ecgi

Does Paying Passive Managers to Engage Improve ESG Performance?

Marco Becht
Solvay Brussels School of Economics and
Management, Université libre de Bruxelles, CEPR
and ECGI
and ECGI

Julian R. Franks
London Business School, CEPR and ECGI

Hideaki Miyajima
Waseda University and RIETI

Kazunori Suzuki
Waseda University and ECGI

This paper can be downloaded without charge from:
http://ssrn.com/abstract_id=4506415

www.ecgi.global/content/working-papers

european corporate governance institute

Does Paying Passive Managers to Engage Improve ESG Performance?

Marco Becht
Julian R. Franks
Hideaki Miyajima
Kazunori Suzuki

Abstract

The paper studies a natural experiment in responsible investment conducted by the Japanese Government Pension Investment Fund (GPIF). In 2018 GPIF gave its largest passive manager a remunerated mandate to engage with portfolio companies to improve environmental, social and governance performance. The fund adopted best-in-class indexes, rewarding companies with high ESG scores with additional equity investment. Using private data and difference-in-differences analysis we show that engagement by the asset manager has improved scores. In an event study, we find that the conditional portfolio tilt significantly impacts share prices. We also provide evidence that ESG scores for the Japanese stock market increased significantly more during the treatment period than most those of companies in other countries.

Marco Becht*
Goldschmidt Professor of Corporate Governance
Université Libre de Bruxelles
42 Avenue F. D. Roosevelt
1050 Brussels, Belgium
phone: +32 2650 4466
e-mail: marco.becht@ulb.ac.be

Julian R. Franks
Professor of Finance and Alexander M Knaster Chair
London Business School
Sussex Place, Regent's Park
London NW1 4SA, United Kingdom
phone: +44 20 7262 5050 x3449
e-mail: jfranks@london.edu

Hideaki Miyajima
Professor
Waseda University
1-6-1, Nishi-Waseda
Shinjuku-Ku, Tokyo 169-8050, Japan, Japan
phone: + 81 3 5286 2019
e-mail: miyajima@waseda.jp

Kazunori Suzuki
Professor of Finance
Waseda University, Waseda Business School
1-6-1, Nishi-Waseda
Shinjuku-Ku, Tokyo 169-8050, Japan, Japan
phone: + 81 3 5286 8411
e-mail: ksuzuki@waseda.jp

Does Paying Passive Managers to Engage Improve ESG Performance?

Marco Becht
Solvay Brussels School, Université libre de Bruxelles
CEPR and ECGI

Julian Franks
London Business School
ECGI

Hideaki Miyajima
Faculty of Commerce, Waseda University
RIETI

Kazunori Suzuki
Graduate School of Business and Finance, Waseda University
ECGI

13 September 2023

Does Paying Passive Managers to Engage Improve ESG Performance?

Abstract

The paper studies a natural experiment in responsible investment conducted by the Japanese Government Pension Investment Fund (GPIF). In 2018 GPIF gave its largest passive manager a remunerated mandate to engage with portfolio companies to improve environmental, social and governance performance. The fund adopted best-in-class indexes, rewarding companies with high ESG scores with additional equity investment. Using private data and difference-in-differences analysis we show that engagement by the asset manager has improved scores. In an event study, we find that the conditional portfolio tilt significantly impacts share prices. We also provide evidence that ESG scores for the Japanese stock market increased significantly more during the treatment period than most those of companies in other countries.

2

1. Introduction

There are over 5,000 investors and service providers that have signed the Principles of Responsible Investment (PRI). The main strategies employed are divestment, portfolio tilting, engagement, and voting. Engagement is considered essential in equity investment, but direct evidence on its impact is scarce because it is usually conducted in private.[1] In addition, the commitments to engage from passive managers are often viewed as lacking in substance. The main issue revolves around cost and incentives: passive managers hold many positions and compete by offering lower fees. In contrast, engaging companies through private contacts and meetings is expensive and ESG requires diverse expertise, for example on energy, biodiversity, human rights, and supply chain management. The practice of bundling engagement and portfolio management gives rise to a potential conflict of interest between beneficial owners and asset managers; for example, passive managers are perceived to have an incentive to minimise stewardship efforts while beneficial owners would like them to monitor and engage.[2] This hypothesis of active ownership passivity is supported by evidence on a lack of outcomes; investment from US PRI signatories has not improved the ESG scores of target companies (Krueger et al., 2020).[3]

Divestment and portfolio tilting are also viewed sceptically. The central argument is that the required portfolio shift is too large to depress the prices of equities of companies with low ESG scores, and as a result prevents any effective signal to management of investor dissatisfaction. Moreover, investors that place little value on ESG, or have a higher ESG risk tolerance, will purchase the shares on any price decline, and mitigate or even eliminate any price effect. The result is that the impact on company behaviour will be minimal (Broccardo et al., 2022; Heinkel et al., 2001), an effect accentuated with traditional remuneration contracts (Davies and Van Wesep, 2018). Calibration on U.S. data suggests that to increase the cost of capital by at least

---

[1] Votes against management, that are publicly observable, are a misleading measure of engagement. When engagement is successful no negative voting is necessary (Becht et al., 2009).

[2] We deliberately distinguish between engagement and voting. It is relatively cheap and easy to comply with the PRI commitment to exercise voting rights by employing a voting manager or by outsourcing to a proxy adviser; engagement that involves time consuming research, contacts, meetings, and specialized staff is a separate and more cost intensive activity. Engagement can be outsourced to equity ownership services, but these operate rather differently from proxy advisors (Becht et al., 2021). In the divestment literature, both empirical and theoretical, "engagement" is often treated as synonymous with voting and can be a source of confusion; see below.

[3] There is some evidence that an equity ownership service achieved engagement outcomes when measured in "milestone progress", a measure devised by the asset manager for self-assessment and client reporting that pre-dates ESG scores (Dimson, Karakaş, and Li 2015). There is also some evidence that asset managers can have an impact on ESG outcomes (Azar et al., 2021; Lin et al., 2023).

1% would require at least 84% of investors to hold "clean stocks" only, in this case by investing in the FTSE 4 Good Select USA Index (Berk and van Binsbergen, 2021).

Empirical evidence on demand shocks from inclusion or exclusion in best-in-class indexes like the FTSE4Good is consistent with the theoretical prediction; returns are generally insignificant (Berk and van Binsbergen, 2021; Curran and Moran, 2007; Hawn et al., 2018).[4] The consensus of this literature is that divestment and portfolio tilting is inferior to engagement, despite the lack of evidence on the efficacy of engagement. The evidence on index inclusion is also far from conclusive because the amounts invested in "best-in-class" indexes are usually quite small.

The objective of this paper is to provide new evidence on ESG engagement and index tilting in the context of a natural experiment conducted by the Government Pension Investment Fund (GPIF) of Japan, the largest pension fund in the world. GPIF developed two related strategies to encourage improvement in ESG. First, in 2018, GPIF gave its largest passive manager (AM One) and an active manager a mandate to "improv[e] the overall market through stewardship activities" by setting medium- to long-term goal for engagement activities, engage and achieve ESG milestone progress.[5] The asset managers are remunerated separately for engagement. Second, in 2017 GPIF adopted two best-in-class indexes from FTSE and MSCI that reward improvements in ESG performance with index inclusion resulting in additional equity investment through portfolio tilting.[6] Index inclusion and exclusion is determined by ESG scores produced by the index providers. While, the engagement mandate required the asset manager to improve the ESG performance of TOPIX constituents in terms of milestones, it was not linked explicitly to the best-in-class indices.[7]

We use private data from the passive asset manager (AM One) to evaluate the impact of the engagement programme using difference-in-differences (DiD), a technique that is widely used in impact evaluation, because it allows us to draw causal inferences (Gertler et al., 2016). We

---

[4] The number of related studies is large and most arrive at similar conclusions (Yilmaz et al., 2020). A recent theory paper argues that tilting is superior to divestment because it provides a financial incentive to take corrective action (Edmans et al., 2022).

[5] GPIF Stewardship Activities Report 2020-21, pg. 30.

[6] GPIF also adopted a thematic "best-in-class" index promoting gender diversity, the MSCI Japan Empowering Women Index (WIN). WIN "aims to include companies that lead their respective [..] sector groups in terms of supporting women's participation and advancement in the workforce and adopting diversity policies" (MSCI WIN Factsheet, 31 January 2023). The index is organized as a tournament; it ranks the companies in each sector using MSCI diversity scores and includes the top 50% from each sector in the index. The WIN index has been studied separately (Mehrotra et al., 2023).

[7] In September 2018 GPIF adopted additional climate related thematic indices, the S&P/JPX Carbon Efficient Index for domestic equities and the S&P Global Ex-Japan Large Mid Carbon Efficient Index for foreign equities. These target indexes use portfolio tilting by overweighting and underweighting the constituents of the underlying index (GPIF, 2018). Best-in-class carbon leader indexes were available at the time (Andersson et al., 2016), but not adopted by GPIF. It is much harder to measure the incentive effects created by target indexes compared to leader indexes.

4

compare the difference in ESG scores between companies that were engaged (the treatment group) with companies that were never engaged by GPIF's asset manager (the control group) before and after the beginning of the engagement programme. If the treatment was successful, we would expect to see a change in the wedge between these two differences. The technique eliminates two potential biases in simple "before and after" impact studies; it controls for fixed effects that might influence the responsiveness to treatment and, for potential times series changes that are unrelated to the treatment but might nevertheless improve the outcome, leading the observer to falsely attribute the improvement to the treatment (Angrist and Pischke, 2009).

Any change in scores potentially will lead to inclusion or exclusion in best-in-class indexes and changes in demand for the company's stock. As a result, we investigate the price reaction of demand shocks caused by ESG index inclusion and exclusion. This allows us to assess if the adoption of best-in-class incentives has provided a financial benefit that would make engagement more likely to persuade management to agree to the demands of the engagement. Note that index inclusion provides an incentive to improve scores for all companies with scores below the inclusion threshold. All companies with low scores should have an incentive to increase their scores, but low score companies that were engaged should have a larger increase in scores than low score companies that were not engaged.

Our definition of the control group reflects the fact that some companies were engaged earlier than others. The practice of neglecting staggered treatment has been very common (de Chaisemartin and D'Haultfœuille, 2022) and involves a misclassification; units that were treated late are assumed to be treated with the first treatment cohort. In many studies this has resulted in biased results (Baker et al., 2022). There are new estimators specifically devised for staggered treatment (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021) and we have applied Callaway and Sant'Anna as a robustness check.[8] The estimates from the dynamic event study confirm that there is a significant treatment effect. In event time, the treatment effect is more pronounced after two or three years. This aligns with the idea that the impact of engagement is incremental rather than instantaneous.

The remunerated engagement programme of the passive manager (AM One) was started in 2018. The program agreed with GPIF is based on 20 ESG engagement themes that were classified into E, S, G and a fourth category "ESG", where the latter is some combination of

---

[8] There are several new approaches to tackling multiple periods and variation in treatment timing with implementations in R and Stata (Roth et al., 2023). These are collected in a special repository by Asjad Naqvi at https://asjadnaqvi.github.io/DiD/.
.

the three separate categories. Progress is self-assessed by the asset manager and measured through an eight-stage milestone system. These stages range from identifying ESG issues (1), to issues being recognised by senior management (4) to plans being implemented (7) and completing the engagement (8). The asset manager is assessed on milestone progress.

One year prior to the remunerated engagement programme, GPIF adopted the "best-in-class" ESG indexes for Japanese domestic equities provided by FTSE and MSCI: the FTSE Blossom Index and the MSCI ESG Leaders Index Japan.[9] These indexes require companies to have an ESG score above a certain threshold for inclusion.[10] GPIF conditionally tilted its portfolio to companies with relatively high ESG scores, as measured by FTSE and MSCI. The additional amounts invested are significant and increased from approximately 1,2 trillion yen in 2017 to 3,9 trillion Yen by the end of fiscal 2021 (approximate investment of 32 billion USD).[11]

There is no global consensus on the definition of good ESG performance, even when taking a risk-based approach. In addition, there is some confusion between the risk management and ethics-based approaches to ESG. As a result, ESG scores from major providers have adopted different methodologies and exhibit a low degree of correlation (Berg et al., 2022). We use ESG metrics from FTSE and MSCI as our measure of ESG performance because these scores were chosen by GPIF when it adopted the FTSE Blossom and MSCI Leaders indexes.[12]

There are some differences between the FTSE and MSCI scores and index methodologies that complicate the analysis. FTSE scores are based on publicly available data and they capture the degree of corporate disclosures. Non-disclosure of a required indicator is penalised with a zero score (Ratsimiveh and Haalebos, 2021); pillar, theme and indicator scores are weighted with a four level "exposure" score (high, medium, low and negligible, or not applicable). In contrast, MSCI scores rely on a mixture of public disclosure with company research. ESG risk is measured through materiality weights that adjust dynamically and capture industry and company specific factors.

---

[9] The FTSE Blossom Sector Relative Index was added in 2022. The MSCI ESG Leaders Index Japan was new; the Blossom was based on the established FTSE4Good Select Japan index. However, unlike the FTSE4Good, the Blossom index does not have an ethics filter; tobacco and fossil fuel companies are eligible for inclusion provided they exceed the required ESG scoring threshold. This modification was necessary because GPIF does not have an ethics or principle-based exclusion list. In this respect GPIF's purely risk-based approach is similar to the average PRI signatory from the United States (Krueger et al., 2020) and is different from the university endowments and faith-based organisations that spearheaded the fossil free divestment movement (Becht et al., 2023).

[10] Index inclusion also requires the absence of "controversies".

[11] In addition, GPIF had invested about 1,246 billion Yen in the thematic MSCI-WIN index at the end of fiscal 2021 (approximately 10 billion USD), a substantial increase from the initial amount in 2017 (388.4 billion Yen).

[12] Inclusion in the MSCI and FTSE ESG indexes also requires the absence of controversies; we will explore the resolution of controversies through engagement in future research.

The index rules also differ. For example, FTSE index inclusions and exclusions are announced publicly, and many companies repeat the announcement through press releases and via social media. In contrast, modifications to the MSCI index are communicated privately to index subscribers, not to the public or the relevant companies. Companies can obtain inclusion and exclusion information from their broker and MSCI will provide an MSCI Leaders Index member with a logo, but only upon request. This might suggest that inclusion in the FTSE Blossom involves a demand shock and a publicity effect; inclusion in the MSCI index may produce primarily a demand shock.[13]

To assess the impact of engagement on ESG performance, we sorted the portfolio into two groups: companies that were engaged at least once and companies that were never engaged by the asset manager. The ESG scores for the FTSE evolved in parallel before AM One started its engagement programme.[14] The overall FTSE scores for the group of companies that were engaged increased significantly relative to companies that were not engaged. In absolute terms E engagements had the largest impact on FTSE E pillar scores, and the treatment effect was most pronounced for the companies with the lowest scores. Overall, there was little or no significant impact on MSCI scores from the engagements.

To explore the question of why AM One's pillar engagements are reflected so inconsistently in MSCI and FTSE pillar scores we undertook two case studies that map the engagement agendas to the theme scores that are then aggregated into pillar scores. We found that in some cases AM One engaged on issues that were considered as relatively unimportant by MSCI. Therefore, progress on these topics is not reflected in improved MSCI scores. We also found that FTSE E scores improve more strongly from better disclosure than MSCI. This could explain the difference in the impact on E scores. We also found that MSCI materiality weights were inversely correlated with milestone progress and higher theme scores. As a result, successful engagement was not reflected in higher pillar scores. In other words, if an engagement was successful on a particular theme, the weight for that theme might be reduced and as a result mitigate any performance improvement. The case studies confirm our conjecture that the difference in results between FTSE and MSCI at the pillar level are likely to be influenced by the scoring methodologies.

We also assess if GPIF's adoption of the FTSE Blossom and the MSCI ESG Leaders Index provides Japanese companies with financial incentives to improve their ESG scores. We use

---

[13] In addition, many FTSE constituents issue a press release when FTSE confirms their continued inclusion.
[14] MSCI for companies that were never engaged showed a slightly divergent trend.

event study methodology to analyse the extent to which inclusion in the Blossom index and the MSCI Leader index generated positive returns. We also analyse exclusions to determine if they generated negative returns. We show abnormal returns to index inclusion and exclusions of around 2% (positive for inclusions and negative for exclusions), although the results for exclusions are less robust. For the Blossom index the effect is most pronounced for inclusions from the TOPIX Small index; for MSCI it is most pronounced from the TOPIX Midcap 400.

Finally, we provide an additional test that avoids some of the complications of our definition of the treated group and the control group within the Japanese market. In this additional test we treat all Japanese companies as the treated group and a set of foreign companies from 26 stock markets as the control. We find that FTSE scores for TOPIX 500 companies improves significantly more than the scores other stock markets, including Germany, France, the United Kingdom and the United States.

The remainder of the paper is structured in six parts. Section 2 provides additional information about GPIF and a short history of its responsible investment programme, with emphasis on passive equity investment. Section 3 sets out the engagement programme put in place by AM One and reports descriptive statistics. Section 4 investigates the impact of engagements on ESG performance using difference-in-differences analysis. It shows that engagement had a measurable impact on FTSE scores, but hardly on MSCI. Section 5 provides a case study to explore why the AM One engagement programme had a different impact on FTSE and MSCI scores. Section 6 explores potential financial incentives from "best-in-class" index inclusion. Section **Error! Reference source not found.** evaluates the joint impact of GPIF's programme. Section 8 concludes. Each sub-section contains a description of the relevant data and methodology.

2. GPIF's Responsible Investment Programme

GPIF is the largest pension fund in the world with an asset value of about USD 1,7 trillion at the end of 2020.[15] It provides supplementary finance for the Japanese public employee pensions system by aiming "to secure a long-term real return […] of 1.7% with minimal risks".[16] GPIF

---

[15] https://www.pionline.com/interactive/worlds-largest-retirement-funds-2021 (accessed 9 February 2022). Norway's Government Pension Fund was a close second with $1,305,920 millions. GPIF was more than three times the size of the largest U.S. pension fund, California Public Employees (CalPERS) with assets valued at $426,247 millions on 20 September 2020.

[16] GPIF Annual Report 2020, pg. 18.

has been instrumental in promoting responsible investment and ESG (environmental, social and governance) in Japan.[17]

GPIF has traditionally sought to achieve this goal through an asset allocation that was heavily skewed to Japanese government bonds. In 2013 the late Prime Minister Abe's cabinet office (*Kantei*) published a Revitalization Strategy for Japan ("Japan is Back")[18] that prompted the review of GPIF's organization and investment policies by an expert committee.[19] Following recommendations from the committee,[20] GPIF's board changed the fund's asset mix and today GPIF has invested half of its portfolio outside of Japan and holds half in publicly listed equities. It is not a coincidence that the fund put its weight behind the development of the Japanese stock market. For example, in 2014 GPIF endorsed the creation of the JPX400 index that was designed to showcase the 400 largest, most liquid, and most profitable Japanese companies. There is evidence that the prestige of inclusion caused companies to improve their margins (Chattopadhyay et al., 2020).

Also in 2014 GPIF started to take a more direct approach to responsible investment.[21] The fund adopted Japan's Stewardship Code and adopted measures that supplemented its monitoring of external managers (Exhibit 1). In 2015 GPIF became a signatory to The Principles of Responsible Investment (PRI) promoting its commitment to investor stewardship.[22] In 2016 it created a business and asset owners' forum and initiated a survey of listed companies. The survey provides GPIF with direct feedback from portfolio companies about their external asset managers' ESG-related dialogue with portfolio companies, voting and other issues.[23]

The law that created GPIF does not allow the fund to engage directly with portfolio companies. Hence the fund exercises its investment and stewardship mandate through external asset managers. In 2020 over 80 percent of its portfolio was managed by passive managers that receive a fixed fee of less than 1 basis points. [24] Active managers receive a base fee equivalent to what passive managers receive plus a performance fee linked to target excess returns.[25] The

---

[17] GPIF's evolution from a traditional investor to an ESG leader is well documented in a recent Harvard Business School case (Henderson et al., 2019). GPIF's former Chief Investment Officer and Executive Director, Hiromichi Mizuno, played a crucial role. He joined GPIF in 2015.

[18] https://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/saikou_jpn.pdf

[19] https://www.mhlw.go.jp/stf/shingi/0000013751.html

[20] Published in November 2013 (https://www.cas.go.jp/jp/seisaku/koutekisikin_unyourisk/houkoku/h251120.pdf).

[21] GPIF has issued annual Stewardship and ESG reports since 2017.

[22] PRI classifies responsible investment into portfolio integration where signatories consider ESG issues when building a portfolio and active ownership or stewardship, where investors seek to improve the ESG performance of portfolio companies, for example through engagement and voting (https://www.unpri.org/an-introduction-to-responsible-investment/what-is-responsible-investment/4780.article; accessed 27 February 2022).

[23] GPIF ESG Report 2017, pg. 13.

[24] GPIF Annual Report 2020, pg. 77.

[25] GPIF Annual Report 2020, pg. 77.

average fee paid to external managers in 2020, both active and passive, was 4 basis points.[26] GPIF asks external managers to "establish a detailed proxy voting policy (guideline) and to report the voting results to the GPIF".[27]

In 2017 GPIF reallocated some of its domestic passive portfolio to companies with a high ESG rating by adopting a group of "best-in-class" ESG indices. Since all these companies were already included in the index portfolios invested in by their passive managers, investment in the ESG indexes would increase both the dollar investment and the weighting of these companies in the GPIF portfolio. An important characteristic of the best-in-class indexes is that they do not screen out companies or sectors based on their activities or sectors, such as tobacco or coal.[28] All companies in the TOPIX All Share Index are eligible for inclusion in these ESG indices providing they exceed certain ESG metrics. GPIF expected "that the use of those selected ESG indices will provide an incentive for Japanese companies to enhance responses to ESG issues so as to lead to an improvement of their corporate value in the long term".[29] It initially invested 4.2 percent of its domestic passive portfolio in three ESG indices in 2017, followed by a carbon "tilted" index (S&P-C) in 2018 (Table 14).[30] In March 2022 (the end of the 2021 fiscal year), about 14 percent of the domestic passive portfolio was invested in the ESG indices.

In 2017 the fund reinforced its commitment to active ownership and stewardship. It shifted its stance towards external managers from monitoring to engagement. GPIF "proactively held dialogues with external asset managers on matters such as exercise of proxy voting rights" to "evaluate that external asset managers are exercising proxy voting rights in compliance with GPIF Proxy Voting Principles by integrating ESG factors into the investment process".[31] GPIF also engaged in a dialogue with ESG analysts that supply the ratings for the ESG indices it adopted. Through the dialogue GPIF sought to positively influence the methods employed by the evaluators and to encourage them to engage in a dialogue with GPIF portfolio companies during the evaluation process (ESG Report 2017, pg. 12).

In 2018 GPIF further promoted its active ownership stance by paying two of its passive managers separately for engaging with companies in the Japanese TOPIX index, Asset

---

[26] GPIF Annual Report 2020, pg. 29, the fee was reduced to 2bp in FY 2021.
[27] GPIF Annual Report 2013, pg. 19.
[28] ESG Report 2018, pg. 4.
[29] GPIF Annual Report 2017, pg. 40.
[30] The "best in class" indices were FTSE Blossom, MSCI-ESG and MSCI-WIN), the "tilted" index was S&P-C.
[31] ESG Report 2017, pg. 11.

Management One (AM One) and Fidelity International (FIL).[32] AM One's asset management mandate is the larger one, managing 25.7 percent of GPIF's passive equity portfolio in 2020, compared with 0.3 percent for FIL (Table 14).[33] In the next sections, we present empirical analysis of the impact of GPIF's engagement initiative.

3.   The AM One Engagement Programme

In this section we focus on AM One, the external manager that received the mandate to develop a dedicated engagement programme focused on ESG in 2018.[34] The intention by GPIF was to "diversify and enhance our approach to stewardship and improve the quality of the entire market through these [engagement] activities".[35] Our analysis of active ownership is based on private information provided to us by AM One. It contains records of the contacts with portfolio companies between the inception of the programme in 2018 and March 2022.[36]

AM One's approach is based on 20 ESG engagement items that are classified into six vertical E, five S, four G categories and five horizontal ESG categories, for example CSR/ESG Management (ESG1), Corporate Misconduct (ESG2) or Digital Transformation (ESG5) (see Figure 1). Progress is self-assessed and measured through an eight stage "milestone" system (Figure 1). The engagements are conducted in writing, through personal meetings and video conferences; the latter became the dominant form of communication in 2020 due to the Covid pandemic. Contacts are often with senior management and board members.

AM One started its engagement programme in its fiscal year 2018 (starting in April 2018) and has engaged with 571 companies by the end of their most recent fiscal year (March 31, 2022). The total number of engagements by theme is 2,292 (Table 1 Panel A). These were divided into "base engagements" when an issue was raised and "follow-up engagements" with multiple contacts and progress measured through the "milestones". The total number of contacts was 3,785 (Table 1 Panel A) and there were up to 17 contacts per company (Panel B). Almost all TOPIX100 large-cap and three quarters of the Mid400 segment were engaged (Panel C).

---

[32] The arrangement is like the mandate given by other foreign and domestic investors to the Japan Engagement Consortium run by Governance for Owners Japan; see Becht, Franks, Mijajima and Suzucki (2022).

[33] AMOne has also been the largest external manager of GPIF's domestic passive equities portfolio since 2017, both as a fraction of the domestic passive portfolio and in terms of the absolute amount under management. AMOne also manages the FTSE Blossom ESG index portfolio.

[34] In addition to providing stewardship services, AM One is also GPIF's largest passive fund manager. In this capacity AM One also votes the GPIF holdings it manages, like all external managers, in close alignment with GPIF's proxy voting policies.

[35] GPIF ESG Report 2020, pg. 25

[36] AM One provided these records under a non-disclosure agreement that gave the asset manager the right to comment on drafts of this paper, but not to censor any of the results. In parallel GPIF was kept informed of our research findings.

11

The most common theme was "G" (1,266), followed by the horizontal "ESG" (726), "E" (176) and "S" (124). Contacts with the company were often at the senior management or board level and sometimes involved the President or the CEO (Table 1 Panel A).

It is possible to measure the impact of engagement on ESG performance through milestone progress, but only for follow-up engagements.[37] There are eight milestone levels (Table 2): (1) identifying an ESG issue; (2) raising concerns/suggestions; (3) issues recognized; (4) issues recognized by senior management; (5) initiatives taken; (6) plans formulated; (7) plans implemented; and, (8) completing the engagement. Progress is not linear. Level (3) is progress from Level (2) because the company starts to recognize the issue. Level (4) is material progress because it indicates that the company's senior management has recognized that there is an issue. Similarly, Level (5) indicates that initiatives have started to be taken and Level (6) indicates that the company is formulating plans to improve the situation. Levels (7) and (8) show that the engagement is basically complete; for almost 40% of engagements these highest milestones were reached by March 2022 (Table 2).

Table 2 also allows us to monitor milestone progress from the start of the engagement to March 2022. There are very few cases where no action was taken after an ESG issue was identified with no initial progress (3.1%). The most common initial milestone was set at levels 3 and 4, with the ESG issue being recognised by management (54.2%). Finally, Panels B and C of Table 2 provide a cohort view by conditioning on the start date of the engagement. Panel B shows that the start dates are spread relatively evenly over the year for base and follow-up engagements. Panel C shows that the fraction of completed follow-up engagements is 55 percent (139/251) for the earliest 2018 cohort.

The initial analysis raises some questions. Was progress fast because the issues identified by AM One were uncontroversial? For example, it might be relatively easy to recognize that there are Board Governance and Accountability or CSR/ESG Management issues and take some form of action. A second question is: Did base engagements have impact? For base engagements it is not possible to measure progress in terms of milestones, but we can measure their impact on ESG scores. If there was an impact on ESG scores, did it lead to inclusion in

---

[37] The milestone methodology was introduced in 1999 by Friends, Ivory & Sime and formed the backbone of its Responsible Engagement Overlay (reo®) engagement service; today the service is owned and provided by BMO Global Asset Management. Milestone progress was also used in previous studies to measure the impact of engagement on ESG (Dimson et al., 2015). (https://www.bmogam.com/uploads/2021/06/67799ebe74f49324b5dfa32bbeed98fb/influencing-for-good-lessons-from-20-years-of-engagement-.pdf). A four stage milestone system is also used by the Federated Hermes Equity Ownership Service (Hermes EOS) (https://sustainability.hermes-investment.com/uploads/2021/02/b69dc219bdbedc8a5188ea7b61bf6569/eos-engagement-plan-2021-2023-public-1.pdf)

the best-in-class indices chosen by GPIF? Most important, did the scores increase more rapidly for companies that were engaged by AM One compared with those that were not engaged by the asset manager, i.e., a control group? Was progress in terms of scores more rapid after the creation of the engagement programme, than before it? We investigate these questions in the next section.

4.   The Impact of Engagements on ESG Performance

In this section we evaluate the potential impact of AM One's engagement on ESG performance. Our sample is drawn from all shares listed on the Japan Exchange Group (JPX) markets between 2009 and the end of March 2022. This includes companies in the TOPIX all shares index (listed on "1st Section") and those listed on the 2nd Market ("2nd Section"), the TSE Growth Section ("Mothers"), JASDAQ ("Standard" and "Growth") and Real Estate Investment Trusts.[38] We provide descriptive statistics for all JPX segments, but the focus of the study is on the TOPIX index. The TOPIX is GPIF's main benchmark index, and it is GPIF's declared goal to improve the ESG performance of TOPIX constituents.

To perform the engagement analysis, we matched the AM One data with JPX segment information followed by FTSE and MSCI ESG scores. In a second step we split the sample into companies that have been engaged by AM One and those that were not engaged by them. In each case we distinguish between ESG, E, S and G engagements and scores.

To assess the FTSE and MSCI coverage by JPX segment we used the list of JPX companies on the last trading date of March of each year, the end of the fiscal year in Japan for most Japanese listed companies as well as for GPIF, and matched it with the most recent ESG scores available, the FTSE scores from the beginning of December of the previous year and the MSCI scores from the beginning of March of the same year.[39]

Table 3 shows the number of JPX companies by segment, the availability of ESG scores by year and the average score by segment. Panel A shows that the TOPIX index is divided into five segments: The Core 30 and Large 70 that together make up the TOPIX 100, the Midcap 400, the Small 1 with around 500 additional constituents and the Small 2 with over 1,000 further constituents after 2017.

---

[38] There is also a small group of companies that will be included in the TOPIX, but have not been assigned to a segment yet.
[39] GPIF portfolio holdings are reported once a year on 31 March. The data includes the number of shares held by GPIF and their market value, but not the fraction of shares issued and outstanding held by GPIF.

13

FTSE and MSCI ESG scores are available for almost all TOPIX 100 companies. The coverage of FTSE scores is almost 80 percent for the Midcap 400 in 2015/16 and just over 50 percent for MSCI; coverage increases to over 90 percent in 2017 for both providers. The variation for small caps is much larger. There are hardly any scores in 2015/16 for the Small 1 segment but coverage increases to almost 90 percent in 2022; for Small 2 the coverage was sparse throughout (Table 3, Panels D and E).

Average ESG scores correlate with market capitalisation; they were higher for the Core 30 throughout the sample period for both FTSE and MSCI (Table 3, Panels F and G). Average FTSE scores for Mid-Caps are lower, while the difference in MSCI scores is less pronounced. We confirmed MSCI Mid-cap scores suffer from sample selection in 2015 and 2016; companies with a higher score were more likely to be rated, resulting in higher average scores. The average scores for small caps are not representative in earlier years because of sparse coverage. FTSE only extended its scoring in 2020 from the FTSE Japan Large-Midcap Index with approximately 500 constituents to the Japan All Cap Index with over 1,000 constituents. MSCI covered 86 percent of the TOPIX Small 1 in 2020.[40] The lack of scores for small-caps in the pre-engagement period and the lack of subsequent engagement lead us to limit the analysis to the TOPIX 500. AM One does not use ESG scores to select engagement targets, but we still wished to know if the mean and median scores for the companies that were engaged were above or below the FTSE and MSCI index inclusion threshold.

Table 4 shows FTSE and MSCI ESG scores for companies that were engaged at the start of each engagement.[41] We used pillar scores for E, S and G engagements and ESG scores for horizontal ESG engagements. FTSE scores are updated twice a year in June and December. Although the MSCI is reported at monthly intervals we use six monthly data from June and December to make it comparable with FTSE. To make sure AM One could observe the score before starting the engagement, we always used the FTSE and MSCI score from the previous period. For reference, an overall ESG score of 3.1 was required for inclusion in the Blossom index until December 2019 and a score of 3.3 thereafter.[42] For MSCI the required score for inclusion was BB or above; to avoid exclusion the score had to remain above CCC.[43] The table

---

[40] The MSCI Japan index and the TOPIX All Share index do not overlap perfectly.

[41] FTSE scores go from 0 to five, while MSCI uses scores from 0 to 10. To put the scores on a similar scale we halved the MSCI score. We did not change the distribution of scores.

[42] MSCI uses a letter rating as a threshold for index inclusion.

[43] In 2017 a B rating translated into an industry adjusted score between 1.5 and 2.8; a BB to 2.9 to 4.2 and a CC to 0 to 1.2; the original MSCI scale is from 0-10. For comparability with FTSE all other MSCI scores were linearly projected onto a scale from 0 to 5.

shows that AM One engaged companies with relatively low scores initially; later engagement cohorts had higher average scores.

Basic Difference-In-Differences

We assess the potential impact of engagement on ESG performance by analysing the difference in differences, a method that is widely used in project and programme evaluation (Gertler et al. 2016; Angrist and Pischke 2009). In this paper, it is not possible to draw causal inferences about the impact of engagement on ESG scores from simple "before and after" comparisons because there are time varying influences that affect the scores of all companies in the TOPIX500 index, such as changes in government. Also, it is likely that there are time invariant factors that make an engagement with a particular company more likely and that some of the company characteristics are unobservable.

The basic model assumes that there are *n* units (*i*) and T=2 periods (*t*). Treatment (engagement) is a binary policy D_{it} and we want to obtain an estimate of the effect of the treatment on outcomes Y_{it}, the average treatment effect on the treated (ATT). The problem is that D_{it} is not randomly assigned. It is possible to identify the ATT assuming parallel trends and non-anticipation of the policy. Parallel trends assume that in the absence of treatment the average outcomes would have evolved on the same path after treatment as before. The ESG scores for companies that were engaged could have different levels from those that were not engaged before the treatment, but without treatment the change in scores would have been the same.

The basic design to estimate the ATT is via a simple two-way fixed effects estimator (TWFE):

$$(1) \qquad Y_{it} = \alpha_i + \gamma_t + D_{it}\beta + \epsilon_{it}$$

where α_i is a unit fixed effect, γ_t is the time fixed effect and D_{it} is the treatment indicator. The parameter estimate of β provides the pooled average treatment effect for the pre-treatment and post-treatment periods.

To control for time varying factors that influence ESG scores for all companies, we compare the difference in scores between those that were engaged with those that were never engaged. For example, the adoption of best-in-class ESG indexes in 2017 affects both groups. In contrast, engagement only affects the treatment group. Without the control group we could not draw causal inferences about the impact of the engagement programme. The difference in the differences also eliminates the influence of factors that are time invariant, including those that are unobserved. Left uncontrolled we might falsely attribute a change in ESG scores to

15

engagement, although the change may be driven by company characteristics that caused the company to be engaged. The technique provides a good estimate of the counterfactual; what would be the change in scores for the engaged in the absence of an engagement?

In our setup the dependent variables are ESG scores from FTSE and MSCI respectively that are available at half-yearly intervals. In this case the most straightforward way to estimate the impact of engagements on ESG scores is through a pooled panel regression model with an interaction term (Baker et al., 2022)

$$(2) \qquad Y_{it} = \beta_1 + \beta_2 \, Treat_i + \beta_3 \, Post_t + \beta_4 \, (Treat * Post)_{it} + \varepsilon_{it}$$

Where Y_{it} are ESG scores, *Treat$_i$* is an indicator variable set to one if a company has been engaged at some point between June 2018 and December 2022 and to zero if it was never engaged. Similarly, *Post$_t$* is set to one for observations in the period June 2018 and December 2022 and to zero otherwise. The coefficient of interest is β_4 that is readily available with robust standard errors.

To make valid inferences it is necessary that the scores of the treatment and the control group evolved in parallel before the start of the treatment and would have continued to evolve in tandem in its absence. If not, the difference in differences would simply pick up the continued divergence in trends. We performed a visual inspection of data availability and of parallel trends by plotting the evolution of the total number of scores and of mean ESG scores for FTSE and for MSCI. The mean MSCI scores were not parallel before June 2016 because engagement was correlated with coverage and scores.[44] We also noted that there was a discontinuity in the raw MSCI G score that was caused by a change in methodology in November 2020. This resulted in a distortion of the unweighted MSCI ESG score. Hence, the MSCI time series only starts in June 2016 and ends in December 2020. The FTSE sample spans the period 2015 to 2021 and includes two rating updates, one in June and one in December. In the cross-section we excluded small-cap firms because they were sparsely engaged and corresponding ESG scores were largely unavailable before 2020 (Table 3).

After these exclusions we proceeded to plot the mean scores again for the selected sub-samples. The analysis confirmed that the means of the engaged and those that were never engaged evolved in parallel before engagement for FTSE (Figure 2, Panel A) and for MSCI (Figure 2,

---

[44] Coverage is not a time invariant factor.

Electronic copy available at: https://ssrn.com/abstract=4506415

Panel B). However, the MSCI series is significantly shorter due to lower data availability prior to 2017 and a material change in methodology in 2020.

We run separate pooled panel regressions for five different types of treatment as defined by AM One (Figure 1): (1) overall ESG scores on any type of engagement (E, S, G or ESG); (2) E scores on environmental engagements (E1-E6); (3) S scores on social engagements (S0-S4); (4) G scores on governance engagements (G0-G3); and (5) ESG scores on ESG engagements (ESG1-ESG5). We also ran pooled quantile regressions at the 10th, 25th, median, 75th and 90th quantile to assess if there was a differential impact for high and low scores companies.

The overall FTSE scores for the group of companies that were engaged increased significantly relative to companies that were not engaged (Table 5). The effect is positive and significant at the 1% level where treatment is defined as any type of engagement (Panel A, Column 1). In absolute terms AM One E-engagements have the largest significant effect on FTSE E scores; the difference in the differences between the engaged and the never engaged increased by 0.3 on the five-point scale (Table 5, Panel A, Column 1). The effect is driven by the impact of the engagement on those companies with low E-scores; the difference improved significantly by 0.5 for scores at the 10th percentile and 0.5 for the 25th percentile (Table 6, Panel 2). The impact of engagements on MSCI scores is negligible, except for G (Table 6, Panel B, Column 4). Again, the result is driven by the impact of the engagement on the companies with the lowest G-scores (Table 6, Panel 4, Column 6).

Multiple Periods and Staggered Treatment

The previous analysis did not take advantage of the time series features of the dataset. It also ignored that companies were engaged at different points in time. To take advantage of the multiple period setup and to account for the staggered rollout, we apply the (Callaway and Sant'Anna, 2021) multiple period estimator.[45] Its building block is "the group-time average treatment effect on the treated *ATT(g,t)*" which captures the average treatment effect at time *t* for the cohort first treated at time *g* (Roth et al., 2023, p. 16). For example, ATT(2018,2020) is the average treatment effect for companies in 2020 that were first engaged in 2018. An estimate of ATT(2018,2020) is obtained by comparing the change in the ESG score of companies in 2020 that were engaged in 2018 with those that had not been engaged yet in 2020. In event time ATT(2018,2020) is equivalent to ATT(2019,2021).

---

[45] The paper devotes considerable attention to robust estimation in the presence of co-variates; we have no co-variates.

The `csdid` package (Rios-Avila et al., 2021) can aggregate ATTs by cohort, in event time or both; the latter yields a single parameter that is comparable to the interaction term from the static regression. The event time aggregation is of particular interest to us; it puts all cohorts onto the same timeline, setting the year of first engagement to zero. This allows us to visualise the impact of the engagements over time. The package also allows us to define the control group as companies that were "never treated" and companies that were "not yet" treated.

The estimator allows us to test the pre-trend assumption and to assess if the potential engagement effect is immediate, increases over time or reverts. Table 8 reports the results of the multiple period regression estimates. Panel 1 shows average treatment effects for the treated *ATT(g,t)* in event time denoted $\theta_{es}(e)$ ((Callaway and Sant'Anna, 2021, p. 209). The ATTs are estimated using all periods relative to the period of the first treatment across all cohorts.[46] For any treatment (E, S, G or ESG) and FTSE ESG scores the pre-treatment period is not significant; the post-treatment period is highly significant. In addition, the treatment effect increases over time, which is consistent with the notion that engagements take time to show effect. It is also consistent with multiple treatments.[47] The pre-treatment effects were also not significant for E, S and G. Consistent with the results from the basic two period model, E engagements significantly improved E scores. However, S engagements significantly decreased S scores after three periods.

In the previous section we also assumed that companies were engaged simultaneously in 2018 or never engaged. In practice companies were engaged earlier or later in the programme. We already showed that the rollout of AM One's engagement programme was staggered (Table 2 Panel B). In addition, in terms of self-reported milestones the number of completed engagements was higher for the first (2018) than for later cohorts (Table 2 Panel C).

Ignoring the cohort effect that results from the staggered rollout can lead to biased estimates (Borusyak et al., 2022; de Chaisemartin and D'Haultfœuille, 2020; de Chaisemartin and D'Haultfœuille, 2022; Goodman-Bacon, 2021; Sun and Abraham, 2021). We classified units that had not been treated yet as treated and compared them to units that were never treated; for example, units that were first engaged in 2020 were already classified as engaged in 2018 and 2019.

---

[46] The Stata estimation command is using `csdid` and `estat event` (Rios-Avila et al., 2021).

[47] As we showed before, in many cases there were multiple contacts; we do not estimate the effect of repeated treatment separately in this paper.

Figure 3, Figure 4 and Table 7 Panel 1 provide further details on the timing of the rollout for different JPX segments and the sample used in the regression analysis. Figure 3 shows the fraction of each index segment that was engaged over time on any of the twenty AM One themes. Figure 4 shows the same evolution for E, S and G themes separately. Engagements started in June 2018 and the fraction of companies engaged at least once increased rapidly. For the TOPIX 100 over 50 percent of the constituents had at least one contact by the beginning of 2019 and nearly all companies had been engaged by the end of the sample period. For the Mid400 over 60 percent of companies were engaged, but only 10 percent for the small cap indexes. There were substantial differences across the E, S and G themes even for the large-cap segments; fewer than 50 percent of TOPIX100 constituents were engaged on S themes and just over half on environmental themes. Table 7 Panel 1 provides an overview of the engagement roll-out for annual cohorts. The sample is limited to companies that were in the TOPIX 500 and had a FTSE score at the end of the year.

Figure 5 Column 1 plots the average treatment effects reported in Table 8 Panel 1 with standard error bars. The delayed impact of the engagements on ESG scores and of E engagements on E scores is clearly visible. Column 2 reports results when the control group is defined as companies that have not yet been engaged. The results are almost identical.

Table 8 Panel 2 reports the results by cohort and the overall effect. Columns 1-5 show ATT estimates by treatment group. Column 1 shows the average across all groups; columns 2-5 show averages for the 2018, 2019, 2020 and 2021 cohorts (`estat group`). The treatment effect is most pronounced for the 2018 and 2019 cohorts because they were the earliest to be engaged. The result is thus consistent with the event time results reported in Panel 1.

Table 8 Column 6 shows average *ATT(g,t)* estimates for all groups across all periods (`estat simple`). Column 7 reports the coefficient of the interaction term in a pooled panel regression for the same sample (TWFE).[48] The overall *ATT(g,t)* estimate is highly significant for overall ESG scores, but not significant for E, S or G. However, the impact on E scores is highly significant for the 2018 and 2019 cohorts. This illustrates that the size and significance of the overall *ATT(g,t)* is dependent on the weighting of the individual building blocks. Overall the staggered difference in differences results while confirming previous results, also yield new insights on the timing of the impact following the first treatment.

---

[48] Coefficient and standard errors are not fully identical to Table 5 because the estimates rely on December scores, while the previous results used bi-annual data for June and December.

FTSE vs. MSCI Scores

Why is the impact of engagement larger in FTSE ESG scores compared with MSCI scores? Although the raw scores increased during the engagement period, the weights for E and S for the engaged sample declined more than for the control sample. The decline in weights largely offset the improvements in raw scores, leaving the overall ESG scores largely unchanged. We formally show this by replacing MSCI scores with materiality weights in the difference-in-differences regressions (Table 5, Panel B, Columns 6-8). The mean materiality E weight for the engaged was 40 before treatment and 37 afterwards; for companies that were never engaged the E weight was 26 before treatment and 23 thereafter. There was a decline in materiality for both groups, but the difference in the differences was more pronounced for the treated. However, the decline is not statistically significant.

We explore this question further in the next section where we analyse an engagement with 7&i Holdings and an anonymous utility company. We show that the materiality weighted-average methodology of MSCI ratings makes it more difficult for the impact of E and S engagements to be captured. Materiality weights adjust dynamically over time at the company level. Hence, paradoxically, engagement success results in lower materiality weights and thereby lowers the weighted scores at the next higher level. In contrast, we show that FTSE scores improved significantly for E and the exposure scores were stable over time. A company targeted by AM One with a high FTSE exposure score that shows an increase in the E score will also show a higher overall score. In contrast, the same company will have an increasing MSCI E score and a decreasing materiality E weight (risk), resulting in an overall score that is barely changed. In addition, FTSE scores consider non-disclosure as a negative signal. Since AM One's engagements often focused on Climate-Related Financial Disclosures (TCFD) and E disclosure improved, it is not surprising that the FTSE E score increased.

5.   Case Studies

To investigate the underlying reasons for the observed divergence in the response of FTSE and MSCI scores to AM One engagements in more detail, we conducted an analysis of two specific engagements conducted by AM One. Through these case studies, we can better understand the impact of the distinct methodologies employed by each scoring system. We chose one engagement with a high S exposure and one company with a high E exposure; in both cases AM One recorded milestone progress.

20

Our analysis confirmed that there are several key factors that contribute to the sensitivity of FTSE scores to engagement outcomes. Firstly, FTSE assigns a zero score to items that are not disclosed, resulting in a reduction of the average score. Conversely, MSCI attributes a missing value to non-disclosure, which has no effect on the average score. This important disparity in scoring methodology influences the sensitivity of FTSE scores to engagement outcomes.

Secondly, MSCI's materiality weights are company-specific and involve significant fluctuations over time. In contrast, FTSE's exposure scores (equivalent to materiality) tend to be more stable. Moreover, improvements in engagement scores may coincide with lower materiality weights according to MSCI's approach, thereby dampening the effect of a successful engagement on their scores. The lower weights may reflect a lower perceived risk, and this may be negatively correlated with improvements in the scores. We show that this issue does not arise with FTSE, that keeps exposure scores relatively constant.

While we have not undertaken a comprehensive comparison of FTSE and MSCI, we believe that for asset managers constructing engagement agendas, FTSE scores may offer some advantages over MSCI scores. However, we report both for completeness. For a comparison of the FTSE and MSCI scoring methodologies, see the Internet Appendix. We now turn to the two case studies.

5.1 Case Study 1: 7&i Holdings

Table 9 and Table 10 establish a connection between the engagement agendas pursued by AM One for Seven & i Holdings, a prominent retail operator in Japan known as "7&i," and the corresponding FTSE and MSCI scores.

In Table 9 we focus on 7&i and engagements at the theme level pertaining to three areas: E5 Water and Pollution, S2 Labour Practices / Health and Safety, and ESG4 CSR Supply Chain Management. These engagements were initiated in July, November, 2018 and January 2019, respectively. Given the nature of 7&i's business all three themes involve material risk.

The company responded immediately regarding S2, as reported by AM One. According to their assessment, the milestone of 7, denoting the successful implementation of plans, was quickly achieved. Also, by March 2021, the issues relating to E5 and ESG4 also reached a milestone of 7. Our examination of AM One's engagement records suggests that their evaluation for a milestone of 7 is primarily based on the retailer's disclosure of their own response to each agenda item.

21

The FTSE scores relating to the three engagement areas show significant improvements. Specifically, within the "Pollution and Resources" category, which falls under the E pillar, there was an increase from 3 in March 2019 to 4 in December 2021. The corresponding exposure measure for this theme was 2, denoting a medium level. However, despite this positive change, the overall E score exhibited a smaller proportionate change between March 2019 and December 2021, rising from 3.3 to 3.7. This is primarily because the "Climate Change" score, an aspect that AM One did not engage on, declined from 4 to 3. The exposure measure for this theme was also medium.

The S score, which encompasses both the "Labour Standards" (AM One's S2 theme) and the "Social Supply Chain" (ESG4 theme), both showed an improvement, rising from 3 to 4 and 3 to 5, respectively. Consequently, the overall S score experienced an increase from 3.2 to 4.5. Given the nature of 7&i's business, the exposure scores for both themes were 3, signifying a high level of importance. As a result of these developments, the company's FTSE overall ESG rating rose from 3.2 to 4.2 between March 2019 and December 2021. It is important to note that the progress in pillar scores can be attributed to AM One improving S theme scores rated as having high exposure: a theme score of 5 with high exposure results in 71-100% of the indicator points feeding through to the pillar score. Importantly, the exposure to the theme remained high throughout the engagement period; the S performance of 7&i improved and given the nature of its business, there is high exposure to S risk.

Table 9 Panel 2 focuses on MSCI scores. As previously explained, MSCI assigns weights to each theme level score based on the company and the industry. In the case of 7&i the (latest) weight assigned by MSCI to the E score at the beginning of the engagement period was relatively low at 18%, while the weight assigned to the S score was significantly higher at 63%.

For the agenda items associated with the three engagement themes of AM One, for the S pillar, S2 is linked to the MSCI theme "human capital theme score." In January 2019, this score stood at 3.8, carrying a weight of 21%, but by December 2021, the score had declined to 1.5, with a weight of 15%. The large decline in the score was mitigated by its lower weight in the overall score. The overall S score experienced only a slight change, decreasing from 4.3 (with a weight of 63%) to 4.1 (with a weight of 52%).

As for the E5 engagement theme that focused on waste and pollution, MSCI's related theme is "natural resource usage" and showed an improvement in scores from 2.4 to 2.7. However, this change had limited influence on the overall ESG index due to MSCI's allocation of less than

10% weight to this theme. For AM One's ESG4 theme, this relates to the MSCI theme, CSR supply chain management, but this was assigned zero weight by MSCI, thereby having no impact on their scores.

The overall weighted average MSCI ESG score for the retailer changed only marginally from 4.4 in January 2019 to 4.5 in December 2021, a much smaller increase than that reported by FTSE.

5.2 Case Study 2: A Utility Company

Table 10 provides an overview of the scores for an anonymous utility company, highlighting the outcomes of its engagement with AM One in relation to a single theme, E1 Climate Change. The engagement commenced in June 2018 and reached a milestone of 7, denoting the implementation of plans, by September 2019. The primary focus of this engagement revolved around enhancing the disclosure practices concerning climate change-related matters, such as the company's roadmap towards achieving carbon neutrality by 2025 and its increased use of renewable energy sources.

Even after reaching milestone 7, AM One chose to continue the engagement and the ongoing dialogue with the company, extending beyond September 2019. This suggests a continued interest in improvements and progress in the area of climate change.

The FTSE scores for the company shown in Table 10 suggest significant improvement in June 2019, coinciding with AM One reaching the milestone of 7. Notably, the Theme Score for "Climate Change" showed a large increase from 3 in December 2018 to 5 in June 2019. Given the industry's nature, the exposure score for this theme was rated as 3, indicating a high level. Given the theme score carried a high exposure it contributed significantly to the significant increase in the E Pillar score, which rose from 2.8 to 3.5.

Upon examining the indicators that made up the climate change score, we find that the improvement in the theme score stemmed from enhanced disclosure practices. Indicators that previously received a zero-score denoting "no-disclosure", with disclosure showed a significant improvement, reaching a score of 3, representing "good practice" or higher. The percentage of indicators disclosed under the "Climate Change" Theme increased from 69% in December 2018 to above 80% in June 2019, consistent with AM One's emphasis on raising the company's commitment to transparency and disclosure practices. Given the nature of the utility's business, the exposure measure was high throughout the period.

23

Table 10 Panel 2 focuses on MSCI scores, and shows a more pronounced pattern compared with the FTSE scores, discussed above. Initially, MSCI assigned a high materiality weight of 74% to the E pillar for this utility company. The E pillar comprises five E themes, with two themes directly relevant to AM One's E1 Climate Change engagement: the MSCI "Climate Change" Theme and the "Natural Resource Use" Theme.

In 2018, the climate change score was recorded at 6.9, accompanied by a theme weight of 20%. Similarly, the natural resource use score stood at 5.7, with a materiality weight of 17%. The climate change score decreased from 6.9 to 5.7 by July 2021, while the natural resource use score experienced an increase, from 5.7 to 6.4. However, both theme scores suffered a decline in materiality, dropping to 14% for both themes. The result was that, although the E Pillar score rose from 5.8 to 6.0 during the period, the pillar weight declined from 74% to 54%. This contributed to a decline in the weighted ESG score from 5.1 to 4.9.

ESG risks and materiality maps are industry specific. In addition, the MSCI methodology emphasizes the forward looking, dynamic and company specific nature of ESG risks and opportunities (MSCI, 2023). In the two cases we discussed this translates into materiality weights that change relatively quickly over time. Paradoxically, this makes MSCI ratings, a weighted average of scores and materiality weights, less suitable for assessing the impact of engagements. Successful engagements increase ESG scores but decrease company specific ESG risk, resulting in a weighted average that changes very little.

24

## 6.	ESG Index Inclusion and Exclusion

In this section we estimate the impact of index inclusion and exclusion on stock returns. As we described above, GPIF adopted two FTSE ESG indexes, the Blossom Index and the Blossom Sector Relative Index, and in addition a third index, the MSCI Japan ESG Select Leaders Index. All three are "best-in-class" and use ESG scores to determine inclusion or exclusion. Since the adoption of the Blossom Sector Relative Index falls outside of our sample period, we concentrate on the MSCI ESGL and the FTSE Blossom index.[49] We also study the FTSE4Good Japan index, the predecessor of the Blossom index.

Data and Methodology

An important incentive for portfolio companies to engage with AM One is that inclusion in an ESG index will increase GPIF's holdings in the company and, also potentially the holdings of other asset managers that also track the indexes for investment purposes. Leader indexes provide a tangible financial incentive for companies to improve their MSCI and FTSE scores if the extra investment is associated with positive excess returns. Equally, there will be a potential penalty from lower scores if exclusion is linked to negative abnormal returns. FTSE inclusions and exclusions are announced publicly so there is also a potential reputation effect.[50] Although MSCI does not make public announcements, companies have an incentive to announce inclusion in the MSCI ESG Leaders (ESGL) Index.[51] Finally, the Blossom index is almost identical to the FTSE4Good index that was started in 2004. This allowed us to compare inclusions and exclusions before and after the adoption of the Blossom index by GPIF.[52]

We use constituent data from FTSE and MSCI to identify the dates when stocks were included or excluded from the relevant indexes. We concentrate on the window twenty days before the effective date of the inclusion/exclusion and the five subsequent days. For FTSE and MSCI the announcement date is nine business days before the effective date.

---

[49] GPIF also adopted two further domestic ESG indexes, an S&P Carbon index and the MSCI WIN index. However, we do not include them in our study because they are more like a theme-based index, the former being about carbon emissions and the latter about gender diversity; see Mehrotra et al. 2023 for a study on the effects of inclusion using the WIN index. In addition, the S&P Carbon is a target index that tilts by changing weights, not a leader index that tilts portfolios through inclusions and exclusions.

[50] The impact of the JPX400 index on profit margins is attributed predominantly to a "shame and aspiration" reputation effect (Chattopadhyay et al., 2020).

[51] Many companies also report the continued inclusion in the FTSE and/or the MSCI index.

[52] There is no obvious control group in Japan, so we could not perform a difference-in-differences analysis.

25

For the FTSE4Good Japan, Blossom and MSCI index we have daily constituent data for individual companies in the indexes going back to the respective inception dates.[53] The Quick Astra Database provides daily data on stock price returns, index returns, and the number of shares issued and outstanding. We merged the daily data for the three ESG indexes with the daily data on JPX market segments (1st Section "一部", 2nd Section "二部", JASDAQ) and TOPIX index sub-segments using local identifier codes.

We wished to avoid inclusions and exclusions that were unrelated to changes in scores, for example, bankruptcies takeovers, and IPOs. For inclusions we dropped IPO companies included between index revisions. We also dropped cases when companies were included "simultaneously" in a TOPIX index and one of the ESG indexes; we defined "simultaneous" as 50 days or less before the ESG index effective date. For exclusions, we dropped takeovers or bankruptcies that resulted in delisting. Finally, we restricted inclusions to companies that were constituents of the TOPIX All Share index, the engagement universe of AM One. Hence, we excluded REITs and a small number of JASDAQ and 2nd Section market segment companies.

ESG Index Inclusion and Exclusion Returns

We estimate four different types of abnormal returns around effective inclusion and exclusion dates: raw returns, constant mean adjusted, market adjusted (MA) and market model (MM) adjusted (MacKinlay 1997). Results are reported in Table 11 and Table 12. Estimates from the market models suggest that inclusion abnormal returns are positive for the Blossom index and for the MSCI Leader index and symmetrically negative for exclusions. There are no significant inclusion and exclusion returns for the FTSE4Good Japan in the period preceding the GPIF programme (2009-2016).

For inclusions in the index the returns for the FTSE Blossom are 3.1% for raw returns and abnormal returns of 1.9% for the MM and 1.6% for the MA model (Table 11, Panel A).[54] The abnormal returns are higher for the MSCI ESG Leader (ESGL) index with raw returns of 1.8%

---

[53] For the FTSE index we also had index inclusion and exclusion announcements from September 2018 to March 2023 that contained the name of the company, the Japanese securities code, the announcement, and the effective date. We compared these announcements to the inclusion/exclusion history we had built from the daily constituent file.

[54] The return is significantly different from zero with 1% significance, except for the market model where the return is significant at the 5% level. The return is over a window from 20 days before the effective inclusion date to 5 days after inclusion. We do not consider the large number of simultaneous inclusions when the Blossom index was created in July 2017.

and abnormal returns of 2.1% for MM and 2.6% for MA model (Table 12, Panel A); although the difference between the model abnormal returns for the two indexes are economically small.

For exclusions the pattern of returns is predictable. For FTSE Blossom the raw returns are -4.4% and the abnormal returns are -3% for MM and -4% for the MA model. For the MSCI, the raw returns are -1.8% and the abnormal returns are lower than FTSE Blossom at -1.8% for the MM and -2.3% for the MA model.

The FTSE Blossom result is largely driven by 18 inclusions from the TOPIX Small index; for these events the average abnormal return is 5% for both the market adjusted and the market model (Table 11, Panel B). For MSCI ESGL inclusion the effect is driven by 35 inclusions from the TOPIX Mid Cap 400 segment). Inclusions from the TOPIX Small segment (Panel A4) were not significant. These results are consistent with a demand shock interpretation; the impact of the additional investment is particularly large for small and mid-caps stocks. However, it is surprising that the effect is not larger for mid-cap inclusions in the Blossom and small-cap inclusions in the MSCI ESGL index. Given the additional publicity we might have expected the inclusion effect to be larger for FTSE than for MSCI; we interpret the larger returns for MSCI as additional evidence in favour of a pure demand shock effect, possibly the result of the relatively larger amounts GPIF has invested in the MSCI ESGL (Table 14).

For MSCI exclusions look symmetric to inclusions (Figure 7) and larger than inclusions for FTSE (Figure 6). There are only 20 exclusions from the Blossom Index but 100 from the MSCI ESGL over the same period (Table 11, Panel B; Table 12; Panel B). The difference is due to a combination of factors: for the Blossom there is an exclusion "buffer zone", so it is difficult to be excluded once a company is included; for MSCI ESG scores are relatively stable but the exclusion threshold was raised from CCC in September 2017 to BB in May 2022, resulting in more exclusions; finally, FTSE scores increase over the period, making exclusions less likely.

It is plausible and likely that the prospect of inclusion in one of the ESG indexes adopted by GPIF also provides non-financial incentives to improve ESG scores. Equally the humiliation of exclusion provides an incentive to keep up the scores, especially vis-à-vis firms in the same industry.[55] We did find evidence that companies issue press releases after index inclusion but never after exclusion. Adoption of the relevant index by GPIF is explicitly mentioned. However, it is difficult to measure the impact of these non-pecuniary incentives.

---

[55] Inclusion and exclusion in the MSCI ESGL and the Blossom Sector Relative is decided by industry sector.

27

The inclusion/exclusion effects we document are consistent with the increase in average ESG scores we observe for all TOPIX companies. However, we cannot repeat the difference-in-differences analysis because we do not have a control group; all companies in the MSCI and FTSE universe were eligible for index inclusion.[56] In the next section we use international companies as a control group since they were not "treated" through GPIF's domestic programme.

7. Combined Impact of Index Tilting and Engagement

This section discusses the combined impact of the GPIF stewardship programme. The previous section showed that GPIF's adoption of ESG leader indexes created financial incentives to improve ESG ratings. The analysis of the engagement programme showed a positive net-impact on FTSE ESG and E scores. However, these results were net of any potential impact the adoption of the leader indexes had on the control group. Specifically, they excluded the impact of the introduction of the leader indexes on all Japanese companies. Also, they ignored the possibility that GPIF's engagement programme has stimulated other asset managers to engage more vigorously. We are informed by GPIF that they monitored ESG engagements of all their asset managers more intensively from 2018 onwards, also those that were not separately remunerated. This would contribute to an increase in overall scores but would not be reflected in our estimate of the impact of the remunerated engagement programme.

To test the impact of GPIF's overall programme on the Japanese market, we repeated the difference in differences analysis considering all Japanese companies as treated and using companies from foreign stock markets as the control group. Table 13 Panel A shows the number of available FTSE scores for large and mid-cap companies from various countries. The scores are counted annually on the last day of each year. The Japanese data primarily come from TOPIX500 companies. We define eight control groups: all companies for all countries, and seven additional groups based on the largest number of available country scores: Australia, Germany, France, United Kingdom, Hong Kong, South Korea and the United States.

To assess the parallel trends assumption, we plot the mean evolution of Japanese scores and the control groups (Figure 8). Two vertical lines highlight GPIF's adoption of the FTSE Blossom

---

[56] A related study of the WINS index uses the lowest scoring companies as a control group, assuming that they can never increase their score sufficiently to achieve index inclusion (Mehrotra et al., 2023). For ESG there are cases when companies made a "leap" from the lowest ESG score decile into the index. In the previous section we showed that the impact of remunerated engagement was particularly strong for the lowest quantiles, although engagement was not specifically targeted at low score companies.

and MSCI ESG Leader indexes and the first engagement by AM One. Japanese scores were significantly lower than for the other countries in the pre-treatment period, except for Hong Kong and South Korea. The graphs also show parallel trends, with the exception of Germany and France. The gap between the mean score for Japan and other countries visibly narrowed post-treatment; the exception was Korea where the gap widened, while scores for Hong Kong continued to move in parallel.

Table 13 Panel B provides parameter estimates from a simple pooled panel regression model with an interaction term. The use of ESG indexes by GPIF in July 2017 is considered the start of the treatment. After this date, the period dummy is set to one. The table allows for comparisons between Japan's scores and various control countries, offering insights into the specific impact of GPIF's programme. The impact of the programme was positive relative to all control groups, except Australia where the treatment effect was not significant and Hong Kong, where average scores improved relatively more than in Japan.[57]

Estimates of the dynamic treatment effect relative to the first treatment year (2017) are even more significant. Figure 9 plots the change in FTSE scores over time, starting with GPIF's adoption of the FTSE Blossom and MSCI ESG Leader index. Using the model from Callaway and Sant'Anna (2021), the figure shows event time impact estimates for the 2017 cohort. The results are consistent with the pooled cross-section regression but show how the treatment effect increased over time. This is consistent with the engagement results and the increased amounts GPIF invested in ESG indexes. It is also plausible that it takes time for companies to respond to the "treatment", for example by improving disclosure, by changing strategy or behaviour. The evidence suggests that GPIF's combined programme has significantly increased ESG scores in Japan when compared with other countries.

8.   Conclusions

GPIF pursues a dual strategy to improve the ESG performance of Japanese listed companies. The fund provides financial incentives by rewarding companies with high ESG scores which lead to entry into either of two ESG indexes, with additional equity investment provided by GPIF. It also encourages its asset managers to engage directly with the portfolio companies to improve their ESG policies by paying them a separate fee for engagement. Thus, the programme addresses two of the main criticisms of active ownership and portfolio choice by

---

[57] The Hong Kong Stock Exchange has required all Hong Kong listed companies to issue ESG reports since 2016 (Hang Seng Indexes Company Limited, 2023).

passive managers, the absence of remuneration for stewardship services beyond a standard asset management fee and the widespread use of target indexes that adjusts the weight of the whole portfolio over leader indexes that rely on inclusions and exclusions. Since the introduction of the two ESG indexes in 2017, FTSE ESG scores in Japan have increased significantly. Our evidence suggests that GPIF's natural experiment in responsible investment has contributed to this improvement in scores.

Notwithstanding, some changes might be considered to the current programme. The engagement programme by the asset manager is not explicitly linked to ESG scores, and thus only indirectly to inclusion or exclusion from GPIF's chosen ESG indexes. On the other hand, the current scoring methodologies are complex and raise significant issues. This is illustrated by the conflicting signals between the two ESG scores, where one suggests engagements by AMOne have been successful in significantly improving scores while the other indicates no significant improvement. The conflicting signals should not be a surprise since the scoring methodologies show considerable differences. For example, FTSE gives points for disclosure while MSCI does not. In addition, one index reduces the weight of a particular ESG component when the score goes up; in other words, successful engagements may reduce the reward from higher scores.

One solution is for asset managers to consider a simpler scoring method than can be used to set engagement objectives and measure engagement outcomes. These engagement compatible scores could then be used as the basis for portfolio choices. More effort and greater resources would be focused on the most material issues. Climate risk clearly stands out. A study commissioned by the Government of Singapore Investment Corporation (GIC) estimates that the cumulative loss on a mixed bond-equity portfolio would be ten percent from reaching net-zero and forty percent from a failed transition, echoing a similar assessment from the Financial Stability Board (Thomas and Houlder, 2023). The uncertainty and size of the risk warrants greater and more concentrated efforts.

Greater emphasis on tackling the largest risks might be helped by using an alternative to leader indexes that promise additional investment when certain conditions are met. For carbon emissions net-zero aligned indexes conditionally divest the highest emitters in line with the Paris Accord agreement and Japan's net-zero commitment.[58] Engagement by the asset manager

---

[58] For example the FTSE JPX Net Zero Japan Index Series, the MSCI World Climate Paris Aligned Index, the S&P Paris-Aligned & Climate Transition (PACT) index or a bespoke index developed for GPIF; see Bolton, Kacperczyk, and Samama (2022).

would aim to persuade high-carbon emitters to avoid publicly announced exclusion; and encourage excluded companies to reduce emissions sufficiently to benefit from re-inclusion.

31

9.   References

Andersson, M., Bolton, P., Samama, F., 2016. Hedging Climate Risk. Financial Analysts
    Journal 72, 13–32. https://doi.org/10.2469/faj.v72.n3.4
Angrist, J., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion.
    Princeton University Press, New Haven.
Azar, J., Duro, M., Kadach, I., Ormazabal, G., 2021. The Big Three and corporate carbon
    emissions around the world. Journal of Financial Economics 142, 674–696.
    https://doi.org/10.1016/j.jfineco.2021.05.007
Baker, A.C., Larcker, D.F., Wang, C.C.Y., 2022. How much should we trust staggered
    difference-in-differences estimates? Journal of Financial Economics 144, 370–395.
    https://doi.org/10.1016/j.jfineco.2022.01.004
Becht, M., Franks, J., Mayer, C., Rossi, S., 2009. Returns to Shareholder Activism: Evidence
    from a Clinical Study of the Hermes UK Focus Fund. The Review of Financial
    Studies 22, 3093–3129. https://doi.org/10.1093/rfs/hhn054
Becht, M., Franks, J.R., Miyajima, H., Suzuki, K., 2021. Outsourcing Active Ownership in
    Japan (ECGI Finance Working Paper No. 766/2021). ECGI.
    https://doi.org/10.2139/ssrn.3864310
Berg, F., Kölbel, J.F., Rigobon, R., 2022. Aggregate Confusion: The Divergence of ESG
    Ratings. Review of Finance 26, 1315–1344. https://doi.org/10.1093/rof/rfac033
Berk, J., van Binsbergen, J.H., 2021. The Impact of Impact Investing (Stanford University
    GSB Research Paper No. 22– 008). Stanford University.
    https://doi.org/10.2139/ssrn.3909166
Bolton, P., Kacperczyk, M., Samama, F., 2022. Net-Zero Carbon Portfolio Alignment.
    Financial Analysts Journal 78, 19–33.
    https://doi.org/10.1080/0015198X.2022.2033105
Borusyak, K., Jaravel, X., Spiess, J., 2022. Revisiting Event Study Designs: Robust and
    Efficient Estimation. https://doi.org/10.2139/ssrn.2826228
Broccardo, E., Hart, O., Zingales, L., 2022. Exit versus Voice. Journal of Political Economy
    000–000. https://doi.org/10.1086/720516
Callaway, B., Sant'Anna, P.H.C., 2021. Difference-in-Differences with multiple time periods.
    Journal of Econometrics, Themed Issue: Treatment Effect 1 225, 200–230.
    https://doi.org/10.1016/j.jeconom.2020.12.001
Chattopadhyay, A., Shaffer, M.D., Wang, C.C.Y., 2020. Governance through shame and
    aspiration: Index creation and corporate behavior. Journal of Financial Economics
    135, 704–724. https://doi.org/10.1016/j.jfineco.2019.07.005
Curran, M., Moran, D., 2007. Impact of the FTSE4Good Index on firm price: An event study.
    Journal of Environmental Management 82, 529–537.
    https://doi.org/10.1016/j.jenvman.2006.02.010
Davies, S.W., Van Wesep, E.D., 2018. The unintended consequences of divestment. Journal
    of Financial Economics 128, 558–575. https://doi.org/10.1016/j.jfineco.2018.03.007
de Chaisemartin, C., D'Haultfœuille, X., 2020. Two-Way Fixed Effects Estimators with
    Heterogeneous Treatment Effects. American Economic Review 110, 2964–2996.
    https://doi.org/10.1257/aer.20181169
de Chaisemartin, C., D'Haultfœuille, X., 2022. Two-way fixed effects and differences-in-
    differences with heterogeneous treatment effects: a survey. The Econometrics Journal.
    https://doi.org/10.1093/ectj/utac017
Dimson, E., Karakaş, O., Li, X., 2015. Active Ownership. Review of Financial Studies 28,
    3225–3268. https://doi.org/10.1093/rfs/hhv044
Edmans, A., Levit, D., Schneemeier, J., 2022. Socially Responsible Divestment (Working
    Paper No. 823/2022), ECGI Finance Working Paper. ECGI, Brussels.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L., Vermeersch, C.M.J., 2016. Impact Evaluation in Practice. International Bank for Reconstruction and Development, Washington DC.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. Journal of Econometrics, Themed Issue: Treatment Effect 1 225, 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

GPIF, 2018. GPIF Selected Global Environmental Stock Indices [WWW Document]. URL https://www.gpif.go.jp/en/%28Full%20version%29GPIF%20Selected%20Global%20 Environmental%20Stock%20Indices.pdf

Hang Seng Indexes Company Limited, 2023. Introduction of ESG [WWW Document]. Hang Seng Indexes. URL https://www.hsi.com.hk/eng/esg/esg-introduction (accessed 8.3.23).

Hawn, O., Chatterji, A.K., Mitchell, W., 2018. Do investors actually value sustainability? New evidence from investor reactions to the Dow Jones Sustainability Index (DJSI). Strategic Management Journal 39, 949–976. https://doi.org/10.1002/smj.2752

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

Henderson, R., Serafeim, G., Lerner, J., Jinko, N., 2019. Should a Pension Fund Try to Change the World? Inside GPIF's Embrace of ESG. (Harvard Business School Case No. 319– 067).

Kaspereit, T., 2022. Eventstudy2: Stata module to perform event studies with complex test statistics. Statistical Software Components.

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

Lin, K., Kimura, Y., Inoue, K., 2023. Selection and Effects of Environment and Social Engagements by Institutional Investors.

MacKinlay, A.C., 1997. Event Studies in Economics and Finance. Journal of Economic Literature 35, 13–39.

Mehrotra, V., Roth, L., Tsujimoto, Y., Wiwattanakantang, Y., 2023. Empowering Women by Index Membership: Evidence from a Unique Experiment from Japan. ECGI Finance Working Paper (forthcoming).

MSCI, 2023. ESG Industry Materiality Map [WWW Document]. URL https://www.msci.com/our-solutions/esg-investing/esg-industry-materiality-map (accessed 8.8.23).

Ratsimiveh, K., Haalebos, R., 2021. ESG scores and beyond. Part 2: Contribution of themes to ESG Ratings: a statistical assessment (Index Insights), Index Insights.

Rios-Avila, F., Callaway, B., Sant'Anna, P.H.C., 2021. csdid: Difference-in-Differences with Multiple Time Periods in Stata. Presented at the Stata Conference, p. 47.

Roth, J., Sant'Anna, P.H.C., Bilinski, A., Poe, J., 2023. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. Journal of Econometrics 235, 2218–2244. https://doi.org/10.1016/j.jeconom.2023.03.008

Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics, Themed Issue: Treatment Effect 1 225, 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Thomas, N., Houlder, V., 2023. Lex in depth: How investors are underpricing climate risks. Financial Times.

Villa, J.M., 2016. Diff: Simplifying the Estimation of Difference-in-differences Treatment Effects. The Stata Journal 16, 52–71. https://doi.org/10.1177/1536867X1601600108

Yilmaz, M.K., Aksoy, M., Tatoglu, E., 2020. Does the Stock Market Value Inclusion in a
Sustainability Index? Evidence from Borsa Istanbul. Sustainability 12, 483.
https://doi.org/10.3390/su12020483

10.   Tables

Table 1. AM One Engagements – Number of Themes and Contacts

Panel A shows the number of themes AM One engaged on and the total number of contacts per theme. The table also shows the level of the contact at the company: A+ - CEO and/or President; A – board member or senior management; B – middle management; C – employee. There were 7 contacts where the contact level was not recorded; these are not reported in the table. Panel B shows the number of engagements per company distinguishing between any engagement, base engagements and milestone engagements. The total number of companies engaged per category do not add up because the many companies have base and follow-up engagements.

Panel A – Engagement Themes and Level of Contact

Engagement Theme		Number of Themes	Number of Contacts	Level of Contact				
				A+	A	B	C	Total
E1:	Climate Change							
E2:	Deforestation							
E3:	Water Resource Management							
E4:	Biodiversity							
E5:	Waste & Pollution							
E6:	Resource & Energy Management							
E1-E6:	Subtotal	176	482	55	146	274	7	482
S0:	Diversity							
S1:	Human Rights							
S2:	Labour Practices / Health & Safety							
S3:	Product Liability & Safety							
S4:	Local Community							
S0-S4:	Subtotal	124	253	29	84	136	4	253
G0:	Board Governance & Accountability							
G1:	Capital Efficiency							
G2:	Takeover Defence Measures							
G3:	Risk Management							
G0~G3:	Subtotal	1,266	1,617	225	638	714	34	1,611
ESG1:	CSR/ESG Management							
ESG2:	Corporate Misconduct							
ESG3:	Regional Revitalization							
ESG4:	CSR Supply Chain Management							
ESG5	Digital Transformation (DX)							
ESG1~5	Subtotal	726	1,433	160	419	845	8	1,432
Total		2,292	3,785	469	1,287	1,969	53	3,778

Engagement Type	Number of Themes	Number of Contacts	A+	A	B	C	Total
Base	1,670	1,686	241	651	757	33	1,682
Follow-up (with milestones)	622	2,099	228	636	1,212	20	2,096
Total	2,292	3,785	469	1,287	1,969	53	3,778

Panel B – Number of Engagements per Company

Number of Engagements	Engagement Type		
	Any	Base	Follow-up
1	117	148	64
2	98	109	120
3	94	74	68
4	74	58	22
5	46	36	4
6	34	30	1
7	36	25	
8	18	9	
9	23	10	
10	12	4	
11	7	4	
12	2	2	
13	4		
14	2	1	
15	2	1	
16	1	1	
17	1		
Total	571	512	279

Panel C – Engagements per JPX Segment

JPX Segment	Total
TOPIX Core30	30
TOPIX Large70	68
TOPIX Mid400	264
TOPIX Small 1	127
TOPIX Small 2	75
TSE_2nd	1
Mothers	1
JASDAQ_S	1
REIT	4
Total	571

Table 2. AM One Engagements – Milestone Progress and Start Dates

Panel A shows a matrix of initial milestones and the highest milestone achieved by the end of the engagement or the end of the sample period. In more than 90% of all cases the issue raised by AM One was immediately recognised. In 50% of all cases plans had been implemented or the engagement was completed in the observation window. Singular engagements to not have milestones, by definition. Panels B and C show engagement start dates, completion status and the date of the last contact. Panel B reports the year and the month of the first contact. Panel C splits the sample into ongoing engagements and completed engagements. An ongoing engagement has a Milestone of 1-6; a completed engagement has a Milestone of 7 or 8. Year refers to the year of the last contact with the company when the engagement theme was discussed.

Panel A – Date of Base Engagements by Year and Month

Initial Milestone level		Highest Milestone level reached								Total	%
		1	2	3	4	5	6	7	8		
identifying ESG issue	1	19		3	1			3	1	27	4.3
raising concern/suggestions	2		3		1	3	4	12	2	25	4.0
issues recognised	3			34	21	13	21	34	8	131	21.1
issues rec. by senior management	4				62	35	28	63	18	206	33.1
initiatives taken	5					31	30	55	16	132	21.2
plans formulated	6						14	12	9	35	5.6
plans implemented	7							26	27	53	8.5
completing engagement	8								13	13	2.1
Total		19	3	37	85	82	97	205	94	622	
%		3.1	0.5	5.9	13.7	13.2	15.6	33.0	15.1		

Panel B – Start of Follow-Up Engagements by Year and Month

	Start Month (First Contact)												Total
Start Year	1	2	3	4	5	6	7	8	9	10	11	12	
2018				10	33	34	8	26	19	29	37	55	251
2019	23	27	22	13	12	10	4	8	5	12	11	8	155
2020	6	3	4	3	9	6	1	12	13	7	13	21	98
2021	16	12	21	2	4	3	1		7	8	6	21	101
2022	5	6	6										17
Total	50	48	53	28	58	53	14	46	44	56	67	105	622

Panel C – Engagement Completion Status and Date of Last Contact

	Ongoing Engagements						Completed Engagements (Milestone 7 or 8)						
	2018	2019	2020	2021	2022	Total	2018	2019	2020	2021	2022	Total	All
Start Year													
2018	30	20	8	28	26	112	6	17	11	62	43	139	251
2019		44	3	13	14	74		5	8	43	25	81	155
2020			6	39	11	56			2	22	18	42	98
2021				48	19	67				15	19	34	101
					14	14					3	3	17
Total	30	64	17	128	84	323	6	22	21	142	108	299	622

Table 3. TOPIX Index Segments and ESG Scores

The table shows the number of TOPIX index constituents over time for the last trading day in March, the last opportunity to adjust the portfolio before the end of the fiscal year. Panel A reports the number of constituents by subindex. Panel B reports the number of FTSE ESG scores available by subindex and Panel C the average FTSE score from the beginning of December of the previous year, the most recent FTSE score available. Panel D and Panel E report the number of MSCI scores and the average score from the beginning of March of the same year, the most recent MSCI score available.

	2015	2016	2017	2018	2019	2020	2021	2022
Panel A: Number of TOPIX Constituents by Subindex								
TOPIX Sub-Index								
TOPIX Core 30	30	30	30	30	30	30	29	30
TOPIX Large 70	70	70	70	70	71	70	70	70
TOPIX Midcap 400	402	403	399	401	399	399	397	400
TOPIX Small 1	499	498	496	498	495	500	493	496
TOPIX Small 2	857	936	1,002	1,062	1,129	1,160	1,197	1,179
TSE 2nd Section	538	538	529	512	491	478	471	470
Mothers	199	212	225	235	271	311	343	424
REIT	50	53	58	60	62	63	61	61
JASDAQ Standard	786	735	707	700	675	662	660	651
JASDAQ Growth	45	44	42	40	37	37	37	34
TOPIX Unassigned	18	10	11	19	13	5	1	1
Total	3,494	3,529	3,569	3,627	3,673	3,715	3,759	3,816
Panel B: Number of FTSE Scores								
TOPIX Sub-Index								
TOPIX Core 30	30	30	30	30	30	30	29	30
TOPIX Large 70	68	68	70	70	70	70	70	70
TOPIX Midcap 400	332	328	385	394	393	395	394	397
TOPIX Small 1	20	18	196	227	224	410	432	432
TOPIX Small 2	1	1	12	19	19	175	190	214
TSE 2nd Section			8	8	5	11	10	11
REIT								40
Total	451	445	701	748	741	1,091	1,125	1,194
Panel C: Number of MSCI Scores								
TOPIX Sub-Index								
TOPIX Core 30	30	30	30	30	30	30	29	30
TOPIX Large 70	70	70	70	70	71	70	70	70
TOPIX Midcap 400	206	210	372	386	394	384	395	398
TOPIX Small 1	18	19	67	153	183	181	261	427
TOPIX Small 2	23	22	33	34	44	25	56	163
TSE 2nd Section	16	13	18	21	19	5	3	7
Mothers	3	1	3	4	4	6	9	31
REIT	8	8	35	38	41	41	43	52
JASDAQ Standard	9	8	16	21	19	7	7	17
TOPIX Unassigned					1			
Total	383	381	644	757	806	749	873	1,195
Panel D: FTSE Coverage								
TOPIX Sub-Index								
TOPIX Core30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
TOPIX Large70	0.97	0.97	1.00	1.00	0.99	1.00	1.00	1.00
TOPIX Mid400	0.83	0.81	0.96	0.98	0.98	0.99	0.99	0.99
TOPIX Small 1	0.04	0.04	0.40	0.46	0.45	0.82	0.88	0.87
TOPIX Small 2	0.00	0.00	0.01	0.02	0.02	0.15	0.16	0.18
Panel E: MSCI Coverage								
TOPIX Sub-Index								
TOPIX Core30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
TOPIX Large70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
TOPIX Mid400	0.51	0.52	0.93	0.96	0.99	0.96	0.99	1.00
TOPIX Small 1	0.04	0.04	0.14	0.31	0.37	0.36	0.53	0.86
TOPIX Small 2	0.03	0.02	0.03	0.03	0.04	0.02	0.05	0.14
Panel F: Average FTSE Score								
TOPIX Sub-Index								
TOPIX Core30	2.6	2.6	2.8	3.1	3.0	3.3	3.4	3.5
TOPIX Large70	2.3	2.3	2.6	2.8	2.9	3.3	3.2	3.3
TOPIX Mid400	1.6	1.7	1.8	2.0	2.0	2.3	2.4	2.5
TOPIX Small 1	1.3	1.4	1.3	1.5	1.3	1.6	1.6	1.7
TOPIX Small 2	0.8	0.8	1.1	1.2	1.1	1.4	1.5	1.5
Panel G: Average MSCI Score								
TOPIX Sub-Index								
TOPIX Core30	2.4	2.4	2.4	2.5	2.5	2.6	2.5	2.7
TOPIX Large70	2.4	2.5	2.5	2.5	2.5	2.6	2.5	2.6
TOPIX Mid400	2.3	2.3	2.3	2.3	2.3	2.4	2.3	2.4
TOPIX Small 1	2.3	2.2	2.2	2.3	2.2	2.3	2.1	2.1
TOPIX Small 2	2.3	2.4	2.4	2.4	2.3	2.0	2.0	1.9

Table 4. Engagements and ESG Scores

The table shows the number of companies engaged by engagement category and the respective ESG score at the start of the engagement. E, S and G engagements were matched with the respective FTSE and MSCI pillar scores; ESG engagements were matched with the aggregate score. In the temporal dimension, the start of the engagement was matched with the most recent available score; for FTSE the beginning of June and December; for MSCI the beginning of each month. As a reference, an ESG score of 3.1 was required for inclusion in the Blossom index until December 2019 and a score of 3.3 thereafter. For MSCI the required score for inclusion BB or above; to avoid exclusion the score had to remain above CCC. A MSCI score of 2 in the table would translate to a BB; a score of 2.5 to a BBB and a score of 3 into an A; a 0.5 would be CCC.

Year	Engagements	FTSE			MSCI		
	Count	Scores	Mean	Median	Scores	Mean	Median
Environmental Engagements and Scores							
2018	39	39	2.9	3.0	39	2.8	2.8
2019	43	43	3.2	3.3	42	2.7	2.7
2020	21	21	3.6	3.9	21	2.7	2.7
2021	52	51	3.0	3.0	52	2.6	2.7
2022	21	21	2.9	3.0	21	2.7	2.9
Total	*176*	*175*	*3.1*	*3.2*	*175*	*2.7*	*2.8*
Social Engagements and Scores							
2018	17	17	1.2	0.9	17	2.8	2.8
2019	19	19	2.9	3.0	19	2.4	2.4
2020	35	35	2.8	3.1	34	2.5	2.5
2021	36	36	2.6	2.8	36	2.6	2.5
2022	17	17	3.2	3.4	17	2.6	2.7
Total	*124*	*124*	*2.6*	*2.8*	*123*	*2.6*	*2.6*
Governance Engagements and Scores							
2018	108	90	2.5	2.5	90	1.9	2.0
2019	285	231	2.8	2.9	229	2.0	2.1
2020	335	304	3.0	3.0	255	2.1	2.2
2021	353	306	3.1	3.2	289	2.0	2.0
2022	185	165	3.2	3.3	158	2.1	2.1
Total	*1,266*	*1,096*	*3.0*	*3.0*	*1,021*	*2.0*	*2.1*
ESG Engagements and Scores							
2018	151	146	2.0	2.0	144	2.4	2.4
2019	159	139	2.7	2.8	137	2.4	2.4
2020	189	182	2.9	3.1	169	2.5	2.5
2021	170	164	3.0	3.3	164	2.4	2.5
2022	57	55	3.0	3.3	54	2.5	2.5
Total	*726*	*686*	*2.7*	*2.9*	*668*	*2.4*	*2.5*
Total	**2,292**	**2,081**			**1,987**		

Table 5. Two-Period Difference-in-Differences Regressions

The table shows difference-in-differences estimates for companies that were TOPIX500 constituents (Core30, Large70, Mid400) at any one point over the sample period. TOPIX500 companies were AM One's main engagement. In Column 1 treated is defined as any engagement by AM One in the period 2018-2022. In Columns 2 and 6 treatment is confined to E engagements; in Columns 3 and 7 to S engagements; in Columns 4 and 8 to G engagements and in Column 5 to horizontal ESG engagements as defined by AM One. The dependent variables match the treatment. MSCI were rescaled from 0-10 to 0-5 to match the FTSE scores; MSCI materiality weights range from 0 to 100. FTSE scores are measured on a scale from zero to five (0 no disclosure, 3 good practice, 5 best practice); exposure scores are measured on a scale from zero to three ("negligible" or not available, low, medium, high). FTSE scores are for the period January 2015 to March 2022; for MSCI the sample is confined to December 2016 to October 2020, due to insufficient data coverage in the preceding quarters and a disruptive change in methodology in November 2020. Means and Standard Errors are estimated by linear regression with robust standard errors. Inference: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. Estimation was carried out with the Stata diff module Villa (2016).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	ESG Score	E Score	S Score	G Score	ESG Score	E Weight	S Weight	G Weight
Treatment	Any	E	S	G	ESG	E	S	G
Panel A: FTSE								
Diff-in-Diff	0.21***	0.29***	0.09	0.01	0.16***	-0.02	0.01	0.00
	(0.049)	(0.064)	(0.074)	(0.043)	(0.046)	(0.029)	(0.027)	(0.010)
Observations	6,609	6,609	6,609	6,609	6,609	6,609	6,609	6,609
R-squared	0.10	0.10	0.07	0.12	0.09	0.09	0.01	0.01
Mean control t(0)	1.81	1.99	1.61	2.20	1.89	2.20	2.43	2.18
Mean treated t(0)	2.11	2.70	1.87	2.36	2.12	2.62	2.51	2.21
Diff t(0)	0.30	0.71	0.25	0.16	0.23	0.42	0.07	0.03
Mean control t(1)	2.14	2.16	2.17	2.81	2.29	2.32	2.47	2.17
Mean treated t(1)	2.65	3.15	2.50	2.98	2.67	2.72	2.55	2.21
Diff t(1)	0.50	0.99	0.34	0.17	0.39	0.39	0.08	0.03
Panel B: MSCI								
Diff-in-Diff	0.02	-0.08	-0.00	0.10***	0.03	-0.56	-0.82	0.64
	(0.028)	(0.064)	(0.064)	(0.035)	(0.026)	(1.344)	(1.218)	(0.804)
Observations	4,241	4,241	4,241	4,241	4,241	4,241	4,241	4,241
R-squared	0.02	0.01	0.01	0.00	0.01	0.10	0.01	0.01
Mean control t(0)	2.25	2.48	2.40	2.04	2.30	26.07	41.93	27.99
Mean treated t(0)	2.35	2.73	2.45	1.99	2.34	39.76	46.26	27.44
Diff t(0)	0.10	0.26	0.05	-0.06	0.04	13.69	4.34	-0.55
Mean control t(1)	2.28	2.56	2.53	1.99	2.33	23.80	42.13	30.32
Mean treated t(1)	2.41	2.74	2.57	2.03	2.40	36.93	45.65	30.41
Diff t(1)	0.13	0.18	0.04	0.04	0.08	13.13	3.52	0.09

Table 6. Two-Period Difference-in-Differences Quantile Regressions

The table shows difference-in-differences using quantile regressions with bootstrap standard errors. The sample, score and treatment definitions are the same as in the linear regressions table. The table shows for which values of the ESG scores (very low, low, median, high, very high) the treatment had the largest effect, if any. Inference: *** p<0.01; ** p<0.05; * p<0.1. Estimation was carried out with the Stata `diff` module and the `qdid` option that calls on `qreg` and `bsqreg` (Villa, 2016).

Panel 1. ESG Scores, Any Engagement

| | FTSE | | | | | MSCI | | | | |
| | (1) | (2) | (3) | (4) | (5) | (1) | (2) | (3) | (4) | (5) |
Percentile	10th	25th	Median	75th	90th	10th	25th	Median	75th	90th
Diff-in-diff	0.40***	0.30	0.30	0.00	0.00	0.05	0.05	0.05	0.00	0.10
S.Err.	(0.130)	(0.197)	(0.228)	(0.197)	(0.129)	(0.068)	(0.099)	(0.117)	(0.100)	(0.070)
Observations	6,609	6,609	6,609	6,609	6,609	4,241	4,241	4,241	4,241	4,241
Mean control t(0)	0.90	1.20	1.70	2.30	3.00	1.75	2.00	2.25	2.50	2.75
Mean treated t(0)	0.90	1.40	2.10	2.80	3.30	1.80	2.05	2.35	2.60	2.85
Diff t(0)	0.00	0.20	0.40	0.50	0.30	0.05	0.05	0.10	0.10	0.10
Mean control t(1)	0.90	1.40	2.00	2.90	3.60	1.80	2.05	2.25	2.55	2.75
Mean treated t(1)	1.30	1.90	2.70	3.40	3.90	1.90	2.15	2.40	2.65	2.95
Diff t(1)	0.40	0.50	0.70	0.50	0.30	0.10	0.10	0.15	0.10	0.20

Panel 2. E Scores, E Engagement

| | FTSE | | | | | MSCI | | | | |
| | (1) | (2) | (3) | (4) | (5) | (1) | (2) | (3) | (4) | (5) |
Percentile	10th	25th	Median	75th	90th	10th	25th	Median	75th	90th
Diff-in-diff	0.50***	0.50**	0.30	0.20	-0.10	-0.00	-0.15	-0.05	-0.05	-0.15
S.Err.	(0.136)	(0.196)	(0.230)	(0.202)	(0.137)	(0.070)	(0.106)	(0.122)	(0.106)	(0.178)
Observations	6,609	6,609	6,609	6,609	6,609	4,241	4,241	4,241	4,241	4,241
Mean control t(0)	0.30	1.00	2.00	2.90	3.50	1.15	1.65	2.55	3.20	3.70
Mean treated t(0)	1.20	2.00	2.70	3.40	4.10	1.70	2.30	2.75	3.20	3.70
Diff t(0)	0.90	1.00	0.70	0.50	0.60	0.55	0.65	0.20	0.00	0.00
Mean control t(1)	0.30	1.00	2.30	3.20	3.90	1.25	1.80	2.55	3.25	3.85
Mean treated t(1)	1.70	2.50	3.30	3.90	4.40	1.80	2.30	2.70	3.20	3.70
Diff t(1)	1.40	1.50	1.00	0.70	0.50	0.55	0.50	0.15	-0.05	-0.15

Panel 3. S Scores, S Engagement

| | FTSE | | | | | MSCI | | | | |
| | (1) | (2) | (3) | (4) | (5) | (1) | (2) | (3) | (4) | (5) |
Percentile	10th	25th	Median	75th	90th	10th	25th	Median	75th	90th
Diff-in-diff	0.20	0.20	0.10	0.10	-0.20	0.10	0.15	0.05	-0.00	-0.05
S.Err.	(0.151)	(0.224)	(0.259)	(0.221)	(0.148)	(0.124)	(0.119)	(0.138)	(0.119)	(0.165)
Observations	6,609	6,609	6,609	6,609	6,609	4,241	4,241	4,241	4,241	4,241
Mean control t(0)	0.30	0.80	1.50	2.40	3.10	1.55	2.00	2.40	2.85	3.30
Mean treated t(0)	0.40	0.90	1.70	2.90	3.50	1.50	1.90	2.40	2.90	3.50
Diff t(0)	0.10	0.10	0.20	0.50	0.40	-0.05	-0.10	0.00	0.05	0.20
Mean control t(1)	0.50	1.20	2.30	3.00	3.80	1.65	2.05	2.50	2.95	3.45
Mean treated t(1)	0.80	1.50	2.60	3.60	4.00	1.70	2.10	2.55	3.00	3.60
Diff t(1)	0.30	0.30	0.30	0.60	0.20	0.05	0.05	0.05	0.05	0.15

Table 6 – Two-Period Difference-in-Differences Quantile Regressions continued …

Panel 4. G Scores, G Engagement

	FTSE					MSCI				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Percentile	10th	25th	Median	75th	90th	10th	25th	Median	75th	90th
Diff-in-diff	0.10	0.10	-0.20	0.10	-0.20*	0.15**	0.05	0.10	-0.00	0.10
S.Err.	(0.120)	(0.166)	(0.203)	(0.160)	(0.115)	(0.063)	(0.092)	(0.108)	(0.094)	(0.063)
Observations	6,609	6,609	6,609	6,609	6,609	4,241	4,241	4,241	4,241	4,241
Mean control t(0)	1.10	1.50	2.00	3.00	3.30	1.40	1.70	2.05	2.35	2.70
Mean treated t(0)	1.30	1.70	2.30	3.00	3.50	1.30	1.70	1.95	2.30	2.65
Diff t(0)	0.20	0.20	0.30	0.00	0.20	-0.10	0.00	-0.10	-0.05	-0.05
Mean control t(1)	1.70	2.00	2.90	3.50	4.00	1.30	1.65	2.05	2.40	2.65
Mean treated t(1)	2.00	2.30	3.00	3.60	4.00	1.35	1.70	2.05	2.35	2.70
Diff t(1)	0.30	0.30	0.10	0.10	0.00	0.05	0.05	-0.00	-0.05	0.05

Panel 5. ESG Scores, ESG Engagement

	FTSE					MSCI				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Percentile	10th	25th	Median	75th	90th	10th	25th	Median	75th	90th
Diff-in-diff	0.30***	0.30*	0.30	0.10	0.10	0.05	0.05	0.05	0.00	0.10*
S.Err.	(0.116)	(0.178)	(0.206)	(0.177)	(0.118)	(0.063)	(0.089)	(0.106)	(0.091)	(0.061)
Observations	6,609	6,609	6,609	6,609	6,609	4,241	4,241	4,241	4,241	4,241
Mean control t(0)	0.90	1.30	1.80	2.50	3.10	1.80	2.00	2.30	2.55	2.90
Mean treated t(0)	0.90	1.40	2.10	2.80	3.30	1.80	2.05	2.35	2.60	2.85
Diff t(0)	0.00	0.10	0.30	0.30	0.20	0.00	0.05	0.05	0.05	-0.05
Mean control t(1)	1.00	1.50	2.20	3.10	3.60	1.85	2.05	2.30	2.60	2.85
Mean treated t(1)	1.30	1.90	2.80	3.50	3.90	1.90	2.15	2.40	2.65	2.90
Diff t(1)	0.30	0.40	0.60	0.40	0.30	0.05	0.10	0.10	0.05	0.05

Table 7. Multiple Period Engagement Cohorts

The table shows frequency counts for the roll-out of AM One's engagement programme. The table is confined to TOPIX 500 companies where FTSE scores are available at the end of each year. The panel is unbalanced with a total of 3,303 firm year observations. Panel 1 shows the engagement cohorts by year. The first engagement cohort is 2018 when 171 sample companies were engaged on E, S, G or ESG; 163 of these companies were already in the TOPIX 500 and had a FTSE score in 2015. Panel 2 shows the engagements in event time where the year of the first engagement is set to zero. For the 2018 cohort the longest post-treatment observation period is 3 (2019,2020,2021). The year 2022 is not considered because the last outcome measurement is in December 2021. In contrast, there are companies that were engaged in 2022 and event time is set to zero; for these companies it is possible to have seven pre-treatment periods.

Panel 1. Engagement Roll-Out by Cohort

| Treatment Type | First Engaged | Number of Firms | | | | | | |
		2015	2016	2017	2018	2019	2020	2021
E,S,G or ESG	Never	111	128	133	137	138	135	137
	2018	163	167	169	171	166	161	158
	2019	103	105	108	111	108	100	97
	2020	36	40	40	41	41	45	43
	2021	26	27	30	30	30	30	27
	2022	2	2	2	2	1	1	1
	Not Yet				184	72	31	1
E	Never	323	350	361	370	363	354	348
	2018	37	37	38	38	37	37	36
	2019	29	29	30	31	31	29	29
	2020	15	15	15	15	15	15	14
	2021	26	27	27	27	27	26	25
	2022	11	11	11	11	11	11	11
	Not Yet				84	53	37	11
S	Never	360	386	397	404	396	385	376
	2018	12	12	13	14	14	14	14
	2019	14	15	15	16	16	15	15
	2020	25	26	26	27	27	27	27
	2021	21	21	22	22	22	22	22
	2022	9	9	9	9	9	9	9
	Not Yet				74	58	31	9
G	Never	165	186	193	198	197	191	190
	2018	67	67	67	68	65	62	62
	2019	108	111	114	116	113	107	103
	2020	55	59	59	60	60	64	63
	2021	34	34	37	38	38	37	34
	2022	12	12	12	12	11	11	11
	Not Yet				226	109	48	11
ESG	Never	191	208	215	220	218	215	214
	2018	107	111	113	114	111	107	104
	2019	68	71	73	74	71	67	65
	2020	45	48	48	49	49	49	47
	2021	24	25	27	28	28	27	26
	2022	6	6	6	7	7	7	7
	Not Yet				158	84	34	7

Panel 2. Engagements in Event Time

| Treatment Type | Event Time | | | | | | | | | | |
	-7	-6	-5	-4	-3	-2	-1	0	1	2	3
E,S,G or ESG	2	28	65	175	339	347	352	351	309	258	158
E	11	37	53	82	119	120	121	109	80	66	36
S	9	30	55	71	84	85	87	79	56	29	14
G	12	46	101	216	286	290	291	279	235	165	62
ESG	6	30	76	150	261	268	270	260	225	172	104

Table 8. Multiple Period Difference in Differences Estimates

The table shows difference in differences estimates with multiple time periods (Callaway and Sant'Anna, 2021). Panel 1 shows dynamic average treatment effects for the treated (ATTs) in event time. The ATT(g,s) are estimated using all periods relative to the period of the first treatment, across all cohorts using `csdid` and `estat event` (Rios-Avila et al., 2021). In Panel 2 columns 1-5 show ATT estimates by treatment group. Column 1 shows the average across all groups; columns 2-5 for the 2018, 2019, 2020 and 2021 cohorts (`estat group`). Column 6 shows average ATT estimates for all groups across all periods (`estat simple`). Column 7 reports the coefficient of the interaction term in a static two-way fixed effects regression (TWFE). Coefficient and standard errors are not fully identical to Table 5 because the estimates rely on December scores, while the previous results used bi-annual data for June and December.

Panel 1. Dynamic Average Treatment Effects for the Treated (ATT's)

	T-4	T-3	T-2	T-1	T=0	T+1	T+2	T+3
E,S,G or ESG								
Coeff.	-0.01	0.02	-0.04	-0.01	0.10	0.17	0.17	0.16
Std. err.	0.05	0.04	0.03	0.03	0.03	0.05	0.07	0.10
z	-0.29	0.44	-1.11	-0.32	3.02	3.15	2.42	1.69
P>|z|	0.77	0.66	0.27	0.75	0.00	0.00	0.02	0.09
E								
Coeff.	0.25	-0.01	0.08	0.07	-0.05	0.18	0.23	0.22
Std. err.	0.11	0.07	0.06	0.05	0.07	0.09	0.11	0.16
z	2.35	-0.14	1.44	1.43	-0.75	2.02	2.21	1.37
P>|z|	0.02	0.89	0.15	0.15	0.45	0.04	0.03	0.17
S								
Coeff.	0.04	0.06	0.05	0.02	0.04	-0.06	-0.11	-0.45
Std. err.	0.09	0.08	0.07	0.06	0.05	0.08	0.13	0.20
z	0.42	0.73	0.72	0.39	0.79	-0.75	-0.89	-2.22
P>|z|	0.67	0.47	0.47	0.70	0.43	0.45	0.37	0.03
G								
Coeff.	0.02	0.07	-0.08	0.01	0.05	0.01	0.03	-0.07
Std. err.	0.07	0.05	0.04	0.04	0.04	0.06	0.09	0.14
z	0.26	1.48	-2.11	0.25	1.29	0.15	0.34	-0.47
P>|z|	0.79	0.14	0.04	0.80	0.20	0.88	0.73	0.64
ESG								
Coeff.	0.04	0.00	-0.03	-0.02	0.09	0.12	0.11	0.11
Std. err.	0.04	0.04	0.03	0.03	0.03	0.05	0.07	0.10
z	0.92	0.11	-0.84	-0.55	2.72	2.32	1.63	1.10
P>|z|	0.36	0.91	0.40	0.59	0.01	0.02	0.10	0.27

Panel 2. Group ATTs, Aggregate ATTs, Fixed Effect Regression Estimates

	G Aver.	G2018	G2019	G2020	G2021	ATT(g,s)	β TWFE
E,S,G or ESG							
Coeff.	0.14	0.15	0.15	0.08	0.07	0.14	0.20
Std. err.	0.05	0.06	0.07	0.09	0.10	0.05	0.07
z	2.89	2.42	2.18	0.85	0.71	2.86	2.89
P>|z|	0.00	0.02	0.03	0.40	0.48	0.00	0.00
E							
Coeff.	0.02	0.16	0.22	-0.10	-0.35	0.11	0.23
Std. err.	0.07	0.13	0.09	0.15	0.18	0.08	0.09
z	0.31	1.30	2.38	-0.67	-1.89	1.44	2.48
P>|z|	0.76	0.19	0.02	0.51	0.06	0.15	0.01
S							
Coeff.	-0.03	-0.13	0.00	-0.08	0.08	-0.05	0.07
Std. err.	0.06	0.17	0.10	0.09	0.10	0.07	0.11
z	-0.46	-0.72	0.01	-0.91	0.83	-0.76	0.68
P>|z|	0.64	0.47	0.99	0.36	0.41	0.44	0.50
G							
Coeff.	0.02	0.01	0.06	-0.06	0.07	0.02	-0.01
Std. err.	0.05	0.10	0.07	0.08	0.09	0.06	0.06
z	0.43	0.09	0.83	-0.74	0.76	0.41	0.17
P>|z|	0.67	0.93	0.40	0.46	0.45	0.68	0.86
ESG							
Coeff.	0.10	0.08	0.20	0.01	0.09	0.10	0.17
Std. err.	0.04	0.07	0.07	0.08	0.10	0.05	0.07
z	2.34	1.25	2.70	0.10	0.90	2.21	2.57
P>|z|	0.02	0.21	0.01	0.92	0.37	0.03	0.01

Table 9. Seven & i Holdings: Theme Level Engagements, ESG Scores and Risk

The table shows the evolution of the theme level, pillar level and over ESG scores for a major retail company. Panel 1 shows FTSE data; Panel 2 shows MSCI data. Not that AM One's ESG4 CSR Supply Chain Management is not included in the scoring of MSCI ESG scores.

Panel 1. FTSE ESG Theme Scores and Exposure Levels

		Mar-19	Jun-19	Dec-19	Dec-20	Dec-21
ESG Rating		3.2	3.8	3.8	3.6	4.2
Environment Score		3.3	4	4	3.6	3.7
Environment Exposure		2.3	2.3	2.3	2.3	2.3
Climate Change Score		4	5	5	5	3
Climate Change Exposure		M	M	M	M	M
Environmental Supply Chain Score		3	4	4	3	4
Environmental Supply Chain Exposure		H	H	H	H	H
Pollution & Resources Score		3	3	3	3	4
Pollution & Resources Exposure		M	M	M	M	M
Water Security Score						
Water Security Exposure		NA	NA	NA	NA	NA
Social		3.2	3.7	3.7	3.3	4.5
		2.8	2.8	2.8	2.8	2.8
Customer Responsibility Score		3	4	4	3	5
Customer Responsibility Weight		H	H	H	H	H
Human Rights & Community Score		4	4	4	3	4
Human Rights & Community Weight		M	M	M	M	M
Labour Standards Score		3	3	3	3	4
Labour Standards Exposure		H	H	H	H	H
Social Supply Chain Score		3	4	4	4	5
Social Supply Chain Exposure		H	H	H	H	H

Engagement Agenda Item		Date				
E5	Waste & Pollution			Jul-19		Mar-21
S2	Labor Practices / Health & Safety			Nov-19		Mar-21
ESG4	CSR Supply Chain Management			Jan-19		Mar-21
		Milestone				
E5	Waste & Pollution			5		7
S2	Labor Practices / Health & Safety			7		7
ESG4	CSR Supply Chain Management			4		7

Panel 2. MSCI ESG Theme Scores and Materiality Weights

			Jan-19	Jun-19	Dec-19	Dec-20	Dec-21
Overall		IVA Company Rating	BBB	BBB	BBB	BB	BBB
		Industry Adjusted Score	5.0	5.0	4.8	3.5	5.4
		Weighted Average Score	4.4	4.2	4.4	4.6	4.5
Pillar	E	Environmental Pillar Score	3.7	3.7	3.8	3.8	3.4
		Environmental Pillar Weight	18.0	18.0	20.0	15.0	15.0
	S	Social Pillar Score	4.3	4.0	4.3	4.9	4.1
		Social Pillar Weight	63.0	63.0	60.0	52.0	52.0
	G	Governance Pillar Score	5.6	5.5	5.4	4.5	5.6
		Governance Pillar Weight	19.0	19.0	20.0	33.0	33.0
Theme	E	Climate Change Theme Score	4.9	4.9	4.5	4.5	4.0
		Climate Change Theme Weight	9.0	9.0	10.0	8.0	8.0
	E	Natural Res Use Theme Score	2.4	2.4	3.1	3.1	2.7
		Natural Res Use Theme Weight	9.0	9.0	10.0	7.0	7.0
	E	Waste Mgmt Theme Score					
		Waste Mgmt Theme Weight	0.0	0.0			
	E	Environmental Opps Theme Score					
		Environmental Opps Theme Weight					
	S	Human Capital Theme Score	3.8	3.8	4.6	2.3	1.5
		Human Capital Theme Weight	21.0	21.0	20.0	15.0	15.0
	S	Product Safety Theme Score	5.1	4.3	4.3	6.4	5.5
		Product Safety Theme Weight	21.0	21.0	20.0	30.0	30.0
	S	Social Opps Theme Score	3.9	3.9	3.9	3.9	3.9
		Social Opps Theme Weight	21.0	21.0	20.0	7.0	7.0

Engagement Agenda Item		Date				
E5	Waste & Pollution			Jul-19		Mar-21
S2	Labor Practices / Health & Safety			Nov-19		Mar-21
ESG4	CSR Supply Chain Management			Jan-19		Mar-21
		Milestone				
E5	Waste & Pollution			5		7
S2	Labor Practices / Health & Safety			7		7
ESG4	CSR Supply Chain Management			4		7

Table 10. Large Utility Company: FTSE ESG Theme Scores and Exposure Levels

The table shows an E engagement with a large utility company at the theme level. AM One met several times with the company to engage on climate change, with a particular focus on disclosure. Significant milestone progress was recorded. At the outset of the engaged in 2018 FTSE considered the company to be highly exposed to climate risk; MSCI assigned a 20% materiality weight. FTSE did not change its exposure assessment; MSCI reduced its weight from 20 percent to 14 percent.

Panel 1. FTSE ESG Theme Scores and Exposure Levels

	Dec-17	Dec-18	Dec-19	Dec-20	Dec-21
ESG Rating	1.8	2.5	3.0	3.5	3.2
Environment Score	2.4	2.8	3.5	4.0	3.5
Environment Exposure	2.8	2.8	2.8	2.8	2.8
Biodiversity Score	2	3	3	4	3
Biodiversity Exposure	M	M	M	M	M
Climate Change Score	3	3	5	5	5
Climate Change Exposure	H	H	H	H	H
Environmental Supply Chain	0	2			
Environmental Supply Exposure	H	H	NA	NA	NA
Pollution & Resources Chain	4	3	4	4	4
Pollution & Resources Exposure	H	H	H	H	H
Water Security Score	3	3	2	3	2
Water Security Exposure	H	H	H	H	H

Engagement Agenda Item	Date				
E1 Climate Change		Jun-18	Sep-19	Jun-20	Mar-21
	Milestone				
		5	7	7	7

Panel 2. MSCI ESG Theme Scores and Materiality Weights

			Jan-18	Jul-18	Jan-19	Jan-20	Jan-21	Jul-21
Overall		IVA Company Rating	BBB	BBB	BBB	BBB	BBB	BBB
		Industry Adjusted Score	4.80	4.80	4.60	4.70	4.60	4.60
		Weighted Average Score	5.10	4.90	5.00	5.00	4.90	4.90
Pillar	E	Environmental Pillar Score	5.80	5.50	5.80	5.60	6.00	6.00
		Environmental Pillar Weight	74	74	69	64	54	54
	S	Social Pillar Score	3.20	3.20	3.80	3.80	3.80	3.80
		Social Pillar Weight	10	10	15	16	13	13
	G	Governance Pillar Score	3.00	3.30	3.10	4.00	3.40	3.40
		Governance Pillar Weight	16	16	16	20	33	33
Theme	E	Climate Change Theme Score	6.90	5.60	5.50	5.60	5.70	5.70
		Climate Change Theme Weight	20	20	18	16	14	14
	E	Natural Res Use Theme Score	5.70	5.70	6.30	5.50	6.40	6.40
		Natural Res Use Theme Weight	17	17	18	16	14	14
	E	Waste Mgmt Theme Score	7.70	7.70	6.80	6.90	6.80	6.80
		Waste Mgmt Theme Weight	20	20	18	16	13	13
	E	Environmental Opps Theme Score	2.50	2.50	4.20	4.30	5.10	5.10
		Environmental Opps Theme Weight	17	17	15	16	13	13

Engagement Agenda Item		Date					
E1 Climate Change			Jun-18			Jun-20	Mar-21
		Milestone					
E1 Climate Change			5			7	7

Table 11. FTSE Blossom Index Inclusion Returns

The table shows event study returns for inclusions and exclusions in the FTSE Blossom ESG index and the FTSE4Good index. The event window starts 15 days before the effective inclusion/exclusion date and spans to five trading days after. Public announcements of index inclusions and exclusions are made nine business days before the effective date. Each panel reports raw returns, and cumulative average abnormal returns from a constant mean return model, a market adjusted model and a (one factor) market model using the TOPIX 500 index as the benchmark. The estimation window for the market model was [-280,-30]. Panel A1 reports inclusions from the TOPIX All Shares Index; Panel A2 from the TOPIX 100, Panel A3 from the TOPIX Mid 400 and Panel A4 from the TOPIX Small Caps index (Small 1 or Small 2). Panel 5 shows FTSE4Good inclusions from 2009 to 2016, the period prior to GPIF's ESG Index investment. Panel B1 shows Blossom exclusions and Panel B2 FTSE4Good exclusions. ESG inclusions/exclusions that are preceded by a change in the TOPIX segment within 50 days prior to the effective date were dropped, as were exclusions due to delisting. The estimation was carried out with the Stata `eventstudy2` module (Kaspereit, 2022). RAW denotes raw returns; COMEAN constant mean adjusted returns; MA market adjusted returns, and MM the market model. Inference: *** p<0.01; ** p<0.05; * p<0.1.

PANEL A.: INCLUSIONS						
MODEL	Window	Events	CAAR	T-Test	P-Value	Inference
PANEL A1.: BLOSSOM 2017-2022						
RAW	[-15;5]	103	0.031	3.17	0.002	***
COMEAN	[-15;5]	103	0.033	3.39	0.001	***
MA	[-15;5]	103	0.016	1.95	0.052	*
MM	[-15;5]	103	0.019	2.43	0.016	**
PANEL A2.: FROM TOPIX 100						
RAW	[-15;5]	28	0.021	1.33	0.186	
COMEAN	[-15;5]	28	0.019	1.19	0.236	
MA	[-15;5]	28	0.008	0.64	0.521	
MM	[-15;5]	28	0.009	0.72	0.475	
PANEL A3.: FROM TOPIX MID 400						
RAW	[-15;5]	57	0.022	1.66	0.098	*
COMEAN	[-15;5]	57	0.025	1.93	0.055	*
MA	[-15;5]	57	0.013	1.13	0.261	
MM	[-15;5]	57	0.018	1.60	0.110	
PANEL A4.: FROM TOPIX SMALL						
RAW	[-15;5]	18	0.079	2.83	0.005	***
COMEAN	[-15;5]	18	0.087	3.10	0.002	***
MA	[-15;5]	18	0.046	1.90	0.059	*
MM	[-15;5]	18	0.052	2.18	0.030	**
PANEL A5.: FTSE4GOOD 2009-2016						
RAW	[-15;5]	50	-0.048	-2.65	0.009	***
COMEAN	[-15;5]	50	-0.041	-2.27	0.024	**
MA	[-15;5]	50	-0.004	-0.26	0.794	
MM	[-15;5]	50	-0.005	-0.36	0.716	

PANEL B.: EXCLUSIONS						
MODEL	Window	Events	CAAR	T-Test	P-Value	Inference
PANEL B1.: BLOSSOM 2017-2022						
RAW	[-15;5]	19	-0.044	-1.86	0.065	*
COMEAN	[-15;5]	18	-0.046	-1.93	0.055	*
MA	[-15;5]	20	-0.040	-1.91	0.057	*
MM	[-15;5]	20	-0.030	-1.51	0.131	
Panel B2.: FTSE4Good 2009-2016						
RAW	[-15;5]	73	-0.048	-3.07	0.002	***
COMEAN	[-15;5]	73	-0.036	-2.33	0.021	**
MA	[-15;5]	73	-0.008	-0.59	0.559	
MM	[-15;5]	73	-0.012	-0.99	0.324	

Table 12. MSCI ESG Leaders Index Inclusion and Exclusion Returns

The table shows event study returns for inclusions in and exclusions from the MSCI ESG Leaders Index that GPIF adopted in 2017. The event window spans 15 days before the effective inclusion/exclusion date and five days after the inclusion/exclusion. The estimation window for the market model was [-280,-30]. Panel A1 reports inclusions from the TOPIX All Shares Index; Panel A2 from the TOPIX 100, Panel A3 from the TOPIX Mid 400 and Panel A4 from the TOPIX Small Caps index (Small 1 or Small 2). Panel B reports exclusions. The event study does not include cases that involve the simultaneous inclusion in or exclusion from any one of the TOPIX indexes. All other parameters were described in the header of the previous table. Inference: *** p<0.01; ** p<0.05; * p<0.1.

PANEL A.: INCLUSIONS						
MODEL	Window	Events	CAAR	T-Test	P-Value	Inference
PANEL A1.: ALL INCLUSIONS						
RAW	[-15;5]	81	0.018	1.77	0.078	*
COMEAN	[-15;5]	81	0.008	0.80	0.425	
MA	[-15;5]	81	0.026	2.85	0.005	***
MM	[-15;5]	81	0.021	2.37	0.019	**
PANEL A2.: FROM TOPIX 100						
RAW	[-15;5]	19	0.004	0.24	0.813	
COMEAN	[-15;5]	19	0.001	0.04	0.965	
MA	[-15;5]	19	0.011	0.66	0.512	
MM	[-15;5]	19	0.016	1.01	0.313	
PANEL A3.: FROM TOPIX MID CAP 400						
RAW	[-15;5]	35	0.029	1.99	0.048	**
COMEAN	[-15;5]	35	0.026	1.79	0.075	*
MA	[-15;5]	35	0.040	3.09	0.002	***
MM	[-15;5]	35	0.033	2.60	0.010	***
PANEL A4.: FROM TOPIX SMALL-CAP						
RAW	[-15;5]	27	-0.007	-0.38	0.703	
COMEAN	[-15;5]	27	-0.015	-0.79	0.429	
MA	[-15;5]	27	0.025	1.41	0.160	
MM	[-15;5]	27	0.011	0.62	0.533	

PANEL B.: EXCLUSIONS						
MODEL	Window	Events	CAAR	T-Test	P-Value	Inference
PANEL B1.: ALL EXCLUSIONS						
RAW	[-15;5]	100	-0.023	-2.62	0.009	***
COMEAN	[-15;5]	100	-0.019	-2.19	0.030	**
MA	[-15;5]	100	-0.023	-2.95	0.003	***
MM	[-15;5]	100	-0.018	-2.43	0.016	**
PANEL B2.: FROM TOPIX 100						
RAW	[-15;5]	12	-0.034	-1.61	0.109	
COMEAN	[-15;5]	12	-0.039	-1.82	0.070	*
MA	[-15;5]	12	-0.025	-1.44	0.150	
MM	[-15;5]	12	-0.025	-1.50	0.136	
PANEL B3.: FROM TOPIX MID CAP 400						
RAW	[-15;5]	81	-0.029	-2.88	0.004	***
COMEAN	[-15;5]	81	-0.020	-2.04	0.043	**
MA	[-15;5]	81	-0.024	-2.75	0.006	***
MM	[-15;5]	81	-0.020	-2.33	0.021	**
PANEL B4.: FROM TOPIX SMALL-CAP						
RAW	[-15;5]	7	0.023	0.64	0.522	
COMEAN	[-15;5]	7	0.032	0.91	0.362	
MA	[-15;5]	7	-0.001	-0.03	0.978	
MM	[-15;5]	7	0.015	0.47	0.638	

Table 13. Impact of GPIF's Programme on FTSE ESG Scores

The table shows the impact of GPIF's overall programme on FTSE ESG Scores. Panel A shows the availability of FTSE scores for large and mid-cap companies by country on the last day of each year. We used all scores to define "the World" (ex-Japan) and selected the seven countries with the largest number of scores as control group samples. Panel B shows parameter estimates for simple difference in differences regressions. The adoption of ESG indexes by GPIF in July 2017 is considered as the beginning of the treatment. The period dummy is set to one afterwards and zero before. In Column 1 the control group are all countries outside of Japan with FTSE scores. In Columns (2)-(8) the control group is the country mentioned in the header, for example Australia in Column (2).

Panel A. FTSE ESG Scores by County and over Time

ISO Country Code	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
AT	10	7	7	7	7	8	8	7	7	68
AU	97	92	94	93	91	103	109	108	107	894
BE	11	11	12	13	13	12	14	15	15	116
CA	73	65	62	63	61	56	52	52	49	533
CH	44	45	46	46	47	48	47	50	48	421
DE	62	60	70	71	72	79	76	80	81	651
DK	16	16	19	19	18	21	20	20	21	170
ES	27	27	30	32	32	28	25	24	24	249
FI	11	11	13	14	14	15	16	16	15	125
FR	80	78	82	82	82	86	80	79	72	721
GB	130	132	137	136	137	127	119	118	113	1,149
GR	4	4								8
HK	89	88	91	92	85	87	85	88	78	783
IE	4	4	4	4	5	5	4	4	5	39
IL	31	25	27	26	27	26	22	27	35	246
IT	30	28	32	33	35	36	35	34	34	297
JP	462	453	489	500	515	505	493	502	506	4,425
KR	107	114	122	120	123	128	132	133	158	1,137
NL	25	23	25	25	26	26	24	28	29	231
NO	11	11	10	10	11	13	15	17	16	114
NZ	13	13	15	13	11	12	12	13	13	115
PL					13	14	13	12	10	62
PT	5	4	4	4	4	4	4	4	4	37
SE	34	31	34	37	38	45	46	57	53	375
SG	40	39	36	32	32	40	39	35		324
US	632	606	616	609	614	610	575	603	595	5,460
Total	2,048	1,987	2,077	2,081	2,113	2,125	2,066	2,130	2,123	18,750

Panel B. Difference-in-Differences - Japan vs. International Control Groups

Country	(1) World	(2) AU	(3) DE	(4) FR	(5) GB	(6) HK	(7) KR	(8) US
Diff-in-diff	0.18***	0.06	0.39***	0.35***	0.17***	-0.20***	0.21***	0.17***
	(0.0248)	(0.0463)	(0.0588)	(0.0431)	(0.0393)	(0.0515)	(0.0503)	(0.0272)
Observations	37,585	10,676	10,197	10,323	11,172	10,450	11,191	19,793
R-squared	0.11	0.15	0.15	0.24	0.25	0.09	0.08	0.12
Mean control t(0)	2.58	2.66	3.03	3.40	3.14	1.69	1.82	2.42
Mean treated t(0)	1.87	1.87	1.87	1.87	1.87	1.87	1.87	1.87
Diff t(0)	-0.71	-0.79	-1.15	-1.53	-1.26	0.18	0.05	-0.55
Mean control t(1)	3.04	3.24	3.28	3.69	3.60	2.53	2.25	2.89
Mean treated t(1)	2.51	2.51	2.51	2.51	2.51	2.51	2.51	2.51
Diff t(1)	-0.52	-0.73	-0.76	-1.18	-1.09	-0.02	0.26	-0.38

11.   Figures

Figure 1 – Asset Management One ESG Issues and Milestones

The left part of the Figure shows the 23 engagement categories adopted by AM One; they are divided into "E", "S", "G" and a horizonal category "ESG" that combines issues from multiple themes. For example, ESG4 "CSR Supply Chain Management" could involve "E" and/or "S" issues in the supply chain. The right-hand side shows the milestones AM One uses to measure progress. Progress is not linear and reaching Level 4 "issue recognised by senior management" is an important threshold marker. Level 8 marks the completion of an engagement agenda item.
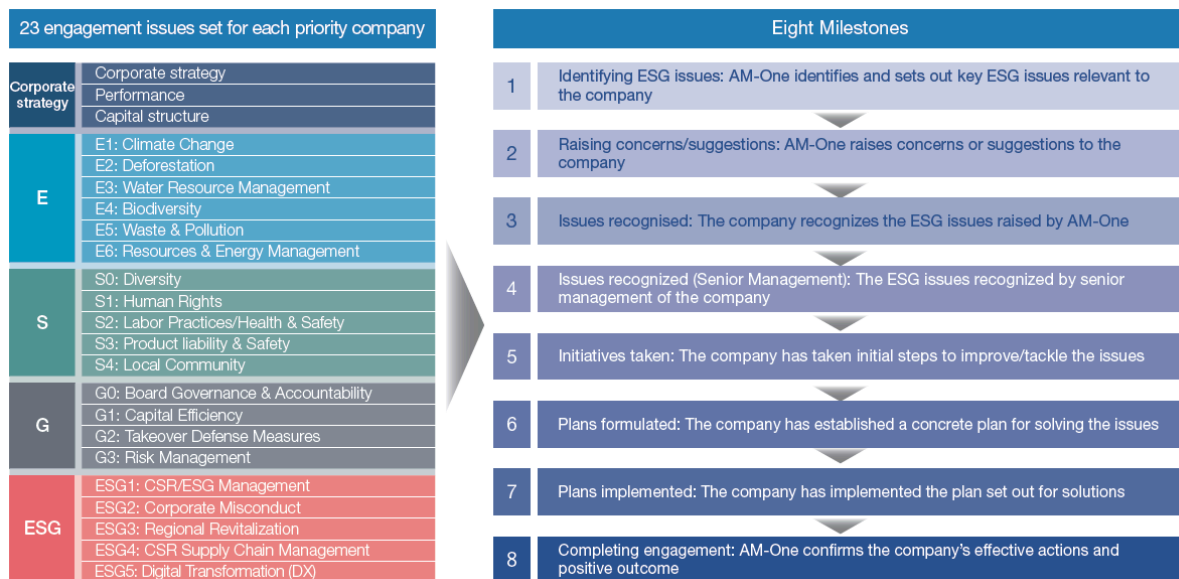
Source: Asset Management One Sustainability Report 2021, pg. 46

Figure 2. Standard Difference in Differences: Parallel Trends

The Figure shows the evolution in the mean FTSE and MSCI ESG scores for companies that were included at least once between 2016 and the end of 2022 in the TOPIX500 index. The index is split into groups: One, companies that were engaged on any topic (E,S,G or ESG) at some point between 2018 and 2022; two, companies that were never engaged. During the pre-engaged period (January 2016 to June 2018) the lines exhibit parallel trends, which is a necessary pre-condition for applying formal difference-in-differences tests. FTSE ESG scores are reported bi-annually; they are not industry weighted. MSCI ESG scores are also reported for June and December. MSCI's methodology for the G pillar score changed at the end of 2020 resulting in a discontinuity for the unweighted score; hence the plots in column (b) stop at the end of 2020.

(a) FTSE ESG Scores – Not Industry Adjusted   (b) MSCI ESG Scores – Not Industry Adjusted

(1)   ESG Scores – E,S,G or ESG Engaged

(2)   E Scores – E Engagement

(3)   S Scores – S Engagement

(4)   G Scores – G Engagement
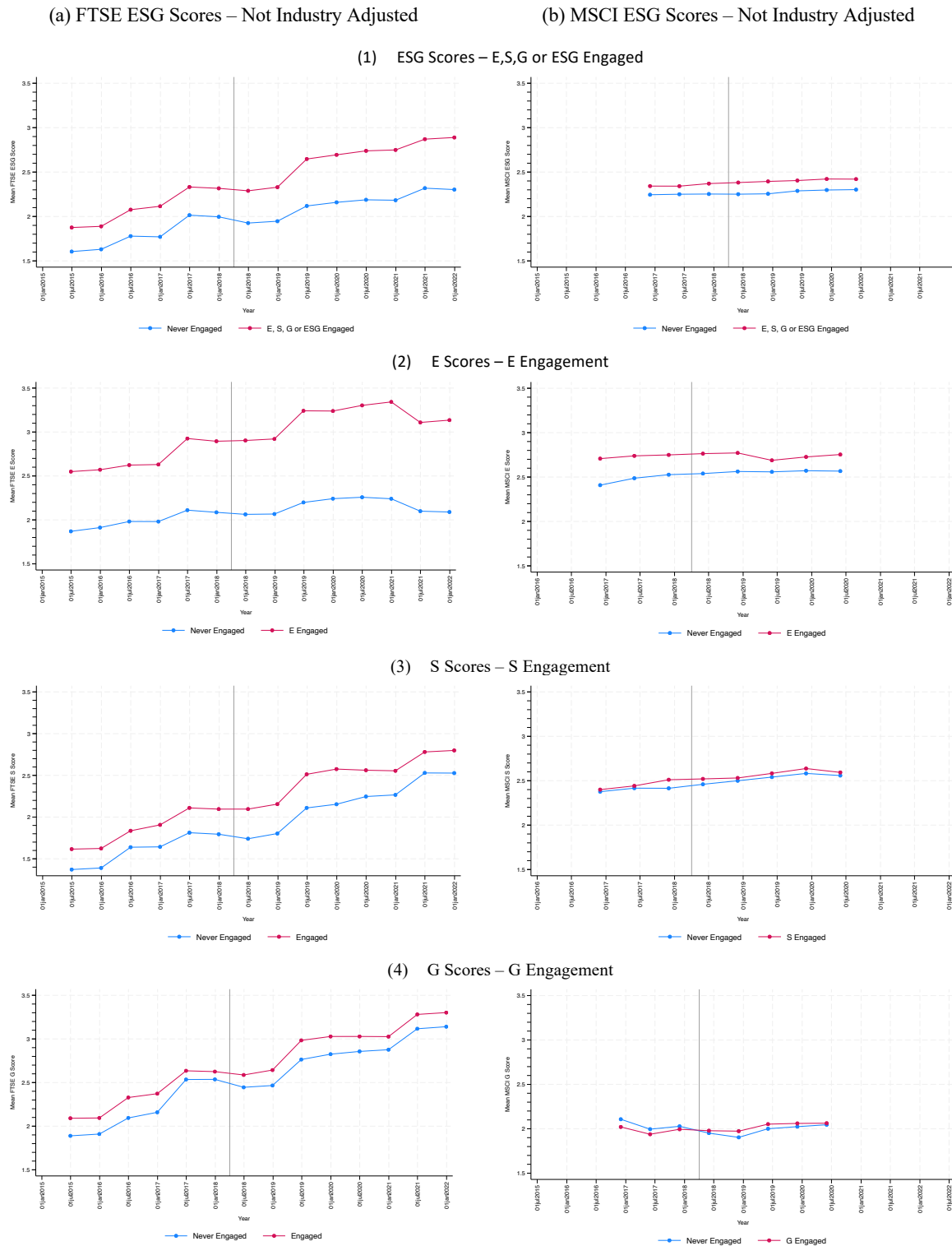
Figure 3. The Evolution of Combined Engagements over Time

The figure shows the evolution of engagements of TOPIX all-share constituents on any issue (E,S,G or ESG) over time in terms of absolute numbers and as a fraction of the respective index segment. Figure (a) traces the TOPIX 100, Figure (b) the TOPIX 400 midcap index and Figure (c) the small cap 1 and small cap 2 indexes. The indicator variable "engaged" was set to one when the first contact with a TOPIX company took place.
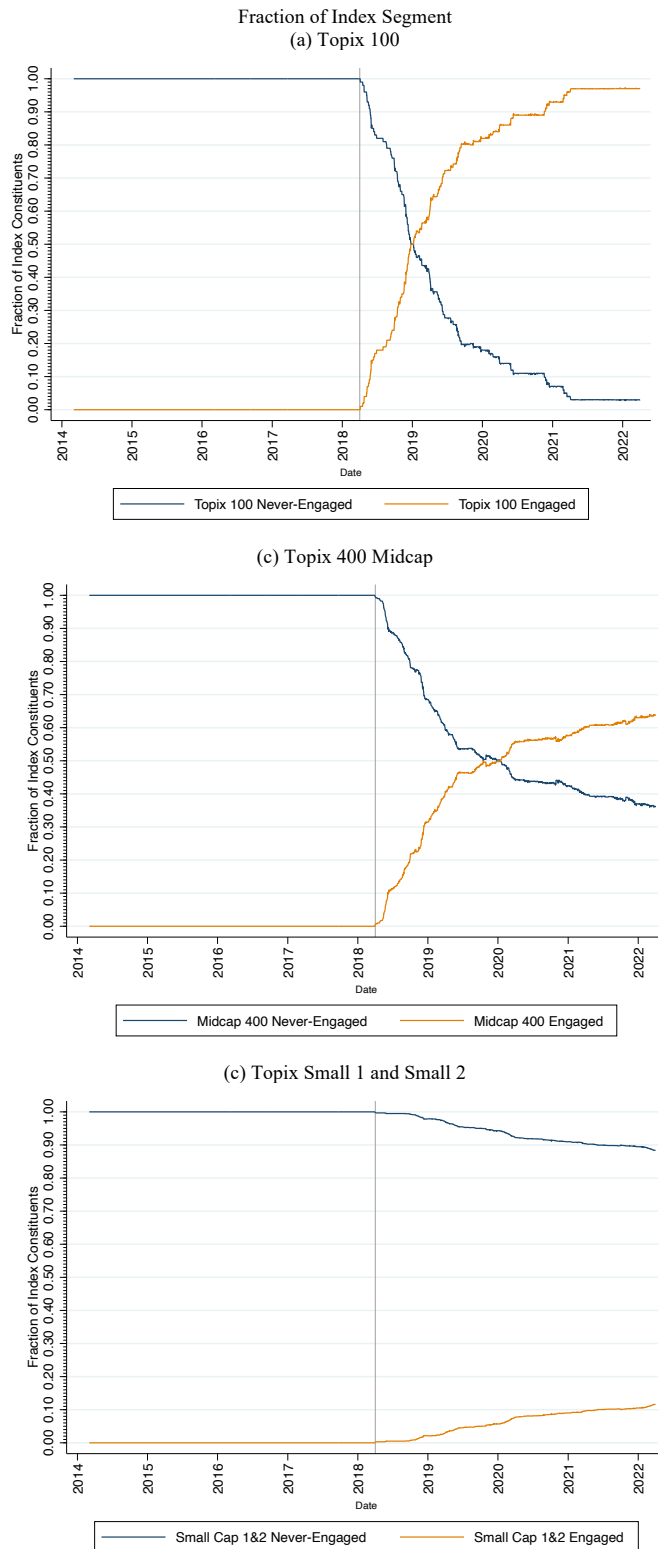
Figure 4. The Evolution of E,S and G Engagements over Time

The figure shows the evolution of engagement of TOPIX all-share constituents on environmental (E), social (S) and governance (G) issues over time as a fraction of all companies in the respective sub-index segment. Figure (a) traces the TOPIX 100, Figure (b) the TOPIX 400 midcap index and Figure (c) the Smallcap 1 and Smallcap 2 indexes. The indicator variable "engaged" was set to one when the first contact on an environmental issue with a TOPIX company took place.

(a) Topix 100
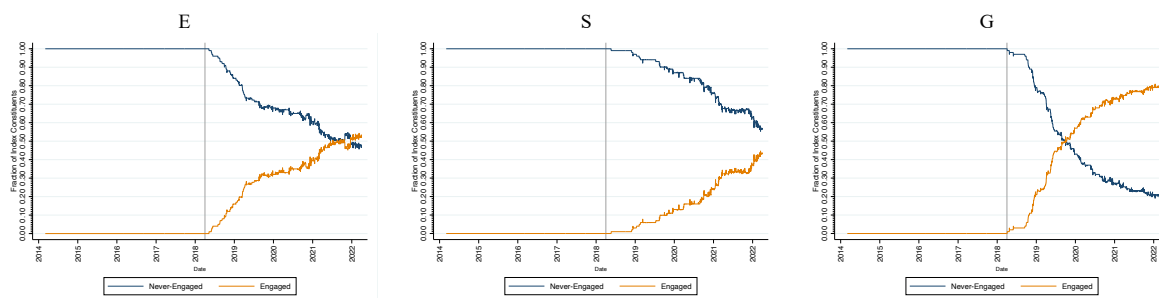
(b) Topix 400 Midcap
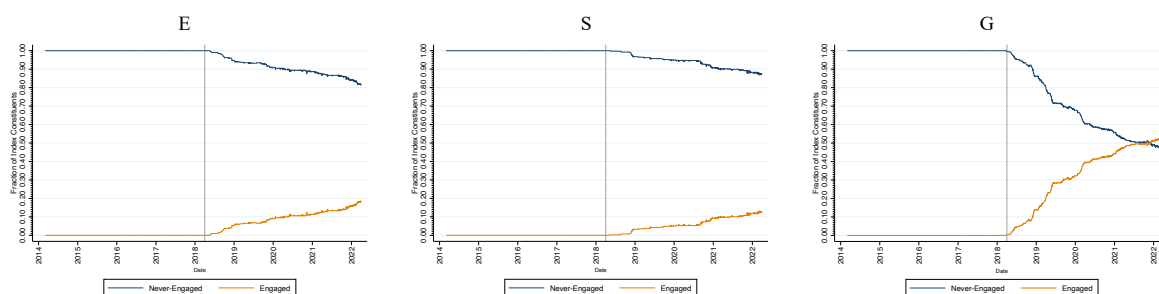
(c) Topix Small 1 and Small 2
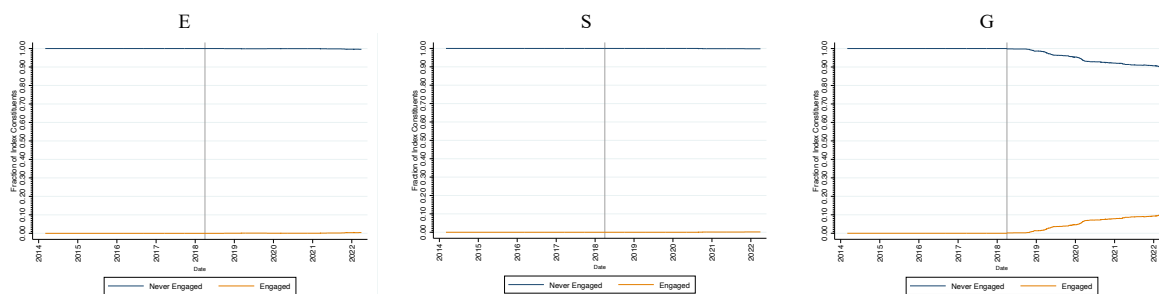
Figure 5. Staggered Difference in Differences: Event Time Analysis

The figure shows event study plots of estimated average treatment effects around the time of the treatment. The dependent variable are the respective FTSE scores (ESG, E, S and G). Since no covariates were specified the parameters were estimated using ordinary least squares regressions with robust and asymptotic standard errors using the *csdid* package (Rios-Avila et al., 2021). In the first column the control group are companies that were never engaged; in the second column the control group are companies that were not yet engaged (Callaway and Sant'Anna, 2021). The capped bars represent a 95% confidence interval. For significance levels see the event estimates in Panel 1 of Table 7.
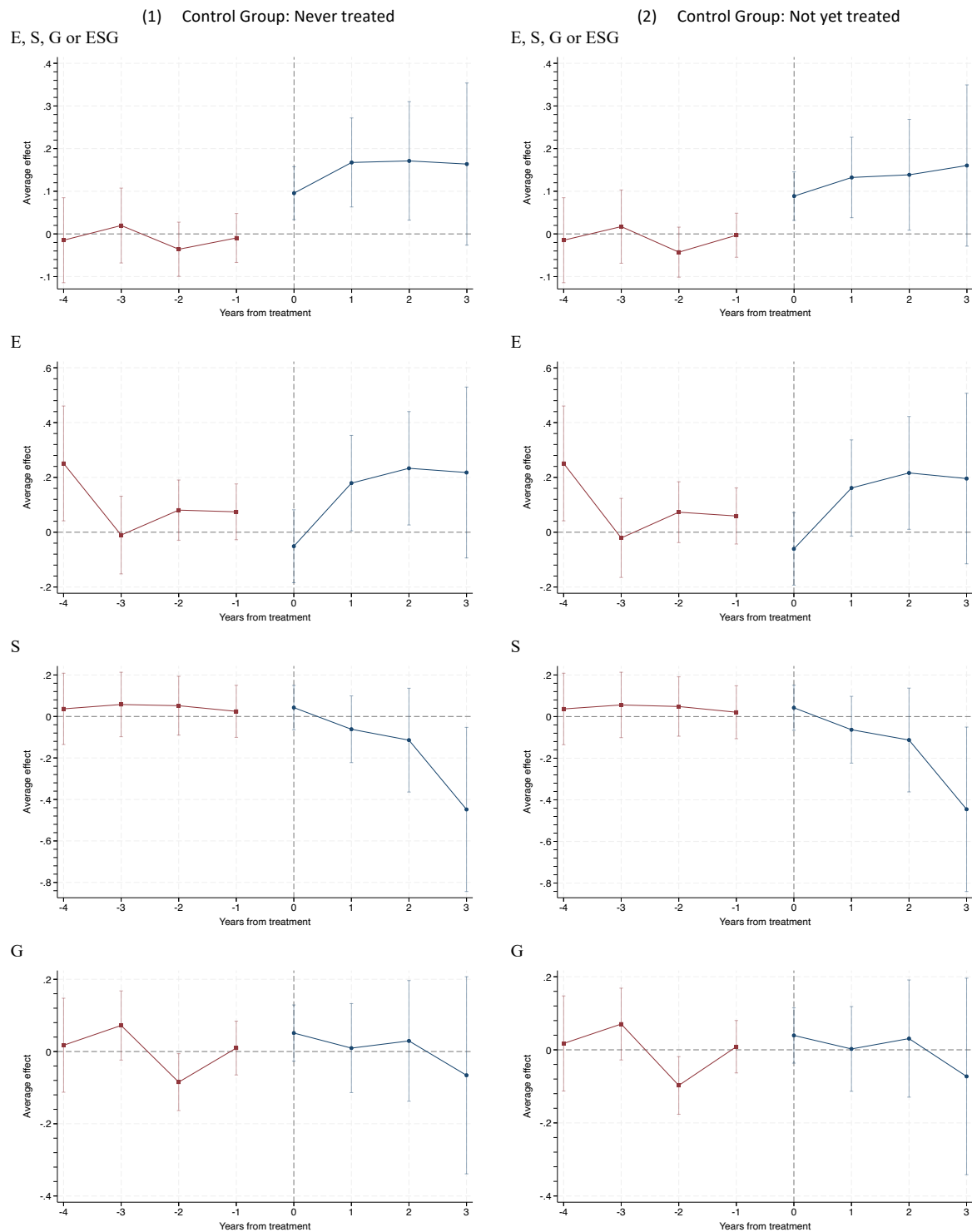
Figure 6. Returns from Blossom Index Inclusions and Exclusions

The figure shows cumulative average abnormal returns for FTSE Blossom inclusions and exclusions between 2017 and 2022 for the market model and the market adjusted model. Event time is set at t=0 for the effective date. Inclusions and exclusions are announced to companies and the market nine trading days in advance. See the event study return table for further details.
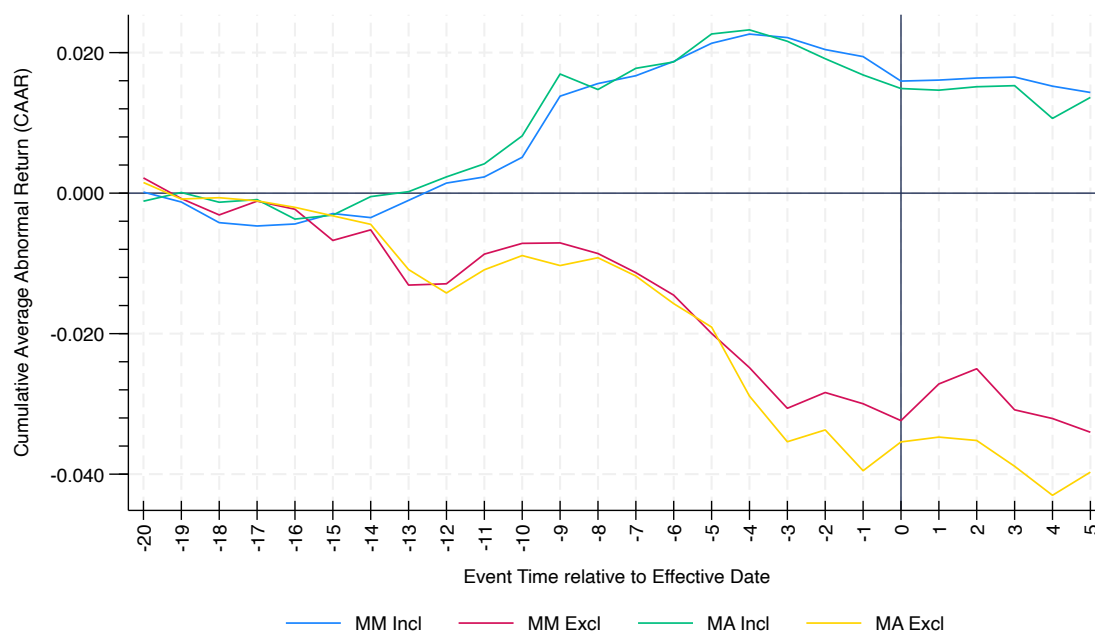
Figure 7. Returns from MSCI ESG Leaders Index Inclusions and Exclusions

The figure shows cumulative average abnormal returns MSCI ESG Leader inclusions and exclusions between 2017 and 2022 for the market model and the market adjusted model. Event time is set at t=0 for the effective date. Inclusions and exclusions are announced to clients nine trading days in advance. See the event study return table for further details.
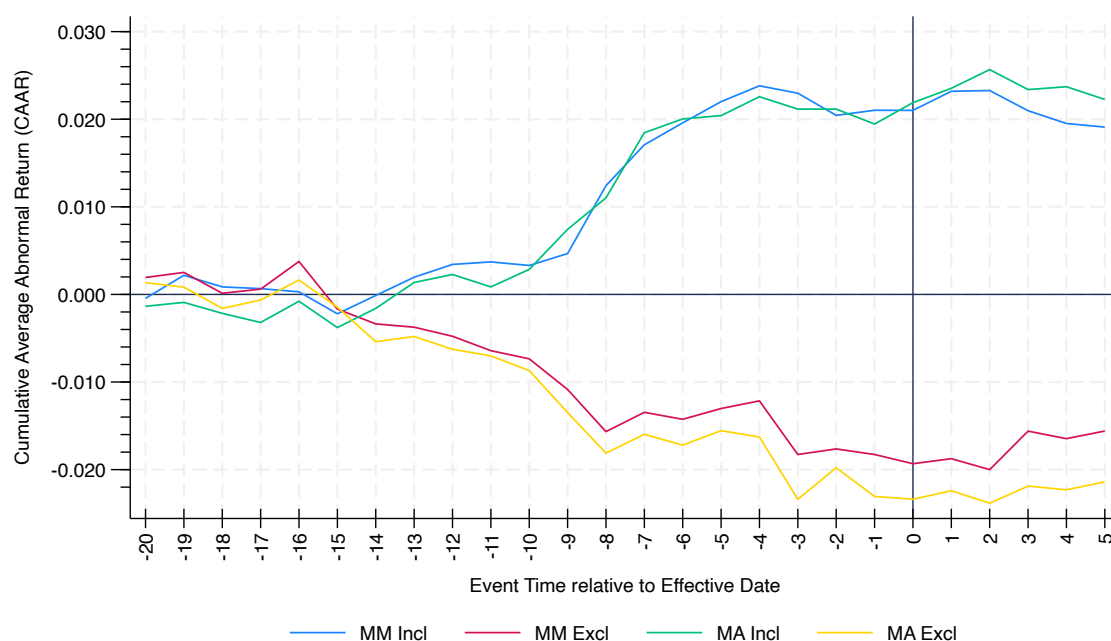
Figure 8. GPIF Treatment and FTSE ESG Scores

The figures show the evolution of FTSE ESG scores over time for different countries. The Japanese sample is largely comprised of TOPIX500 companies that are considered as treated under the GPIF programme. The first vertical line marks the adoption of the FTSE Blossom and MSCI ESG Leader indexes by GPIF; the second vertical line marks the first engagement by AM One. Plot 1 shows the evolution of the average score for all countries relative to the treated (Japan). Scores are plotted bi-annually at the end of June and the end of December of each year.

(1) Control Group: Rest of the World
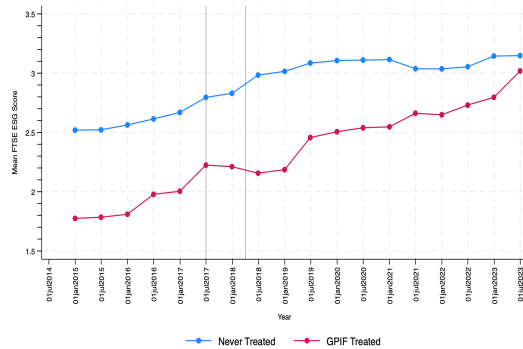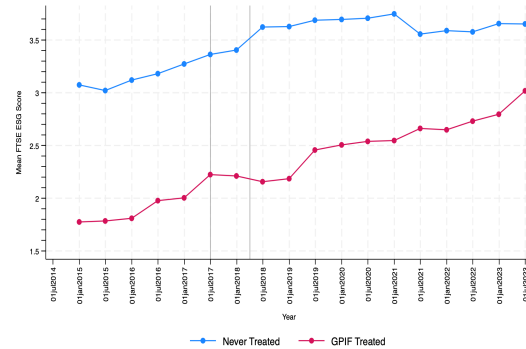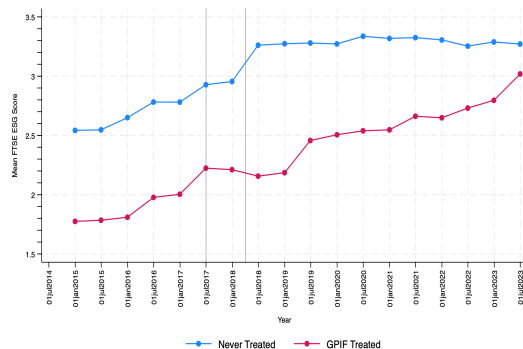
(5) Control Group: Great Britain

(2) Control Group: Australia

(6) Control Group: Hong Kong
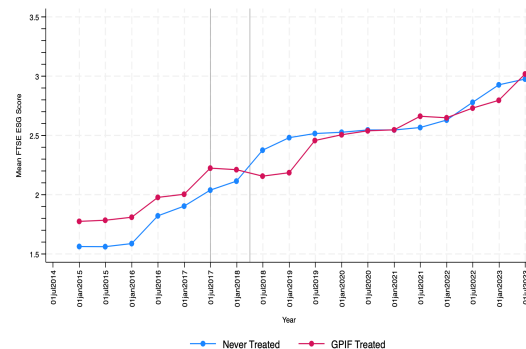
(3) Control Group: Germany
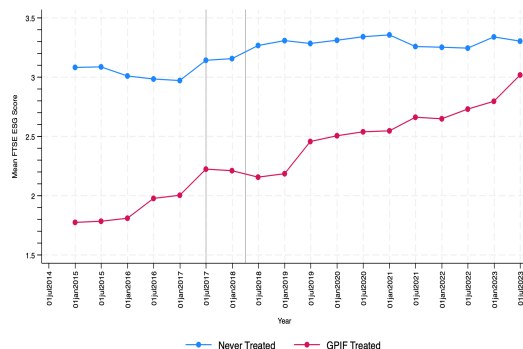
(7) Control Group: South Korea
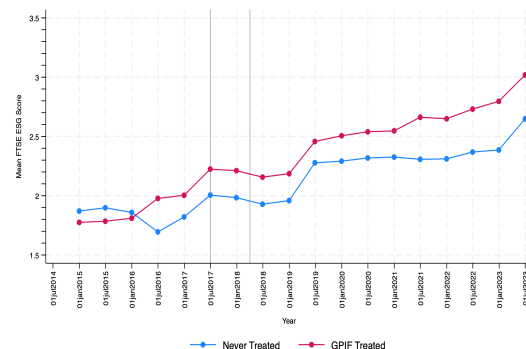
(4) Control Group: France
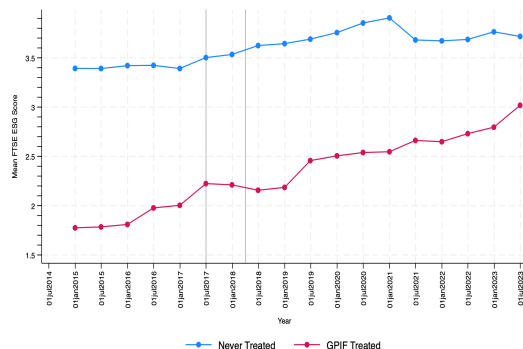
(8) Control Group: The United States
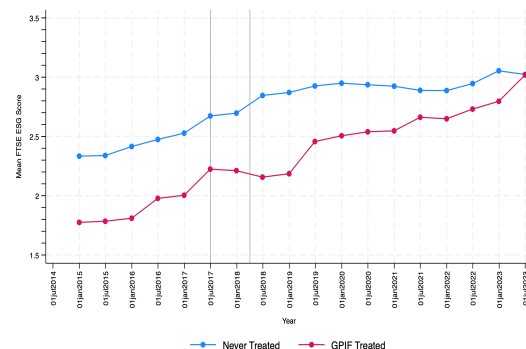
Figure 9. GPIF and ESG Scores: Dynamic Treatment Effects

The figures show the evolution of FTSE scores in event time. Time zero is set as the adoption of the FTSE Blossom and MSCI ESG Leader index by GPIF. The single parameters are estimates of $\theta_{es}(e)$ ((Callaway and Sant'Anna, 2021, p. 209) from a multiple period difference in differences regression. The control group is "never treated" by GPIF under its dometic stewardship programme for companies outside of Japan.
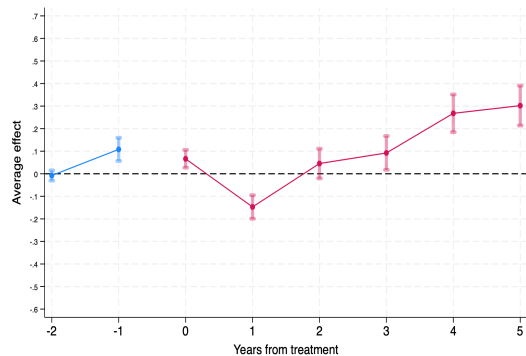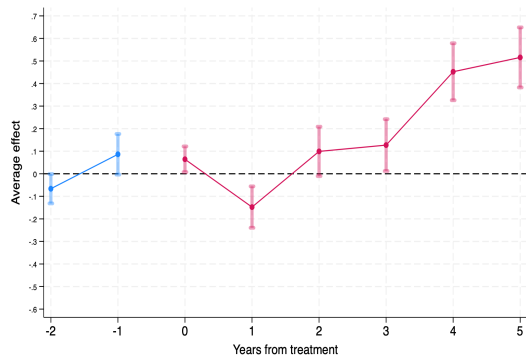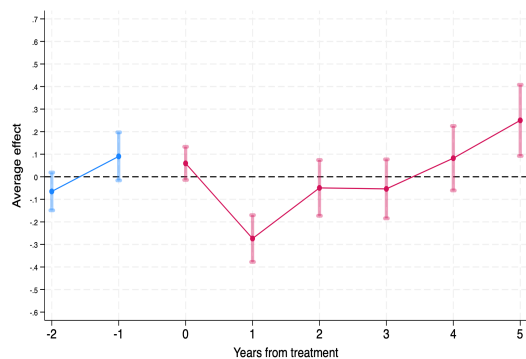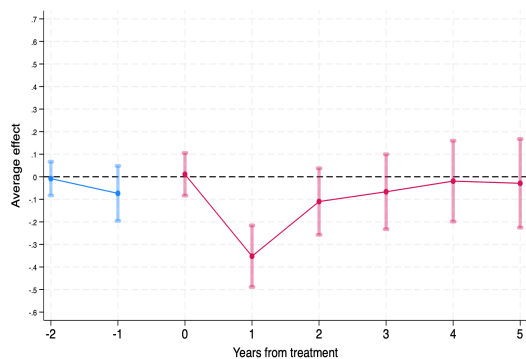

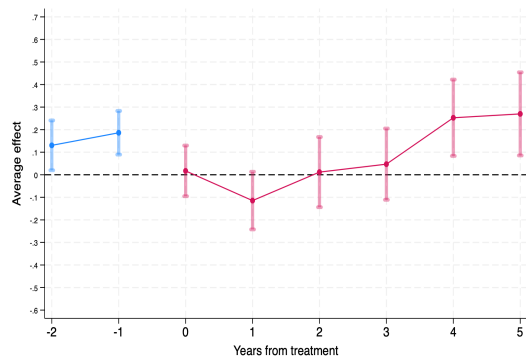(1) Control Group: Rest of the World
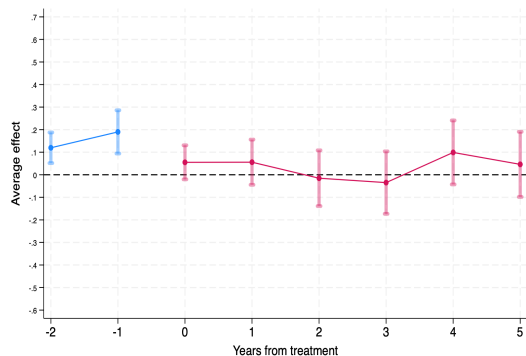(5) Control Group: Great Britain
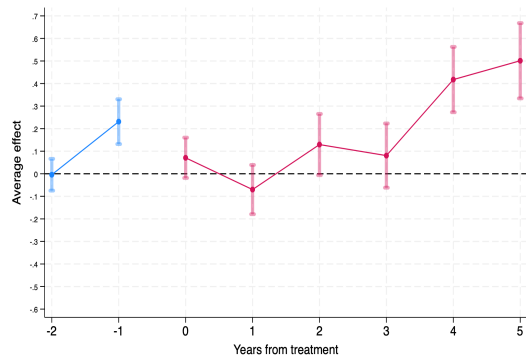(2) Control Group: Australia
(6) Control Group: Hong Kong
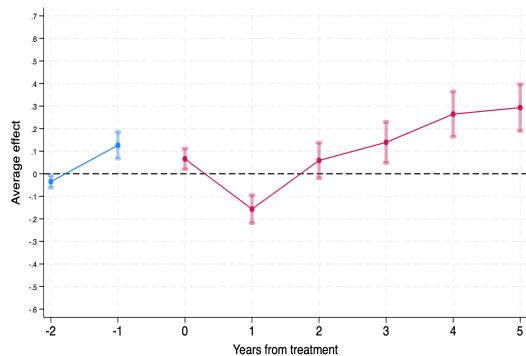(3) Control Group: Germany
(7) Control Group: South Korea
(4) Control Group: France
(8) Control Group: The United States

12.   Appendix: GPIF

Table 14. GPIF Domestic Equities by Benchmark

The table shows the market value of active and passive domestic equity portfolios managed by GPIF's external asset managers. Panel A reports the actively management portfolios and the name of the respective asset managers. Panel B reports the passively managed portfolio. Panel C breaks out ESG assets from Panel B. Acronymns: TOPIX - TOPIX (incl. dividends); RN-P - RUSSELL/NOMURA Prime Index (incl. dividends); RN-V - RUSSELL/NOMURA Large Cap Value Index (incl. dividends); RN-S - RUSSELL/NOMURA Small Cap Index (incl. dividends); RN-SG - RUSSELL/NOMURA Small Cap Growth Index (incl. dividends); MSCI-JS - MSCI Japan Small (incl. dividends); MSCI-IR - MSCI Japan IMI Equity REITS Index (incl. dividends); MSCI-ESG - MSCI Japan ESG Select Leaders Index; MSCI-WIN - MSCI Japan Empowering Women Index (WIN); FTSE-BL - FTSE Blossom Japan Index; SP-C - S&P/JPX Carbon Efficient Index; RAFI - Nomura RAFI Index.

(Unit : ¥billion)	2017	2018	2019	2020	2021
Panel A: Active Management					
Benchmark Index					
MSCI-JS	172.1	241.5	209.4		
RN-S	178.7	56.2	47.8		110.6
RN-SG		101.8	82.1	127.9	116.0
RN-V	804.2	749.2	633.9		356.3
TOPIX	2735.6	2481.1	2232.6	1504.2	2,513.7
Sub-Total	3890.6	3629.8	3205.8	1632.1	3,096.6
Panel B: Passive Management					
Tracking Index					
TOPIX	27,435.2	25,532.2	22,479.6	34,159.4	36,028.4
FTSE-BL	526.7	642.8	931.4	1,490.6	983.0
FTSE-BLSR					800.0
JPX	1,995.7	1,908.8	1,187.9		
MSCI-ESG	622.9	804.3	1,306.1	2,026.8	2,099.0
MSCI-IR				66.7	143.5
MSCI-J	316.2	0.6			
MSCI-WIN	388.4	474.6	797.8	1,236.2	1,245.7
RAFI	1,855.2	1,753.5	1,535.3	1,354.5	1,422.3
RN-P	1,643.5	1,564.8	1,420.7	2,037.1	2,072.6
SP-C		387.8	980.2	1,536.5	1,567.8
SP-G	2,023.9	1,945.6	1,696.9		
Sub-Total	36,807.7	35,015.0	32,335.9	43,907.8	46,362.3
Total	40,698.3	38,644.8	35,541.7	45,539.9	49,458.9
Panel C: Passive domestic equities invested in ESG					
FTSE-BL	526.7	642.8	931.4	1,490.6	983.0
FTSE-BLSR					800.0
MSCI-ESG	622.9	804.3	1,306.1	2,026.8	2,099.0
MSCI-WIN	388.4	474.6	797.8	1,236.2	1,245.7
SP-C	0.0	387.8	980.2	1,536.5	1,567.8
Total	1,538.0	2,309.5	4,015.5	6,290.1	6,695.5
Fraction	4.2%	6.6%	12.4%	14.3%	13.5%
Sources	(1)	(2)	(3)	(4)	(5)

Sources Legend:

(1)   GPIF Annual Report FY 2017, pg. 58
(2)   GPIF Annual Report FY 2018, pg. 66
(3)   GPIF Annual Report FY 2019, pg. 82
(4)   GPIF Annual Report FY 2020, pg. 94
(5)   GPIF Annual Report FY 2021, pg. 91

Table 15. GPIF External Asset Managers - Japanese Domestic Equities

The table shows the market value of assets under management in Yen billions per external manager at year end. Panel A shows active managers that make investment decision. Panel B shows the breakdown for passive managers.

End of FY	Market value under management FY end (billions)					Percentage of total				
	2017	2018	2019	2020	2021	2017	2018	2019	2020	2021
Panel A: Active Managers										
AM One	754.5	697.4	613.2	223.3	345.1	19.4%	19.2%	19.1%	6.8%	11.1%
Capital International	495.4	471.3	434.1	673.2	619.3	12.7%	13.0%	13.5%	20.5%	20.0%
Daiwa SB Investments	408.6	377.7				10.5%	10.4%			
Eastspring Investments	276.2	170.5	138.0			7.1%	4.7%	4.3%		
FIL Investments (Japan)	392.9	352.1	337.3	521.2	498.9	10.1%	9.7%	10.5%	15.8%	16.1%
Invesco AM	217.3	221.7	194.4	378.1	429.6	5.6%	6.1%	6.1%	11.5%	13.9%
JP Morgan AM	395.6	371.5	319.4			10.2%	10.2%	10.0%		
Nikko AM	319.8	205.9	186.4	63.0		8.2%	5.7%	5.8%	1.9%	
Nomura AM	230.5	297.7	257.2	435.4	110.6	5.9%	8.2%	8.0%	13.2%	3.6%
Russell Investments	101.2	187.8	166.0	242.9	231.2	2.6%	5.2%	5.2%	7.4%	7.5%
Schroder Investment	287.2	276.2	245.3	357.5	361.9	7.4%	7.6%	7.7%	10.9%	11.7%
Seiryu AM	11.4					0.3%				
Sumitomo Mitsui			314.5	327.6	356.3			9.8%	10.0%	11.5%
Lazard Japan				66.3	143.7				2.0%	4.6%
Total Active	3,890.6	3,629.8	3,205.8	3,288.5	3,096.6	100.0%	100.0%	100.0%	100.0%	100.0%
Panel B: Passive managers										
AM One	11,842.4	11,419.4	10,500.3	11,299.1	11,488.1	32.2%	32.6%	32.5%	25.7%	24.8%
BlackRock Japan	7,715.8	6,043.0	4,195.7	5,825.4	6,408.2	21.0%	17.3%	13.0%	13.3%	13.8%
FIL Investments (Japan)		98.2	88.9	126.4	231.5		0.3%	0.3%	0.3%	0.5%
Goldman Sachs AM	2,023.9	1,945.6	1,696.9			5.5%	5.6%	5.2%		
Mitsubishi UFJ Trust BK	8,926.2	8,162.5	6,694.9	9,889.4	9,307.5	24.3%	23.3%	20.7%	22.5%	20.1%
Nomura AM	1,855.2	1,753.5	1,535.3	1,421.2	1,565.8	5.0%	5.0%	4.7%	3.2%	3.4%
Resona AM		982.5	2,819.5	6,425.0	7,897.5		2.8%	8.7%	14.6%	17.0%
Sumitomo Mitsui Trust AM	4,444.2	4,610.3	4,804.4	8,921.3	9,463.7	12.1%	13.2%	14.9%	20.3%	20.4%
Total Passive	36,807.7	35,015.0	32,335.9	43,907.8	46,362.3	100.0%	100.0%	100.0%	100.0%	100.0%
Total	40,698.3	38,644.8	35,541.7	47,196.3	49,458.9					
Active %	9.6%	9.4%	9.0%	7.0%	6.3%					
Passive %	90.4%	90.6%	91.0%	93.0%	93.7%					
Sources	(1)	(2)	(3)	(4)	(5)					

Source: GPIF Annual Reports

Sources Legend:

(1)   GPIF Annual Report FY 2017, pg. 58
(2)   GPIF Annual Report FY 2018, pg. 66
(3)   GPIF Annual Report FY 2019, pg. 82
(4)   GPIF Annual Report FY 2020, pg. 94
(5)   GPIF Annual Report FY 2021, pg. 91

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html

www.ecgi.global/content/working-papers