	9.00	0.02	1.01	0.07	-0.11	-0.00
Treated	30.86	9.00	-0.06	1.03	0.16	-0.12	0.01
z-score	-0.59	-0.41	1.59	-1.38	-1.95	-0.28	0.01
N_control	138	138	138	138	138	138	138
N_treated	138	138	138	138	138	138	138

This table presents balance test results using 1-year lagged covariates as of the shock year 2010. Sample in Panel B is obtained by dropping observations that do not have data on Lagged $\alpha_{i,t}$ from the before matching sample (Panel A). Sample in Panel D is obtained by truncating the before matching sample (Panel C) to keep observations with logit propensity score between 0 and 0.1 (inclusive). Sorting by propensity score, each SI fund is matched by the closest one conventional fund. Logit propensity score estimation results are shown in Appendix 6.

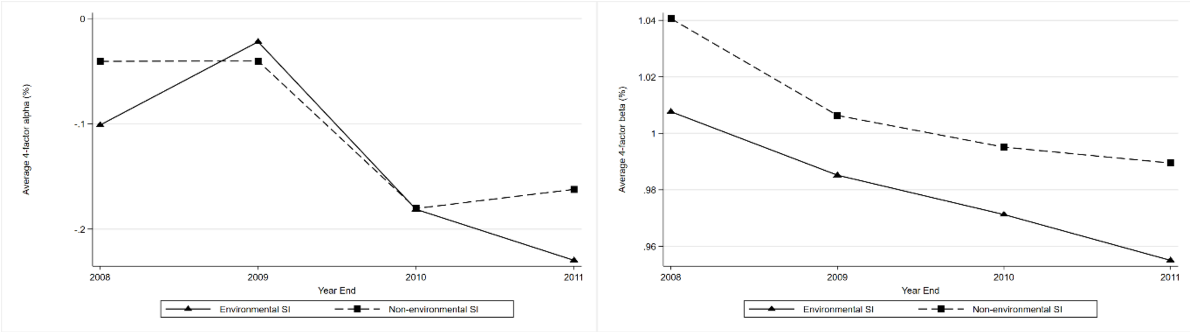


Figure 1. BP oil spill – Parallel trends on 4-Factor alpha (left) and 4-Factor market beta (right)

Table 2. US Paris Agreement withdrawal– covariate balancing test

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	9.46	12.83	-0.16	0.99	0.04	-0.08	-0.03
Treated	9.77	9.83	-0.13	0.94	0.18	-0.07	-0.01
t-stat	-0.32	3.33	-1.04	2.48	-3.01	-0.39	-0.98
Median Test							
Control	9.61	12.00	-0.12	0.99	-0.04	-0.02	-0.00
Treated	10.40	8.00	-0.12	0.94	0.01	-0.00	0.00
z-score	-1.14	3.58	-0.49	3.15	-2.60	-2.02	-1.78
N_control	72	111	80	80	80	80	80
N_treated	143	193	144	144	144	144	144
Panel B. After Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	10.44	10.93	-0.08	0.93	0.12	-0.06	-0.01
Treated	9.71	11.30	-0.13	0.95	0.18	-0.07	-0.02
t-stat	0.86	-0.44	1.78	-0.63	-1.43	0.40	0.35
Median Test							
Control	10.07	11.00	-0.04	0.96	0.09	-0.01	0.01
Treated	10.39	10.00	-0.12	0.94	0.01	-0.01	0.00
z-score	-0.13	0.31	2.74	-0.46	-0.37	-1.14	-0.57
N_control	124	124	124	124	124	124	124
N_treated	124	124	124	124	124	124	124
Panel C. Before Matching (SI vs Conventional Funds)							
Mean Test							
Control	11.27	12.14	-0.14	0.93	0.16	-0.05	-0.03
Treated	9.66	10.92	-0.14	0.96	0.13	-0.07	-0.02
t-stat	2.82	2.19	-0.23	-2.43	1.42	1.31	-0.63
Median Test							
Control	9.67	11.00	-0.14	0.97	0.01	-0.06	-0.00
Treated	9.94	10.00	-0.12	0.97	-0.00	-0.01	0.00
z-score	1.62	1.93	-0.79	-1.51	1.53	0.53	0.57
N_control	8,034	10,523	8,522	8,522	8,522	8,522	8,522
N_treated	215	304	224	224	224	224	224
Panel D. After Matching (SI vs Conventional Funds)							
Mean Test							
Control	10.10	10.71	-0.16	0.92	0.09	-0.05	-0.03
Treated	9.98	12.66	-0.17	0.94	0.08	-0.04	-0.01
t-stat	0.15	-2.30	0.43	-1.14	0.34	-0.49	-0.95
Median Test							
Control	8.24	10.00	-0.14	0.94	-0.02	-0.08	-0.02
Treated	9.86	11.00	-0.14	0.96	-0.01	-0.01	0.00
z-score	-0.64	-2.71	-0.20	-1.09	-0.30	-1.06	-1.83
N_control	161	161	161	161	161	161	161
N_treated	161	161	161	161	161	161	161

This table presents balance test results using 1-year lagged covariates as of the shock year 2017. Sample in Panel B is obtained by truncating the before matching sample (Panel A) to keep observations with logit propensity score between 0.2 and 0.97 (inclusive). Sorting by propensity score, each environmental SI fund is matched by the closest one non-environmental funds.

Sample in Panel D is obtained by truncating the before matching sample (Panel C) to keep observations with logit propensity score between 0 and 0.04 (inclusive). Sorting by propensity score, each SI fund is matched by the closest one conventional fund. Logit propensity score estimation results are shown in Appendix 6.

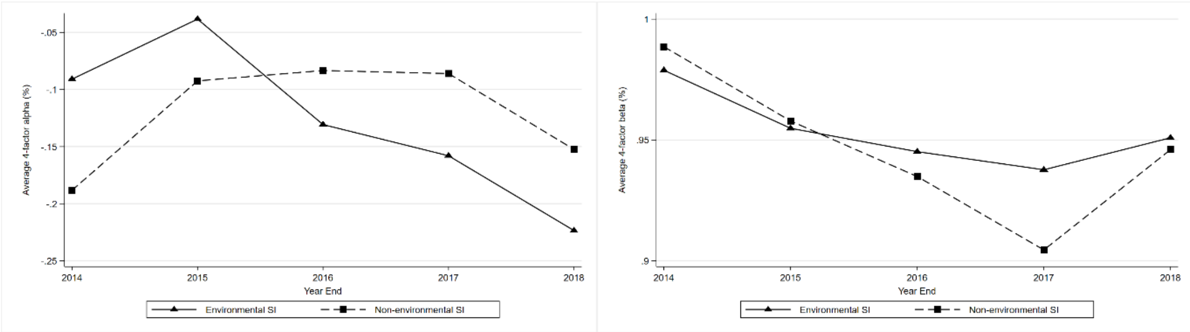


Figure 2. US Paris Agreement withdrawal – Parallel trends on 4-Factor alpha (left) and 4-Factor market beta (right)

Table 3.2016 US presidential election – covariate balancing test

	Lagged $RTN_{i,t}$	$Age_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	-1.64	13.11	-0.11	1.00	0.05	-0.04	-0.01
Treated	-1.23	10.46	-0.03	0.95	0.22	-0.10	0.01
t-stat	-0.56	2.98	-1.73	2.89	-3.52	1.44	-1.11
Median Test							
Control	-0.62	11.00	-0.07	1.00	-0.02	0.07	-0.01
Treated	-1.89	9.00	-0.11	0.97	0.04	-0.01	0.01
z-score	1.43	3.44	0.46	3.58	-3.39	1.01	-0.64
N_control	80	100	81	81	81	81	81
N_treated	146	162	131	131	131	131	131
Panel B. After Matching (Environmental vs Non-Environmental SI funds)							
Mean Test							
Control	-1.44	11.96	-0.15	0.98	0.14	-0.06	0.00
Treated	-0.97	12.19	-0.12	0.97	0.12	-0.03	0.02
t-stat	-0.66	-0.23	-1.05	0.51	0.69	-1.18	-1.59
Median Test							
Control	-0.83	10.00	-0.14	0.99	0.10	0.04	0.01
Treated	-1.09	11.00	-0.12	0.97	0.03	-0.01	0.03
z-score	1.42	-0.42	-0.27	1.62	0.96	-0.64	-0.92
N_control	105	105	105	105	105	105	105
N_treated	105	105	105	105	105	105	105
Panel C. Before Matching (SI vs Conventional Funds)							
Mean Test							
Control	-2.27	11.86	-0.14	0.93	0.17	-0.05	-0.01
Treated	-1.38	11.47	-0.06	0.97	0.15	-0.07	0.00
t-stat	-2.02	0.66	-4.00	-4.09	0.97	1.14	-1.11
Median Test							
Control	-1.78	11.00	-0.12	0.96	0.03	-0.04	0.00
Treated	-0.90	10.00	-0.09	0.99	0.03	0.01	0.01
z-score	-2.50	-0.34	-3.11	-4.41	0.92	0.17	0.60
N_control	8,644	9,875	8,044	8,044	8,044	8,044	8,044
N_treated	226	262	212	212	212	212	212
Panel D. After Matching (SI vs Conventional Funds)							
Mean Test							
Control	-0.89	10.87	-0.05	0.96	0.15	-0.07	0.01
Treated	-0.97	11.95	-0.06	0.97	0.15	-0.08	0.00
t-stat	0.16	-1.28	0.38	-0.63	0.06	0.41	0.27
Median Test							
Control	-1.61	10.00	-0.09	0.98	0.03	-0.03	0.00
Treated	-0.73	10.00	-0.09	0.99	0.03	-0.00	0.01
z-score	-0.95	-2.42	0.58	-0.67	0.40	0.19	0.15
N_control	204	204	204	204	204	204	204

N_treated	204	204	204	204	204	204	204

This table presents balance test results using 1-year lagged covariates as of the shock year 2016. Sample in Panel B is obtained by truncating the before matching sample (Panel A) to keep observations with logit propensity score between 0.3 and 0.9 (inclusive). Sorting by propensity score, each environmental SI fund is matched by the closest one non-environmental funds. Sample in Panel D is obtained by truncating the before matching sample (Panel C) to keep observations with logit propensity score between 0 and 0.1 (inclusive). Sorting by propensity score, each SI fund is matched by the closest one conventional fund. Logit propensity score estimation results are shown in Appendix 6.

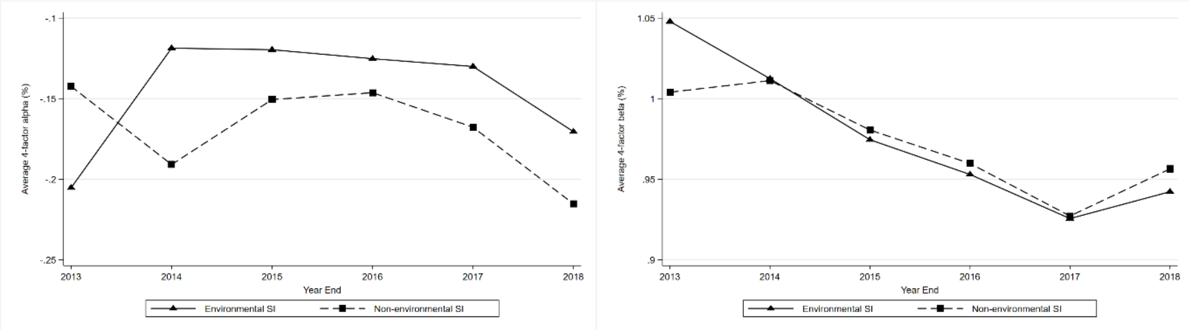


Figure 3. 2016 US presidential election – Parallel trends on 4-Factor alpha (left) and 4-Factor market beta (right)

Table 4. Tax Cut and Jobs Act 2017 – covariate balancing test

	Lagged $\text{RTN}_{i,t}$	$\text{Age}_{i,t}$	Lagged $\alpha_{i,t}$	Lagged $\beta_{mkt,i,t}$	Lagged $\beta_{smb,i,t}$	Lagged $\beta_{hml,i,t}$	Lagged $\beta_{mom,i,t}$
Panel A. Before Matching (SI vs Conventional Funds)							
Mean Test							
Control	11.27	12.14	-0.14	0.93	0.16	-0.05	-0.03
Treated	9.66	10.92	-0.14	0.96	0.13	-0.07	-0.02
t-stat	2.82	2.19	-0.23	-2.43	1.42	1.31	-0.63
Median Test							
Control	9.67	11.00	-0.14	0.97	0.01	-0.06	-0.00
Treated	9.94	10.00	-0.12	0.97	-0.00	-0.01	0.00
z-score	1.62	1.93	-0.79	-1.51	1.53	0.53	0.57
N_control	8,034	10,523	8,522	8,522	8,522	8,522	8,522
N_treated	215	304	224	224	224	224	224
Panel B. After Matching (SI vs Conventional Funds)							
Mean Test							
Control	10.10	10.71	-0.16	0.92	0.09	-0.05	-0.03
Treated	9.98	12.66	-0.17	0.94	0.08	-0.04	-0.01
t-stat	0.15	-2.30	0.43	-1.14	0.34	-0.49	-0.95
Median Test							
Control	8.24	10.00	-0.14	0.94	-0.02	-0.08	-0.02
Treated	9.86	11.00	-0.14	0.96	-0.01	-0.01	0.00
z-score	-0.64	-2.71	-0.20	-1.09	-0.30	-1.06	-1.83
N_control	161	161	161	161	161	161	161
N_treated	161	161	161	161	161	161	161

This table presents balance test results using 1-year lagged covariates as of the shock year 2017. Sample in Panel B is obtained by truncating the before matching sample (Panel A) to keep observations with logit propensity score between 0.3 and 0.04 (inclusive). Sorting by propensity score, each SI fund is matched by the closest one conventional fund. Logit propensity score estimation results are shown in Appendix 6.

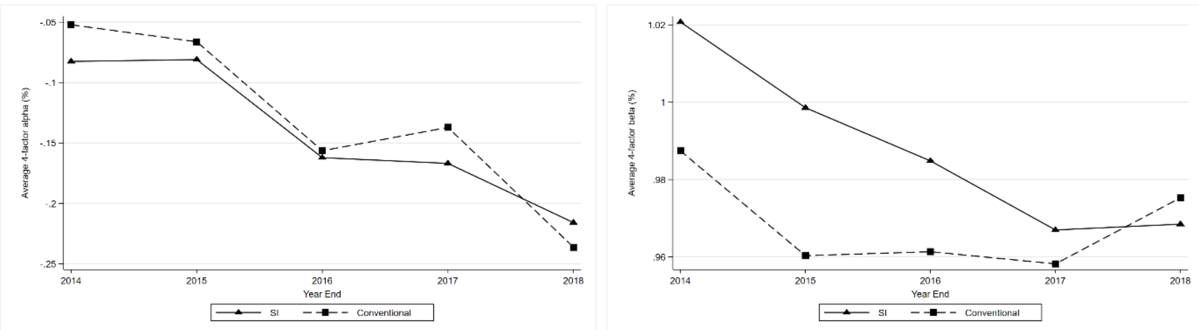


Figure 4. 2016 US presidential election – Parallel trends on 4-Factor alpha (left) and 4-Factor market beta (right)

Appendix 6. Logit Propensity Score Estimation

Table 1. BP oil spill – Logit propensity score estimation

	UK		US	
	Environmental SI funds	SI funds	Environmental SI funds	SI funds
Lagged $RTN_{i,t}$	-0.2016***	-0.0939***	NA	-0.0716***
	(0.07)	(0.02)		(0.02)
$Age_{i,t-1}$	-0.0450	-0.0033		-0.0145
	(0.03)	(0.01)		(0.01)
Lagged $\alpha_{i,t}$	2.4173*	-0.1850		-0.1360
	(1.36)	(0.35)		(0.38)
Lagged $\beta_{mkt,i,t}$	-2.2513	0.7125		-0.9976
	(3.11)	(0.77)		(0.72)
Lagged $\beta_{smb,i,t}$	7.3282***	2.2573***		0.8299***
	(1.70)	(0.44)		(0.30)
Lagged $\beta_{hml,i,t}$	-4.8023***	0.9455**		-2.1496***
	(1.35)	(0.40)		(0.48)
Lagged $\beta_{mom,i,t}$	-8.0079	-2.5013**		-3.6895**
	(4.91)	(1.23)		(1.85)
Intercept	6.0413**	-0.3943		-0.7217
	(2.44)	(0.74)		(0.59)
Log Likelihood	-54.6019	-407.0515		-652.8288
N	127	1534		5770

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is treatment dummy variable: $Treat_i$ = 1 if a fund is an environment-focused SI fund.

Table 2. US Paris Agreement withdrawal– Logit propensity score estimation

	UK		US	
	Environmental SI funds	SI funds	Environmental SI funds	SI funds
Lagged $RTN_{i,t}$	0.0814**	0.0406***	0.0624	-0.1245***
	(0.04)	(0.01)	(0.08)	(0.03)
$Age_{i,t-1}$	-0.0063	0.0096	-0.0292	-0.0092
	(0.02)	(0.01)	(0.02)	(0.01)
Lagged $\alpha_{i,t}$	-3.6642***	-1.0903***	-0.7880	1.8641***
	(0.95)	(0.33)	(1.22)	(0.47)
Lagged $\beta_{mkt,i,t}$	-2.6852**	0.8204	-4.8510***	2.3731***
	(1.11)	(0.56)	(1.71)	(0.60)
Lagged $\beta_{smb,i,t}$	-0.0073	-0.6254***	2.1496**	0.1947
	(0.56)	(0.21)	(0.84)	(0.26)
Lagged $\beta_{hml,i,t}$	-1.6500*	-1.1697***	-1.3856	2.4627***
	(0.92)	(0.35)	(2.01)	(0.65)
Lagged $\beta_{mom,i,t}$	-0.6364	-0.1292	3.2478	-3.2827***
	(1.03)	(0.32)	(2.34)	(0.77)
Intercept	1.7950*	-3.5230***	4.7922***	-4.2013***
	(1.06)	(0.55)	(1.61)	(0.59)
Log Likelihood	-164.6365	-947.4033	-109.7574	-860.4182
N	264	3818	188	7512

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is treatment dummy variable: $Treat_i$ = 1 if a fund is an environment-focused SI fund.

Table 3. 2016 US presidential election - Logit propensity score estimation				
	UK		US	
	Environmental SI funds	SI funds	Environmental SI funds	SI funds
Lagged $RTN_{i,t}$	NA	NA	0.0080	-0.0352
			(0.07)	(0.02)
$Age_{i,t-1}$			-0.0091	-0.0146*
			(0.02)	(0.01)
Lagged $\alpha_{i,t}$			0.2449	1.5356***
			(0.99)	(0.36)
Lagged $\beta_{mkt,i,t}$			-4.8884***	3.1743***
			(1.66)	(0.70)
Lagged $\beta_{smb,i,t}$			2.8878***	-0.7748***
			(0.88)	(0.27)
Lagged $\beta_{hml,i,t}$			-0.9158	-0.3593
			(1.02)	(0.30)
Lagged $\beta_{mom,i,t}$			2.8742	0.3599
			(2.44)	(0.75)
Intercept			5.0763***	-6.2944***
			(1.63)	(0.68)
Log Likelihood			-120.6742	-932.1397
N			205	8094

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is treatment dummy variable: $Treat_i$ = 1 if a fund is an environment-focused SI fund.

Table 3. Oxfam scandal 2011 and Tax Cut and Jobs Act 2017 – Logit propensity score estimation				
	UK – Oxfam scandal		US – Tax Cut and Jobs Act	
	Environmental SI funds	SI funds	Environmental SI funds	SI funds
Lagged $RTN_{i,t}$	NA	0.0643***	NA	-0.1245***
		(0.02)		(0.03)
$Age_{i,t-1}$		-0.0052		-0.0092
		(0.01)		(0.01)
Lagged $\alpha_{i,t}$		-2.8616***		1.8641***
		(0.41)		(0.47)
Lagged $\beta_{mkt,i,t}$		-0.6095		2.3731***
		(0.66)		(0.60)
Lagged $\beta_{smb,i,t}$		-0.4285		0.1947
		(0.42)		(0.26)
Lagged $\beta_{hml,i,t}$		1.2462**		2.4627***
		(0.49)		(0.65)
Lagged $\beta_{mom,i,t}$		-1.4185		-3.2827***
		(0.93)		(0.77)
Intercept		-2.5175***		-4.2013***
		(0.67)		(0.59)
Log Likelihood		-460.3638		-860.4182
N		1753		7512

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is treatment dummy variable: $Treat_i$ = 1 if a fund is an SI fund.

Appendix 7. Tax Cut and Jobs Act December 2017 - Impact on US donors

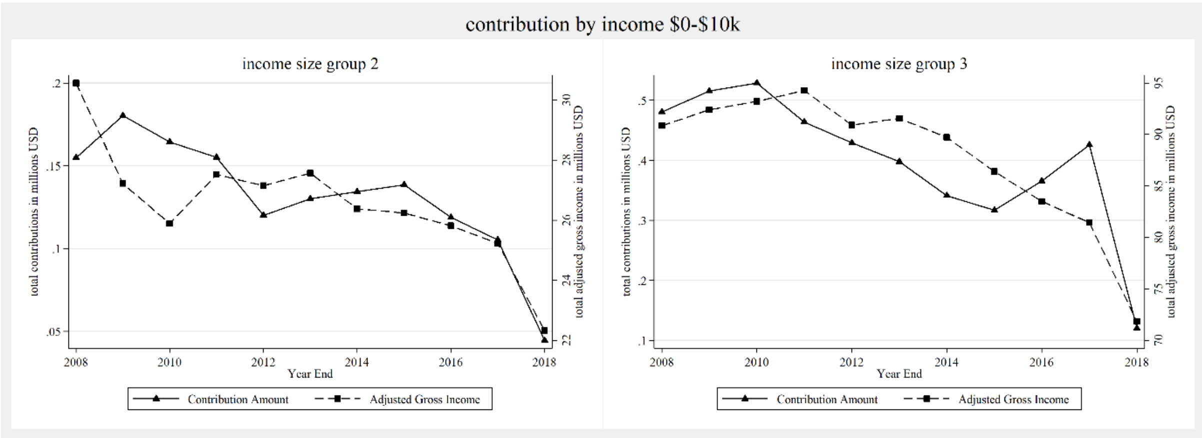


Figure 1. US Charitable contribution by income group ($0-$5k and $5-10k)

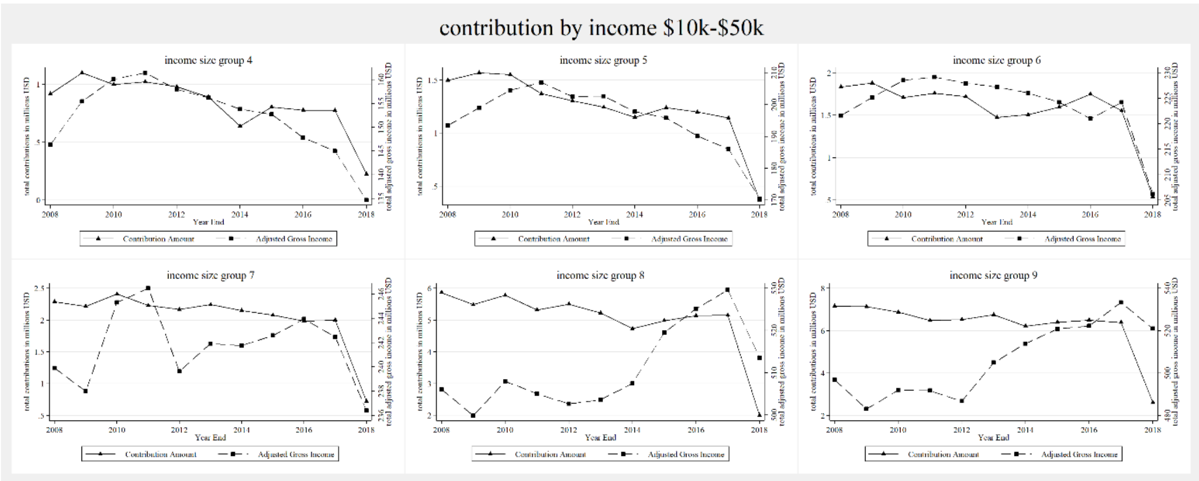


Figure 2. US Charitable contribution by income group ($10k-$50k)

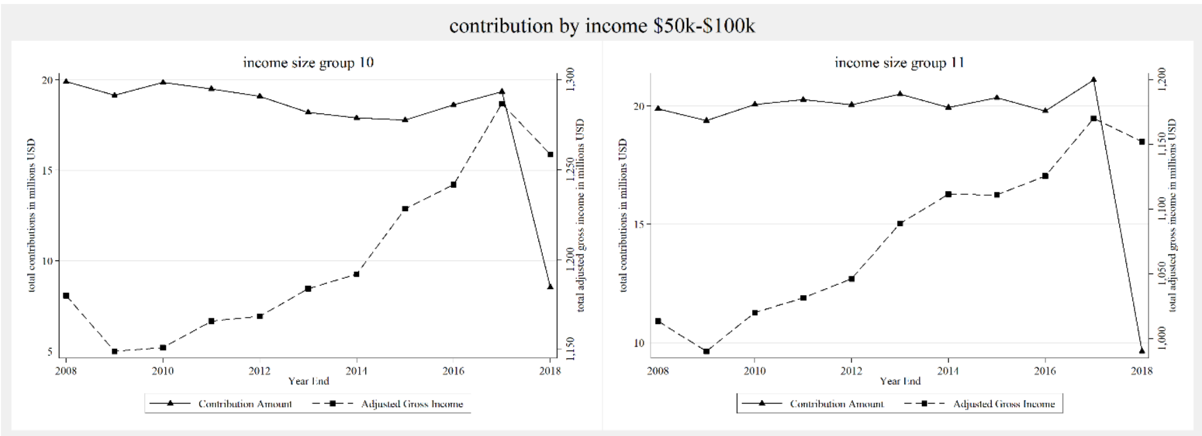


Figure 3. US Charitable contribution by income group ($50k-$75k and $75k-$100k)

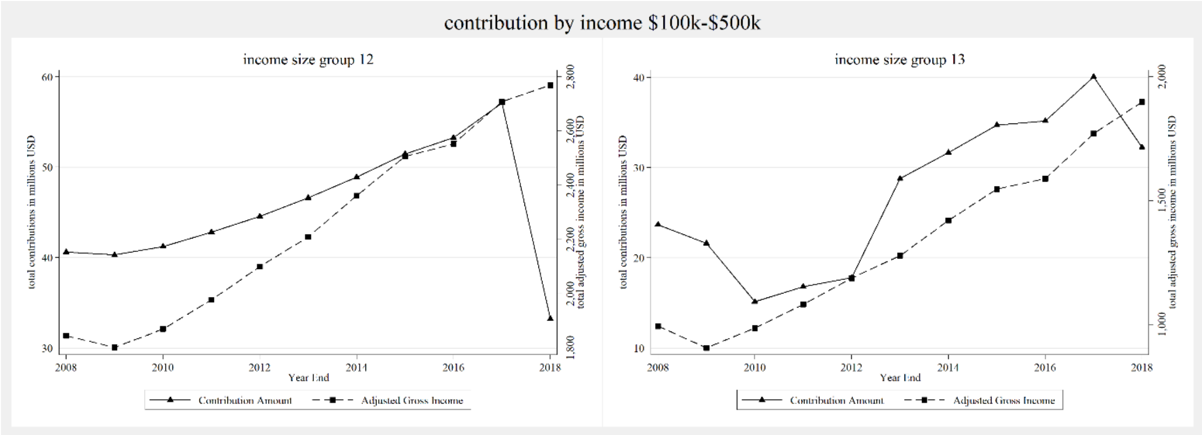


Figure 4. US Charitable contribution by income group ($100k-$200k and $200k-$500k)

Appendix 8. UK Charity Commissions Classification Code

classno	classtext
101	General Charitable Purposes
102	Education/training
103	The Advancement Of Health Or Saving Of Lives
104	Disability
105	The Prevention Or Relief Of Poverty
106	Overseas Aid/famine Relief
107	Accommodation/housing
108	Religious Activities
109	Arts/culture/heritage/science
110	Amateur Sport
111	Animals
112	Environment/conservation/heritage
113	Economic/community Development/employment
114	Armed Forces/emergency Service Efficiency
115	Human Rights/religious Or Racial Harmony/equality Or Diversity
116	Recreation
117	Other Charitable Purposes
201	Children/young People
202	Elderly/old People
203	People With Disabilities
204	People Of A Particular Ethnic Or Racial Origin
205	Other Charities Or Voluntary Bodies
206	Other Defined Groups
207	The General Public/mankind
301	Makes Grants To Individuals
302	Makes Grants To Organisations
303	Provides Other Finance
304	Provides Human Resources
305	Provides Buildings/facilities/open Space
306	Provides Services
307	Provides Advocacy/advice/information
308	Sponsors Or Undertakes Research
309	Acts As An Umbrella Or Resource Body
310	Other Charitable Activities

Source: Charity Commissions at https://register-of-charities.charitycommission.gov.uk/register/full-register-download

Appendix 9. Event Study - CAR plot

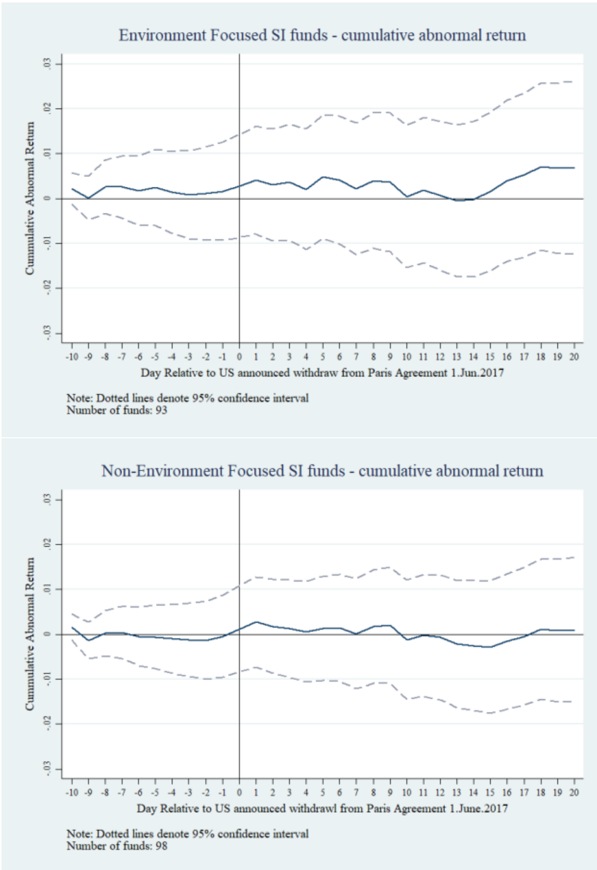
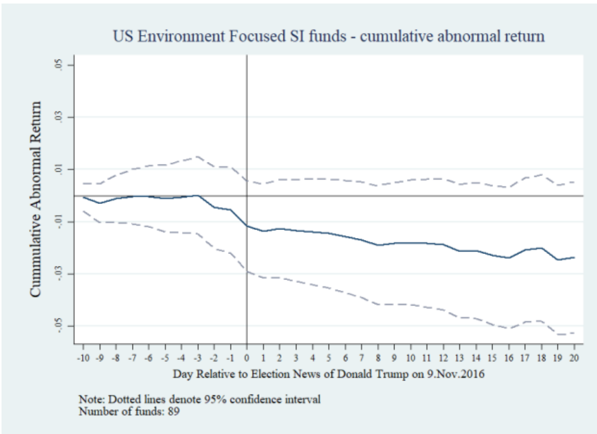


Figure 1. CAR plot on US Paris Agreement withdrawal– US environment(left) and non-environment(right) focused SI funds



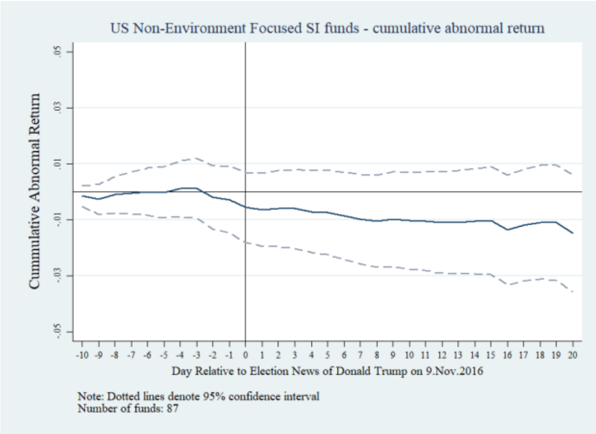


Figure 2. CAR plot on 2016 US presidential election – US environment(left) and non-environment(right) focused SI funds

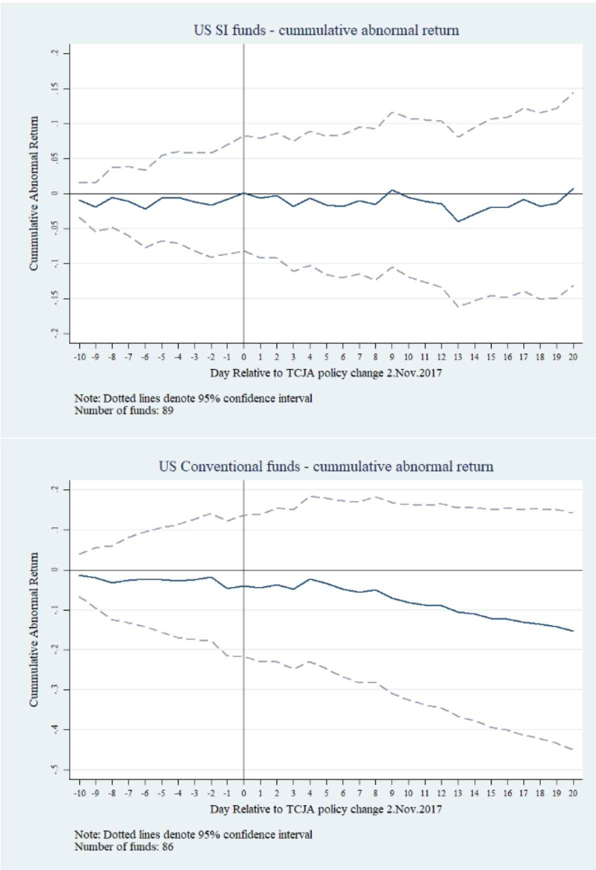


Figure 3. CAR plot on Tax Cut and Jobs Act – US SI (left) and Conventional (right) funds

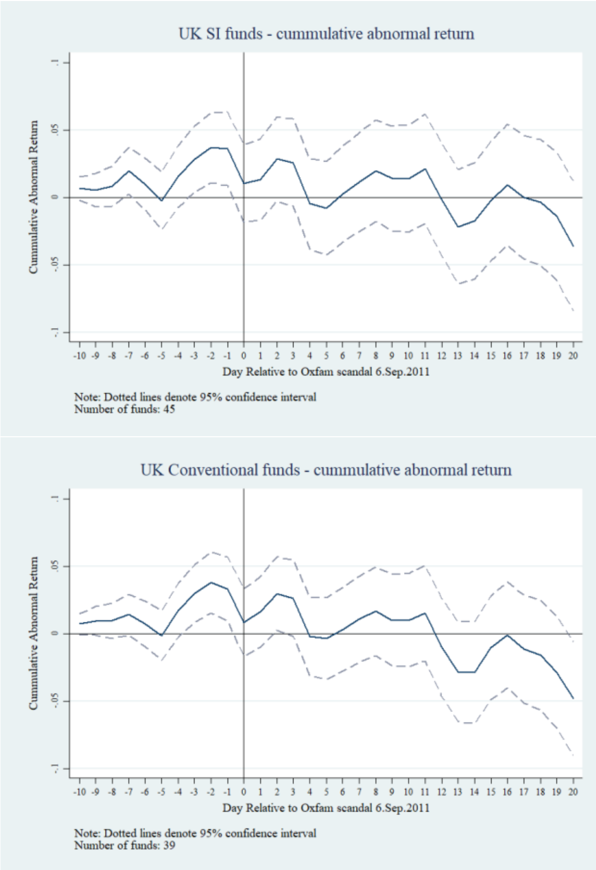


Figure 4. CAR plot on Oxfam scandal – UK SI (left) and Conventional (right) fund

Appendix 11. Variable and Concept Definitions

Variable Definitions	
Notation	*Definition*
Investor	A respondent to the Survey of Consumer Finance (SCF) is defined as an investor if he/she (and/or the respondent's family) holds a non-0 amount in stocks as part of any of the following accounts: stock mutual funds and ETFs (X3822, X6704), combination funds and ETFs (X3830), publicly traded stocks (X3915), or through the following means: Roth IRA (X6551, X6559, X6567), roll-over IRA (X6553, X6560, X6568), regular or other IRA (X6553, X6561, X6569), Keogh account (X6554, X6562, X6570), savings or money market accounts (X3730, X3736, X3742, X3748, X3754, X3760), annuities (X6577); managed investment accounts (X6587); and pension accounts (X11032, X11132, X11332, X11432). In addition, we require any holdings through pension plans to provide a choice about how the pension plan is invested.
Donor	A respondent to the Survey of Consumer Finance (SCF) is defined as a donor if he/she (or the respondent's family) made a charitable contribution of at least $500 in the survey year (X5822) or had a charitable trust or foundation (X7660).
$Inc\_vol_{i,t}$ $Total\_contributions_{i,t}$	As part of total income, the voluntary income received by UK non-profit organization *i* in year *t* (in thousands of GBP) (excludes legacies) As part of total revenue, the total contributions received by US non-profit organization *i* in year *t* (in thousands of USD) from public contributions (i.e., donations).
$Total\_income_{i,t}$ $Total\_revenue_{i,t}$	Total income of UK non-profit organization *i* in year *t* (in thousands of GBP), including donations, legacies, government grants, and income from other trading and charitable activities (e.g. providing services) Total revenue of US non-profit organization *i* in year *t* (in thousands of USD), including contributions as defined above, investment income, sale of assets, and income from other activities (e.g. rental income)
Total functional $expense_{i,t}$	To be deducted from total revenue of US non-profit organization *i* in year *t* (in thousands of USD), the amount of fundraising expenses
Total government $grants_{i,t}$	Gifts, grants, membership fees received by US non-profit organization *i* in year *t* (in thousands of USD) as a proxy for government grants.
$Volunteers_{i,t}$	Number of volunteers of UK non-profit organization *i* in year *t*
$Employee_{i,t}$	Number of employees of UK non-profit organization *i* in year *t*
Fixed $assets_{i,t}$	Total fixed assets of UK non-profit organization *i* in year *t* (in thousands of GBP)
$Total\_assets_{i,t}$	Total assets of US non-profit organization *i* in year *t* (in thousands of GBP or USD)
Total government $grants_{i,t}$	Total government grants received by US non-profit organization *i* in year *t* (in thousands of USD)
Fundraising $expense_{i,t}$	Total fundraising expenses of UK or US non-profit organization *i* in year *t* (in thousands of GBP or USD)
$Post_t$	A dummy variable for the post-shock year. Shock year differs depending on the specification of the diff-in-diff shock; full description available under each regression table note.
$Treated_i$	For our charity flow analysis: a dummy variable for a charity *i,* defined based on shocks examined; full description available under table notes of each regressions For our SI and conventional fund flow analysis: a dummy variable equal to 1 for UK/US open-ended equity fund *i* with a socially conscious mandate as defined by Morningstar, and 0 if fund *i* does not have a socially conscious mandate as defined by Morningstar and is thus a 'conventional' fund.

$Post_t \times Treated_i$	Interaction term with dummy variables $Post_t$ and $Treated_i$
$Flow_{i,t}$	Calendar-year estimated share-class level net flow toward fund *i* (in millions of USD) at year *t* obtained from Morningstar
Lagged $RTN_{i,t}$	1-year holding period return for fund *i* in year *t-1*
$RTN_{i,t}$	1-year holding period return (in %) for fund *i* in year *t*
$Age_{i,t}$	Fund age in the nearest round year measured on 31st December of the pre-shock year *t-1* depending on different diff-in-diff shock.
Lagged $expratio_{i,t}$	Lagged net expense ratio (in %) for fund *i* at year *t-1*, defined by Morningstar as: The percentage of fund assets used to pay for operating expenses and management fees, including 12b-1 fees, administrative fees and all other asset-based costs incurred by the fund, except brokerage costs. Fund expenses are reflected in the fund's NAV. Sales charges are not included in the expense ratio.
$\alpha_{i,t}$	Fama-French four-factor alpha of fund *i* at year *t*, estimated from running Fama-French four-factor model using fund *i*'s past 24 months of excess return and return on four-factor portfolios obtained from Ken-French online data library.
$\beta_{mkt,i,t}$	Market beta of fund *i* at year *t*, estimated from running Fama-French four-factor model using fund *i*'s past 24 months of excess return and return on four-factor portfolios obtained from Ken-French online data library.
$\beta_{smb,i,t}$	Size factor loading of fund *i* at year *t*, estimated from running Fama-French four-factor model using fund *i*'s past 24 months of excess return and return on four-factor portfolios obtained from Ken-French online data library.
$\beta_{hml,i,t}$	Value factor loading of fund *i* at year *t*, estimated from running Fama-French four-factor model using fund *i*'s past 24 months of excess return and return on four-factor portfolios obtained from Ken-French online data library.
$\beta_{mom,i,t}$	Momentum factor loading of fund *i* at year *t*, estimated from running Fama-French four-factor model using fund *i*'s past 24 months of excess return and return on four-factor portfolios obtained from Ken-French online data library.
$Institutional_i$	Defined by Morningstar as an indication that the share class is primarily aimed at institutional investors. $Institutional_i$ = 1 if the Morningstar entry to this field for fund *i* is "Yes".

Appendix 12. UK 2018 Charity Summary Statistics with updated dataset by Charity Commissions UK

Table 2.2 Charity summary statistics

Panel A. UK charity sample (in thousands of £) – mean by year	
	2018
Voluntary income	1008.6
Total income	3641.8
Charitable activity expenditure	2840.7
Total assets	12482.1
Number of volunteers	0
Number of employees	0
Number of charities	603