Controlled Firms, Preferences, and Carbon Emissions

Finance Working Paper N° 983/2024
August 2026

Alexander Dyck
University of Toronto and ECGI

Karl V. Lins
University of Utah

Lukas Roth
University of Alberta and ECGI

Mitch Towner
Arizona State University

Hannes F. Wagner
Bocconi University and ECGI

ECGI Working Paper Series in Finance

Controlled Firms, Preferences, and Carbon Emissions

Working Paper N° 983/2024
August 2026

Alexander Dyck
Karl V. Lins
Lukas Roth
Mitch Towner
Hannes F. Wagner

We thank Morten Bennedsen, Harold Contreras, Art Durnev, Florencio Lopez de Silanes, Beatriz Martínez, Pedro Matos, Luigi Zingales, Helen Zhang, seminar participants at the University of Arizona, University of Delaware, HEC Montreal, Laval University, McGill, Skema Business School, University of Missouri, University of North Carolina at Chapel Hill?

Abstract

Controlled firms represent the dominant ownership structure globally, yet their environmental impact remains unclear. We test this for 3,769 firms from 35 countries using actual carbon emissions data and an environmental preference measure built using retrieval-augmented large language models. Overall, controlling owners' preferences affect emissions but are insufficient to address environmental externalities. Controlled firms with low environmental preferences emit approximately 25% more carbon than widely held firms, while those with high environmental preferences have emissions no different than widely held firms. Controlled firms' carbon performance is even worse in settings where marginal emissions abatement costs are high.

Keywords: Carbon emissions, family ownership, environmental sustainability, preferences

JEL Classification: G32, G34, G54

Alexander Dyck

University of Toronto

Karl V. Lins

University of Utah

Lukas Roth

University of Alberta

Mitch Towner

Arizona State University

Hannes F. Wagner

Bocconi University

Controlled Firms, Preferences, and Carbon Emissions

Alexander Dyck
University of Toronto

Karl V. Lins
University of Utah

Lukas Roth
University of Alberta

Mitch Towner
Arizona State University

Hannes F. Wagner
Bocconi University

Journal of Financial and Quantitative Analysis, forthcoming.

August 27, 2026

Abstract

Controlled firms represent the dominant ownership structure globally, yet their environmental impact remains unclear. We test this for 3,769 firms from 35 countries using actual carbon emissions data and an environmental preference measure built using retrieval-augmented large language models. Overall, controlling owners' preferences affect emissions but are insufficient to address environmental externalities. Controlled firms with low environmental preferences emit approximately 25% more carbon than widely held firms, while those with high environmental preferences have emissions no different than widely held firms. Controlled firms' carbon performance is even worse in settings where marginal emissions abatement costs are high.

Keywords: Carbon emissions, family ownership, environmental sustainability, preferences
JEL Classification: G32, G34, G54

Author contacts: adyck@rotman.utoronto.ca (corresponding author); karl.lins@eccles.utah.edu; lukas.roth@ualberta.ca; mitchtowner@email.arizona.edu; and hannes.wagner@unibocconi.it. We thank Morten Bennedsen, Harold Contreras, Art Durnev, Florencio Lopez de Silanes, Beatriz Martínez, Pedro Matos, Luigi Zingales, Helen Zhang, seminar participants at the University of Arizona, University of Delaware, HEC Montreal, Laval University, McGill, Skema Business School, University of Missouri, University of North Carolina at Chapel Hill, University of Texas at San Antonio, University of Toronto, Waseda University, and Xi'an Jiaotong University, and conference participants at the 2024 ECGI-Baltic Family Firm Institute Conference, 2025 EFMA Conference, 2025 IE Madrid Symposium on Financial Markets in the Era of Climate Change, 2025 Sustainability at Crossroads: The Edinburgh and International Review of Finance Conference, 2025 Paris-Dauphine "Ownership, Control and Performance" Conference, 2025 Santiago Finance Workshop, 2026 Drexel Corporate Governance Conference, 4th HKU Summer Finance Conference (2026), 2026 LSW Academic Sessions on Accounting and Finance, and 2026 Spanish Finance Forum, for helpful comments and suggestions. We are grateful to the Social Sciences and Humanities Research Council of Canada, the BAFFI Centre on Economics, Finance and Regulation, the Johnston Centre for Corporate Governance Innovation, the Michael Lee-Chin Family Institute for Corporate Citizenship, and the Innocenzo Gasparini Institute for Economic Research for financial support. This publication benefited from European Union – Next Generation EU funds, component M4C2, investment 1.1. PRIN 2022 PNRR Challenges in ESG Investing and Sustainable Finance Prot. P2022HTXXM. None of the authors has a conflict of interest to declare.

I. Introduction

If a firm changed its focus from financial value maximization to shareholder welfare maximization would it be more *clean*? Hart and Zingales (2017) assess this and conclude that, for the widely held firm as of today, the answer is no. While investors in such a firm may have non-pecuniary preferences for *clean*, "the market for corporate control will push a board who wants to choose *clean* into a choice of *dirty*: we call this an *amoral drift*" (p. 256). However, for the closely controlled firm insulated from takeover pressure, their answer is different. As they state, if a closely controlled company "has a single shareholder, nobody would suggest that this single shareholder cannot instruct directors to maximize her utility, rather than her financial return." (p. 263).

The likelihood that a controlled firm chooses *clean* depends on whether its owner has *clean* preferences along with additional factors. If the marginal costs of becoming *cleaner* are relatively high, environmental preferences would have to be very strong to overcome these costs. Further, all entrenched owners can consume pecuniary private benefits of control by diverting current cash flows to themselves if they wish, reducing cash available for *clean* investment.

Our paper investigates whether controlled firms are *cleaner* by assessing the actual carbon emissions of 3,769 non-U.S. firms from 35 countries in the years 2023 and 2024. In this global sample, controlled firms are generally the most common type of firm ownership structure.[1] In our sample, 54% of firms are controlled and 46% are widely held. We identify three categories of controlled firms: family, government, and other (42%, 7%, and 6% of sample firms, respectively).

No database currently exists to measure the environmental preferences of a controlling shareholder. A contribution of our paper is to demonstrate that generative AI can be used to construct such a measure of the environmental preferences of a controlling shareholder, at scale.

[1] See Aminadav and Papaioannou (2020) and La Porta, Lopez-De-Silanes, and Shleifer (1999), among others.

We measure these preferences for family-controlled firms, given their prevalence. We construct an index for *clean* using five actions observable by AI: i) the family's personal philanthropy towards environmental causes; ii) the family's public advocacy for environmental issues; iii) the family's participation in environmental NGOs; iv) 'green' investments in the family's personal portfolio outside the firm; and v) the family's support and contributions for environmental policies.

To measure expected marginal costs of reducing carbon emissions we would ideally obtain detailed information on all carbon-reducing projects available to a firm. Because such firm-level data do not exist and cannot be feasibly imputed, we measure marginal abatement costs at the country and industry levels, which we acknowledge has limitations. At the country level, we split firms into two subsamples based on a climate performance index that assesses each country's climate protection efforts relative to international standards. Where climate performance is weak, low-cost options to improve carbon emissions likely still exist. At the industry level, we split firms into two subsamples based on the importance of direct (Scope 1) emissions as a fraction of total industry emissions. We are most interested in direct emissions, as they can be more easily managed by a controlling owner and will likely have drawn greater regulator or social scrutiny, but also need to take account an industry's total emissions. If direct emissions are relatively unimportant, outside pressures to reduce emissions will be muted, and low-cost options to improve emissions likely remain.

We conduct tests that consider ownership type and the strength of environmental preferences. In all tests, we control for private benefits as well as industry, country, and firm characteristics. Our first key finding is that no category of controlled firm has lower carbon emissions than the widely held firm. The results range from no difference for the government-controlled firm to marginally and substantially higher emissions for family-controlled and other-controlled firms.

Our second key finding is that the environmental preferences of controlling shareholders are indeed associated with carbon emissions, but mostly in ways that indicate worse outcomes. In firms controlled by families with low environmental preferences, carbon emissions are on average about 24% higher than those of widely held firms, conditional on controls and fixed effects. In contrast, carbon emissions are similar to those of widely held firms when firms are controlled by families with high environmental preferences. These results indicate that while weak preferences are clearly linked to dirtier outcomes, strong preferences do not seem to overcome the pecuniary private benefit drift to remain *dirty* that affects all controlled firms. A broad implication—perhaps counterintuitive—is that the concept of letting controlling shareholder preferences dictate environmental outcomes is generally associated with worse, not better, environmental performance.

In an extension, we consider the expected marginal costs of becoming *clean*. Test evidence here is suggestive in nature given the difficulty of measuring marginal costs. Where improving carbon efficiency is arguably more costly (that is, in high CCPI countries and high Scope 1 emissions industries), our tests show that the carbon emissions of family-controlled firms with both low and high environmental preferences are never lower than those of widely held firms and are often significantly higher. Where carbon efficiency improvements are arguably less costly, we find that family-controlled firms with high environmental preferences have lower carbon emissions than widely held firms.

Taken together, these findings cast doubt on the hypothesis that moving from financial value maximization to shareholder welfare maximization (Hart and Zingales, 2017) will meaningfully increase the likelihood that firms choose to be *clean*. Controlling-shareholder preferences are a lever that can move their firms in the direction of better carbon performance, but our evidence suggests preferences are not a particularly strong lever. In most circumstances, family-controlled firms—even those with strong environmental preferences—do not have better carbon performance, and low-environmental-preference family control is associated with higher emissions.

Our analyses use actual emissions rather than imputed data to avoid a selective disclosure bias problem. If certain ownership types systematically do not disclose their actual emissions because they have worse emissions, regression estimates that use imputed data will falsely suggest better carbon performance for these ownership types than is actually the case (i.e., selection bias).

Actual carbon emissions have been infrequently reported until recently. For example, emission reporting rates in 2010 are just 18% for LSEG firms from our 35 sample countries that meet a minimum size threshold. We study the years 2023 and 2024 when emission reporting jumps, with disclosure rates of 81% in 2023 and 86% in 2024. This increased disclosure stems in part from the International Sustainability Standards Board, an independent standard-setting body within the IFRS foundation, issuing its first detailed standards in June of 2023, focusing exclusively on climate-related disclosures, particularly CO2-equivalent emissions. Countries that use IFRS and adopt the standards were expected to disclose CO2-equivalent emissions soon after. These new IFRS standards were not a 'shock' to carbon reporting in the strictest sense because they were telegraphed in advance, and not all countries that follow IFRS standards adopted them. Nonetheless, the actual release of these standards arguably represents a 'quasi-exogenous shock,' causing many previously non-reporting firms to begin reporting their carbon emissions.

Our paper makes several contributions. First, we provide new evidence on the way in which environmental preferences influence corporate sustainability by looking at the preferences of controlling shareholders. When family owners lack pro-environmental preferences, their firms produce substantially higher carbon emissions. This result extends earlier findings that outside investors' preferences and social values shape sustainability outcomes (e.g., Hong and Kacperczyk, 2009; Dimson, Karakaş, and Li, 2015; Hart and Zingales, 2017; Dyck, Lins, Roth, and Wagner, 2019; Krueger, Sautner, and Starks, 2020; Shive and Forster, 2020; Pástor, Stambaugh, and Taylor, 2021; Pástor, Stambaugh, and Taylor, 2022; Pedersen, Fitzgibbons, and Pomorski, 2021; Dyck, Lins, Roth, Towner, and Wagner, 2023; Iliev and Roth, 2023; Lins, Roth, Servaes, and Tamayo, 2024; Duchin, Gao, and Xu, 2025).

Second, we provide new evidence on the relationship between concentrated ownership and carbon emissions using data from 2023 and 2024 when the vast majority of firms report their actual carbon emissions. We find that controlled firms generally have higher emissions compared to widely held firms. A related paper, Borsuk, Eugster, Klein, and Kowalewski (2024), reports lower carbon emissions for family firms using 2010-2019 data, though their sample relies heavily on imputed emissions (approximately two thirds of observations) and U.S. firms. Potential selective disclosure bias in samples from this period could explain their different conclusions.[2]

Our third contribution is a methodological innovation demonstrating that retrieval-augmented LLMs can quantify otherwise unobservable owner characteristics by synthesizing dispersed public information across languages and sources. Prior work has used text-based and AI methods to extract signals from specific inputs—for example, earnings calls and disclosures

[2] The one test they conduct using actual emissions data is based on a much smaller sample than ours: 1,723 unique firms over a ten-year panel, approximately half of which are U.S. firms. In 2010 15% of U.S. firms that meet our sample selection criteria disclose their emissions. By 2024, this had only increased to 48% (vs. 86% for non-U.S. firms).

(Hassan, Hollander, van Lent, and Tahoun, 2019; de Kok, 2025; Siano, 2025). Our approach is situated within the rapidly evolving AI-in-finance literature (see, for example, the survey by Li, Wang, Ding, and Chen, 2024, and the regulatory perspective in IOSCO, 2025), emerging evaluations of financial-domain retrieval-augmented generation (RAG) systems (Zhao, Singh, Bhathena, Ramos, Joshi, Gadiyaram, and Sharma, 2024; Wang, Tan, Dou, and Wen, 2024; Choi, Kwon, Ha, Choi, Kim, Lee, Sohn, Lopez-Lira, 2025). Demonstrating that our RAG-derived measure explains differences in corporate outcomes, we show that otherwise unobservable owner values can be quantified at scale.

II. Theoretical Considerations

We begin our analysis by postulating that a controlled firm can behave in the way that Hart and Zingales (2017) suggest, acting as if there is only a single shareholder, and directing the firm to pursue the controller's utility maximization irrespective of what other investors may prefer if they had a voice. We make this assumption because a controlled firm is protected from takeover and activist pressures, and its board is effectively entrenched.

If a controlled owner has a strong non-pecuniary preference for *clean,* this should lead their firm to lower emissions all else equal. Two countervailing forces, however, may hamper the implementation of *clean* preferences. The first of these potential countervailing forces is the marginal cost of carbon emissions abatement. A strong preference for *clean* is likely to matter for emissions so long as the marginal costs of becoming *cleaner* are reasonably low. When marginal costs of becoming *cleaner* are relatively high, a controlling owner's environmental preferences would have to be very strong to overcome these costs. It is thus possible that there will be little differentiation in carbon emission outcomes between controlled firms with low and high environmental preferences when expected incremental carbon abatement costs are quite high.

The second countervailing force is the impact of pecuniary private benefits on a firm's likelihood to pursue *clean*. A significant body of literature focuses on how the control of voting power in firms gives controlling shareholders the potential to reward themselves at the expense of all other investors in the firm (e.g., Shleifer and Vishny, 1997). These control rights create financial benefits for controlling shareholders and are significant around the world (e.g., La Porta et al., 1999; Lins, 2003; Dyck and Zingales, 2004). If the controlled firm's owner uses their control rights for pecuniary private benefits, this shifts a firm towards being *dirty*—what we call a 'pecuniary private benefit drift.' This is because consuming a firm's current cash flow leaves less available for environmental investments.

Do controlled firm types differ in their expected non-pecuniary private benefits for the controlling owner, and thus the likelihood of overcoming these two countervailing forces? This question of the relative strength of environmental preferences across controlled owners is ultimately an empirical question, but theory offers some guidance.

Family-controlled firms can plausibly have strong environmental preferences and are more likely to act upon them than other types of controlled firms. Families may value reputation and have longer time horizons (e.g., Anderson and Reeb, 2003; Bennedsen and Fan, 2014), both of which can generate high and consistent levels of trust.[3] These long-term assets of the family can be impaired if a firm is perceived to not be *clean* enough, and given that much of the family's wealth is typically concentrated in their own firms, the family is also more sensitive to long-term

[3] For example, a family firm with such trust is better positioned to make implicit contracts with employees and contractors to reward them for making firm-specific investment decisions, thus generating value.

firm-specific risks such as a sudden clampdown on carbon emissions. These potential incentives may lead families to choose *clean*.

We do not expect government owners to have stronger preferences for *clean* compared to owners of a widely held firm. From an environmental standpoint, if a government controls a firm and believes that *clean* is important, it can enact regulations and force all firms to meet higher standards equally, rather than having only government-controlled firms outperform.

Other-controlled firms in our sample include multiple owners where no single owner is dominant, opaque-controlled firms in which there is a controlling shareholder but the ultimate controlling shareholder cannot be identified, and firms controlled by financial owners (private equity, hedge funds, venture capitalists). There are good reasons to expect weaker environmental preferences in such firms. Hart and Zingales (2017) suggest that an owner cares about a *clean* or *dirty* outcome only "if he feels responsible for the action in question" (p. 250). When no owner dominates, no owner is likely to feel the responsibility. When the owner is opaque, there is no reputational benefit from being seen to be *clean*. When an owner is financial-controlled and focused on short-term financial value maximization, there is little likelihood they will invest today to reduce carbon emissions for long-term payoffs.

In summary, this section motivates our tests that control for private benefits, focus on different types of controlled owners, introduce measures of controlled owners' environmental preferences, and assess both high and low expected marginal costs of carbon emissions abatement.

III. Sample and Summary Statistics

A. Sample Design

Our sample is a two-year cross-section of non-U.S. firms in the years 2023 and 2024. In these years there is extensive reporting of actual carbon emissions due to the IFRS carbon reporting mandate. To illustrate, Figure 1 plots the percentage of non-financial firms with a minimum market capitalization of $100 million from the 35 countries in our sample that report actual carbon emissions from 2010 to 2024. There is a substantial 12 percentage point increase in carbon emission reporting between 2022 and 2023, consistent with the increased pressure from IFRS. In 2023, 81% of firms report emissions and in 2024, 86% report. In contrast, in 2010 only 18% of firms reported emissions. By using data from both 2023 and 2024, we improve precision and reduce reliance on a single emissions year.[4]

We note that the use of data from earlier years would be inappropriate for our research questions due to a concern that, with voluntary disclosure, firms with excessive carbon emissions will be less likely to report. In earlier years, with actual emissions data sparse, researchers would have to make extensive use of estimated (imputed) emissions to maximize sample size. However, if certain ownership types systematically do not disclose their actual carbon emissions because they have worse carbon performance, regression estimates will then falsely suggest better carbon performance for these firms than is actually the case. Imputation is also problematic because Aswani, Raghunandan, and Rajgopal (2024) show that "estimated emissions seem to be a nearly deterministic function of size, sales growth, industry, and time" (p. 77). By using simple parameters, the commercial data vendor imputation process would not capture any independent effect of ownership and/or preferences of controlling owners on emissions.

B. Sample Construction

Our starting point is the universe of non-financial publicly traded firms with ESG data coverage by the London Stock Exchange Group (LSEG) as of 2023. We also include year 2024

[4] We do not use the two-year cross section to introduce firm fixed effects and explore within-firm analysis. As we report below, there is insufficient variation in preferences and ownership over time to support such tests.

data for these same firms if such data is available. LSEG, previously called Refinitiv, is one of several commercial ESG data providers used by both practitioners and academia. We require firms to have non-missing assets and a minimum market capitalization of $100 million and exclude firms from countries where we have less than ten observations. We exclude firms incorporated in Russia and China because political realities make it almost impossible for non-government-affiliated blockholders to establish effective control of firms. We exclude firms from the U.S. because they were not subject to the IFRS 'shock' to reporting actual carbon emissions and their reporting frequency is below 50% even in 2024 (see Figure A1 in Appendix A), making them unsuitable to use given the sample selection bias issues discussed earlier.[5]

Our full sample comprises 7,232 observations from 3,769 firms across 35 countries. Carbon emissions tests are done on a subsample of firms that report actual emissions in 2023 or 2024. This subsample comprises 6,434 firm-year observations.

C. Environmental Performance Metrics

The emissions metric used in all tests is actual total CO2-equivalent (CO2e) emissions. This metric includes all GHG emissions and converts methane and other non-CO2 emissions into carbon equivalent emissions following the Greenhouse Gas Protocol. These carbon emissions are the sum of Scope 1 emissions (direct emissions from firm-owned or controlled sources) and Scope 2 emissions (indirect emissions from the generation of purchased energy). Unlike other ESG

[5] We considered using the U.S. EPA's Greenhouse Gas Reporting Program (GHGRP) to construct a U.S. sample sufficient to conduct carbon emissions tests. We found, however, that this approach results in a relatively small, specific-industry-concentrated sample and that it captures only part of firms' company-wide emissions (e.g., it covers only Scope 1 emissions from U.S. facilities above applicable reporting thresholds). As such, the U.S. sample we can construct would not be comparable to the broad sample, company-level analyses of total Scope 1 and 2 emissions that we conduct.

metrics, this metric is highly standardized.[6] For our baseline tests, we also report results using just Scope 1 emissions.

Panel A of Table 1 reports that the median firm in our sample emits an average of 83,950 tons of CO2e per year over the 2023 and 2024 reporting years. Median Scope 1 emissions are roughly 26,000 tons, while Scope 2 emissions are roughly 30,000 tons. The large standard deviation in emissions suggests the need to control for the intensity of activity used to generate these emissions, and to control for industry differences—we do both in our tests. The pooled sample helps to address any concerns about unusual reporting of either emissions or revenues in a specific year.

D. Establishing Controlling Owners

We identify firms' ownership structures as of the year-end 2022, the year prior to the years when we measure carbon emission performance.[7] We take great efforts to identify controlled firms. We start with commercially available data sources such as Bureau van Dijk's (BvD) Orbis, LSEG, Datastream, and Worldscope, all of which provide data to trace the ownership of firms. We then manually research, categorize, and verify each firm's ultimate ownership with data from a variety of additional sources, including annual reports, internet searches, and country guides. This manual verification is essential for addressing the substantial classification errors in commercially available ownership data documented by Aminadav and Papaioannou (2020).

We construct four categories of ownership: family-controlled, government-controlled, other-controlled, and widely held. We classify a firm as family-controlled if: i) the sum of the

[6] We compare reported CO2e values across multiple data providers and find virtually identical values. This consistency of CO2e across data providers is in line with the results in Busch, Johnson, Pioch, and Kopp (2018) who report a correlation of approximately 0.99 for this metric across five data providers.
[7] Prior papers have assembled information on ownership, but even the more recent ones do not provide hand-collected cross-sections for 2022; for example, family ownership around the world is measured for the year 2002 by Masulis, Pham, and Zein (2011), and up to the year 2012 by Aminadav and Papaioannou (2020).

shares owned by the family members exceeds those of any other shareholder and is greater than 20%; ii) the sum of family stakes exceeds those of any other shareholder, is greater than 10%, and family members hold the CEO or chair position; or iii) the sum of family stakes exceeds those of any other shareholder, is greater than 10%, and the firm has multiple voting class shares. Government-controlled firms are firms where the largest shareholder is the government and they own at least 20% of the shares.[8] Other-controlled firms are firms with: i) 20% shareholdings by financial owners; ii) 20% shareholdings by multiple owners where no one owner is dominant; or iii) firms in which there is a 20% shareholder, but that owner is opaque. Widely held firms are all remaining firms that are not controlled.

In identifying controlling shareholders, our searches often involve multiple entities and voting stakes. To illustrate, consider two examples. The first is Pfeiffer Vacuum Technology AG, a small-cap German company specializing in vacuum technology. As of December 2022, it has a single class of common equity, with its largest shareholder, Pangea GmbH, holding a controlling 63% stake. Pangea is an investment vehicle owned by Busch SE, which is wholly owned by Busch GBR, which in turn is entirely owned by members of the Busch family. Through this layered structure, the Busch family exercises indirect control over Pfeiffer Vacuum Technology AG, despite the presence of seemingly separate corporate entities. Another example is Canadian Utilities Limited, a large Canadian electric utility, where we identify Sentgraf Enterprises to be the controlling shareholder based on its voting power (97.3%) rather than its economic stakes (2.5%) We identify the ultimate owner of Sentgraf to be the Southern Family, led by CEO Nancy Southern, who is the daughter of the company's founder, Ron Southern.

---

[8] We use voting stakes where available, and economic stakes when not. Further, in identifying the largest shareholder we exclude from our consideration those owners that are widely diversified asset managers (e.g., Vanguard) as they are largely passive and unlikely to contest control.

Panel A of Table 1 shows summary statistics for the ownership types in our sample. We find that 46% of firms are widely held, 42% are family-controlled, 7% are government-controlled, and 6% are other-controlled. This large frequency of controlled firms provides power for empirical tests to assess the impact of ownership structures on environmental performance. Other-controlled firms include 3% that are controlled by financial owners, 2% that have unidentified controlling owner(s), and 1% that have multiple large owners. Our empirical tests do not analyze the sub-categories of other-controlled firms given the small number of such firms in each sub-category.

In Panel B of Table 1, we report, by country, the incidence of the ownership types. There is substantial variation in how common controlled firms are around the world. For example, family ownership is highest in the Philippines and South Korea, where 89% and 83% of firms are family-controlled, respectively, and lowest in Australia, Ireland, Japan, Taiwan, and the U.K., where family firms represent less than 20%. Figure 2 provides a country map of the incidence of family control around the world.

While we have two years of data, we identify ownership once, using data from end of 2022. Previous research has found that control structures are remarkably stable over time (see, e.g. Franks et al., 2012, who report that family-controlled firms transition to being widely held at an annual rate of 0.5-0.9%). We confirm the stickiness of ownership in our sample. Using a procedure described in Appendix B, we reclassify ownership status for firms for year-end 2024. 3,709 of 3,769 firms retain their 2022 ownership status (98.4%), and the annual rate of transition away from family control is just 0.6% per year.[9]

[9] This lack of variation in ownership over time precludes the use of tests based on ownership changes because estimates with firm fixed effects would be identified only from the very small number of firms experiencing control changes.

E. Measuring Environmental Preferences of Controlling Families

The strength of environmental preferences of controlled owners is a private characteristic for which no commercial database exists. Since we cannot elicit those preferences directly via a survey or similar mechanism, we rely on the publicly observable data footprint that controlled owners leave across various languages in the public domain.

To that end, we focus specifically on the category of family-controlled firms, where we can identify a specific set of individuals with firm ownership. We proxy for a family's environmental preferences using scores from five different types of activities: i) the family's personal philanthropy towards environmental causes; ii) the family's public advocacy for environmental issues; iii) the family's participation in environmental NGOs; iv) 'green' investments in the family's personal portfolio outside the firm; and v) the family's support and contributions for environmental policies.

Prior research supports the use of indicators of such types of activities. For example, field experiments track donations and show that individuals' donations to environmental causes reflect intrinsic motivations (Alpízar, Carlsson, and Johansson-Stenman, 2008) corporate finance researchers measure directors' affiliations with nonprofit organizations and interpret these as signals of prosocial preferences (Masulis and Reza, 2015; Cai, Xu, and Yang, 2021; Kim, Minton, and Williamson, 2025). Other work shows that public advocacy for environmental issues emerges in response to weak regulation and reflects strong environmental preferences among citizens (Daubanes and Rochet, 2019). As a further example, voluntary participation is systematically related to pro-environment attitudes and values (Lowry, 1998), making NGO involvement a credible, observable indicator of underlying environmental commitment.

Tracking multiple components of preferences provides a more comprehensive measure of revealed environmental preferences than focusing on a single component. For example, a family that cared deeply about environmental outcomes might well signal that by doing all of the following: making donations to environmental causes that are reported in the news, by being vocal in advocating personally for pro environmental outcomes that generate coverage, by participating in organizations devoted to environmental improvement that are visible, by making visible investments in green firms, and by political contributions that support pro environmentally policies. Other families may prefer to speak only with money and avoid public engagement, or alternatively only with public engagement and not money.

Manually constructing scores of these preferences measures for 35 countries and across languages is infeasible. A LLM can at low cost consider a wide range of sources in various languages, which is helpful as no single source tracks all five indicators across countries.

To assess a family owner's environmental preferences, we identify the family name of the controlling owner using manual searches. To confirm our manual classification and to assess environmental preferences of the family, we then take advantage of the capabilities of two benchmark-setting LLMs. We now describe our approach in detail and provide additional information about the 'prompts' we use in Appendix C.

We use OpenAI's GPT-4o to identify controlling family names and board members because of its strong performance in parsing corporate ownership structures, and Perplexity's Sonar-pro for preference scoring because its retrieval-augmented architecture provides current citations that enable verification. While other advanced models exist, these represent state-of-the-art capabilities during our data collection from late 2023 to mid-2026. Our use of LLMs to score families' environmental preferences proceeds in two sequential stages, executed through the vendors' REST APIs. First, OpenAI's GPT-4o (trained on data through October 2023) receives a structured context block that includes the firm's name, country, and our assessment of family control, and returns the dominant family surname plus up to five family directors. Second, the family identifiers are fed into Perplexity's Sonar-pro, a search-augmented multilingual LLM that retrieves current web content and assigns a score of 0 to 10 on each of the five dimensions described above and an equally weighted sum, with a maximum of 50. The prompt expressly instructs Sonar-pro to ignore the environmental performance of the target firm to avoid tautological inference, and the model records its own confidence level and provides citation-linked sources (capped at 10 sources). We weigh all five dimensions equally since we lack theoretical or empirical guidance for differential weighting.

Although this architecture offers broad coverage with fast processing times, it raises potential concerns that we acknowledge and mitigate where possible. First, the open-web RAG approach complicates replicability as the LLM is not restricted to any static dataset. We mitigate this by publishing, in Appendix C, the exact prompts. Second, LLMs can hallucinate. By separating family identification (GPT-4o) from preference scoring (Sonar-pro) we reduce the risk that errors in one task contaminate the other. Additionally, Sonar-pro's citation-linked output enables ex-post manual verification. Table C2 in Appendix C provides examples of families with high environmental preferences to illustrate how the scores are compiled.

Third, although we try to mitigate look-ahead bias by limiting Sonar-pro's web access to items timestamped on or before December 31, 2023, there could be residual leakage. While we cannot fully verify such leakage, it is unlikely to be material, as families' environmental preferences are expected to be persistent over time. Fourth, the LLM could base its analysis primarily on the environmental choices of the firm we are analyzing. To mitigate this concern, we explicitly direct the LLM to ignore the target firm in producing its evaluation. Fifth, a generative model can always return an answer even when the underlying evidence is thin. We therefore discard observations in which the model's self-reported confidence falls below 40 (on the model's 0-100 self-reported confidence scale).

Together, these design choices yield a transparent environmental preference measure, and—subject to the caveats described—robust to the customary concerns about hallucination, look-ahead bias, and low-information noise. At the same time, we leverage the breadth, multilingual reach, and online search capability of modern LLMs to measure a latent family characteristic across our sample that is otherwise unobservable.

Panel A of Table 1 shows the average family's environmental preference score is 3.7 out of a maximum of 50; the median is two. About one third of scored families receive a zero score. The model assigns a zero when it does not find any affirmative evidence of personal environmental engagement. The maximum score is 36 for the Herlins (Finland). We use these scores to construct low and high preference indicators to make sure inferences are based on large differences in preferences, to give weight to all five dimensions of environmental preferences, and for better distributional properties. We classify families with scores above 2 (i.e., above the median) as high E family, and those with scores at or below two as low E family. The LLM cannot reliably assign a score for some families (or has low confidence that it has established the environmental preferences of the family with precision)—we label them as inconclusive E family. We posit that family firms categorized as inconclusive are unlikely to have high environmental preferences. As a share of family-controlled firms, 32% of family firms are classified as low E family, 25% are high E famil*y*, and 43% are inconclusive E family.

To illustrate Sonar-pro's assessment of families, we consider several family firms in our sample here and in Appendix C. An example of a firm classified as low E family is Grupo México, a Mexican mining and metals group controlled by the Larrea family. Sonar-pro assigns the family the lowest possible score of 0, with high self-reported confidence, reasoning that "publicly available evidence as of December 2023 shows little to no direct personal environmental engagement by the Larrea family's controlling individual, […]. The available material focuses on […] Grupo México's controversies, not on personal philanthropy, advocacy, NGO roles, green investments, or policy support for environmental causes." Importantly, the LLM explicitly ignores the corporate environmental activities of Grupo México in its assessment.

An example of a firm classified as high E family is Kone, a Finnish elevator manufacturer controlled by the Herlin family, which receives the highest score in our sample, 36 out of 50. The LLM's reasoning is that "the Herlin family […] shows strong personal environmental preferences. Ilkka Herlin has repeatedly supported Baltic Sea and climate-related work, including founding and serving in the Elävä Itämeri Foundation and working on carbon-sequestration agriculture at […]. [Their] foundation has an explicit emissions-cutting environmental program." The assessment draws on family-specific, verifiable activity outside the firm, foundations, personal projects, and public advocacy, rather than on Kone's corporate record.

Finally, as noted already, our models cannot classify every family's environmental preferences. As an example, for the Chang and Choi families that control Korea Zinc, the model reports low confidence and hence, the family preference is labeled as inconclusive. Appendix Table C2 shows the model's full assessments for these and other families across the high, low, and inconclusive classifications.

We assess the reliability and stability of the preference scores in several ways: test-retest correlations across identical runs; sensitivity to semantically equivalent prompt rewordings; agreement with an independent re-estimation by a different vendor's model; and stability of the classification when the information cutoff is moved from December 2023 to December 2022. The classification that enters our regressions is highly stable across runs and vintages, and our central results are unchanged under the alternative measures. We report further details of these assessments in Appendix C.

F. Measuring Expected Marginal Costs of Improving Carbon Emissions

We use two approaches to identify the expected marginal abatement costs firms face to improve their carbon emissions, exploiting variation in expected costs at the country level and at the industry level. The results of these tests should be viewed as suggestive, given that it is impossible to construct marginal abatement cost estimates at the firm level for all of the firms in our sample.

At the country level, we form two subsamples based on a country's Climate Change Performance Index (CCPI) score. The CCPI, produced by Germanwatch, assesses each country's performance in four categories, capturing differences in climate protection efforts relative to international standards. The index is based on GHG Emissions (40% of the overall ranking), Renewable Energy (20%), Energy Use (20%) and Climate Policy (20%). Around 80% of the CCPI score is based on quantitative data from the International Energy Agency (IEA), PRIMAP, the Food and Agriculture Organization, and the national GHG inventories submitted to the United Nations Framework Convention on Climate Change.[10] Twenty percent of the CCPI is based on

[10] According to Germanwatch, "the CCPI quantitative categories are each defined by four indicators: Current Level, Past Trend, Well-Below-2°C Compatibility of the Current Level, and Well-Below-2°C Compatibility of the Countries'

qualitative indicators of National Climate Policy and International Climate Policy assessed annually with a comprehensive questionnaire from over 450 climate and energy policy experts from the evaluated countries.

The CCPI is a comprehensive measure capturing the effects of policies that penalize emissions, such as tax policy, as well as proactive incentives for emission abatement. Low CCPI countries have less emphasis on policies and standards to reduce carbon emissions. Firms in these countries likely have more low-cost options to further improve their emissions performance. The CCPI has been used extensively in prior carbon emission research (see, for example, Bolton and Kacperczyk, 2023, and Eskildsen, Ibert, Jensen, and Pedersen, 2025).

A country's CCPI score has a possible range from zero (worst) to 100 (best). We group firms in subsamples based on the sample median score of 60, with all countries below this score identified as low CCPI countries.[11] CCPI scores are highest in Denmark, and lowest in South Korea. The CCPI metric has a low correlation with GDP per capita. For example, Australia, Canada, and Austria are low CCPI countries that have relatively high levels of GDP per capita.[12]

To help to validate the CCPI as a carbon abatement cost measure, we assess its correlation with two prominent, but narrower, country-level metrics the OECD uses to gauge the relative costs of improving carbon emissions across countries. First, the OECD calculates and uses effective carbon rates as a summary statistic to measure the extent to which governments have used the tax system to incentivize firms in their country to consider their carbon emissions. The Effective Carbon Rates statistic "summarizes the price signals from ETSs (emission trading systems), carbon

---

2030 Target." It assesses "the extent to which each country is taking appropriate action in the areas of Emissions, Renewable energy and Energy use in order to achieve the climate goals set in Paris."

[11] CCPI scores are not available for Israel, Peru, and Singapore.

[12] We also conduct robustness tests of this country-level measure by constructing subsamples based on a country's score on just the GHG emissions category (that has a 40% weight in the overall CCPI score). This is based on assessments of countries' GHG emissions per capita, trends in emissions, and current emissions and reduction targets compared to a well-below -2°C compatible pathway.

taxes and fuel excise taxes, and is expressed in EUR/tCO2e." This is available at the country level and for a subset of industries they define as 'energy use sectors' (i.e., industries that account for the greatest energy use).[13]

In our sample of countries, the effective carbon rates range from 0.5 to 151 Euros/ tCO2e in 2023.[14] Consistent with the contention that high CCPI countries have high marginal costs of abatement, there is a positive and significant correlation between effective carbon rates and CCPI scores of 0.43. In high CCPI countries the median (mean) effective carbon rate is 89.7 (77.2) and it is 26.0 (41.6) in low CCPI countries.[15] We next build cross-country variation by looking only at a country's energy use sectors. We find a similar positive and significant correlation between the CCPI and effective carbon rates in all 'energy use sectors' (0.43), and in two of the most important energy use sectors 'industry' (0.43) and 'road transportation' (0.36).[16]

Environmental patenting also relates to carbon abatement costs, albeit less directly. The idea here is that firms are more likely to patent environmentally efficient technologies if a country provides greater incentives to reduce emissions. The OECD also provides country level statistics on patenting, and we focus on environmental patents as a percentage of all technical patents (thus accounting for a country's overall patenting propensity).[17] This environmental patenting metric ranges from 4.2% to 22.3% in our sample of countries. We find a positive and statistically

[13] OECD, Carbon Pricing and Energy Taxation Database, OECD Data Explorer (accessed in May 2026).
[14] These rates capture how "Countries deploy taxes on energy use, carbon taxes and greenhouse gas emissions trading systems in view of policy objectives related to climate change, public revenue raising, energy affordability and cost of living, energy security and competitiveness." OECD reports data for 33 out of 35 countries in our sample.
[15] These coefficients are statistically different with *p*-values of 0.026 (0.023). We further note that results are similar using an alternative OECD metric of net effective carbon tax rates, which in addition takes into account fossil fuel. By this metric, the median (mean) net effective carbon rate is 89.1 (71.4) in high CCPI countries and the median (mean) net effective carbon rate in low CCPI countries is 26.0 (34.5). These differences are also statistically significant.
[16] The OECD defines energy-use sectors as industry, road transportation, electricity, buildings, agriculture and fisheries and off-road transportation.
[17] OECD, STI Micro-data lab: Intellectual Property Database, OECD Data Explorer (accessed in April 2026).

significant correlation between this measure and CCPI scores of 0.21. The median (mean) percentage of environmental patents is greater in high CCPI than low CCPI countries, however this difference is not significantly different.

Next, we identify marginal emission abatement costs using variation in expected costs across industries. Regulators and society are likely to hold owners accountable for a firm's emissions and will pay attention to both total emissions and direct emissions. We conjecture, however, that they will be particularly interested in a firm's direct emissions as these are largely determined by an owner's operating choices. In industries with a high proportion of direct emissions, firms will arguably have exhausted their low-cost carbon abatement opportunities.

We split our sample into two subsamples based on the intensity of direct carbon emissions (Scope 1) in total carbon emissions (sum of Scope 1 and Scope 2) for the seventy-two industries as defined by the Sustainability Accounting Standards Board (SASB) in our sample.[18] We classify an industry as high Scope 1 if at least 51% of an industry's total carbon emissions are Scope 1 emissions. High Scope 1 industries have vastly larger total emissions, with average log total CO2e emissions in tons of 12.14 compared to 10.38 for low Scope 1 industries (see Table 1). For low Scope 1 industries, total emissions themselves are relatively small, and direct emissions are even smaller. In these industries, firms are likely to have escaped scrutiny, leaving them with a larger supply of low-cost carbon abatement opportunities available if they want to implement them.

[18] We also examine industry splits based solely on direct emissions, irrespective of total emissions. As described in the next section, the pattern of results is similar.

IV. Results

A. Emissions Disclosure and Ownership Type

As reported in Table 1, 89% of firm-year observations that meet our sample selection criteria disclose their carbon emissions. With this high level of reporting, any selective-disclosure biases are likely to be small. To validate this conjecture, we test whether the disclosure of carbon emissions is linked to ownership type using the full sample in Table 2.

To isolate the impact of ownership on carbon emissions reporting (and actual carbon emissions in the tables that follow) all regressions include firm size (log of assets), cash, asset tangibility, leverage, profitability, and dual class shares as control variables. We include firm size as prior literature has shown it to be related to ownership structures, and larger firms may be subject to more external pressures. Hong, Kubik, and Scheinkman (2012) suggest that financial slack helps explain the adoption of sustainability-oriented policies. We thus include cash, asset tangibility, and leverage to capture credit constraints and profitability to capture the impact of performance. We control for expected private benefits of control using an indicator for dual class shares to capture the wedge between control rights and cash flow rights. Given the substantial variation in carbon emissions across countries and industries, we include country and industry (72 SASB industry codes) fixed effects along with year fixed effects. We cluster standard errors by country.[19]

In Table 2 we find a negative coefficient for family control and government control, but both are insignificant. Thus, in our sample period there is no evidence of selection bias related to ownership type.

[19] Because our sample contains 35 countries, we also assess the robustness of our inferences using a wild-cluster-bootstrap at the country level. The resulting bootstrap p-values are very similar to those we obtain using our conventional country-clustered standard error approach and do not change the interpretation of our main findings.

B. Controlled Firms, Carbon Emissions, and Preferences

We begin with baseline regressions examining the relationship between ownership type and carbon emissions. Then we examine the relationship between environmental preferences and carbon emissions, with preferences available for family-controlled firms. In an extension, we examine whether the expected marginal costs to improve carbon emissions also play a role. In all tests we estimate models using the following baseline specification:

$$Log(Emissions_i) = \alpha + \beta' X_i + \gamma' Y_i + \Lambda + \varepsilon_i , \quad (1)$$

where the dependent variable is either the log of total CO2e emissions or the log of Scope 1 CO2e emissions, scaled by revenue or unscaled, for firm i. As independent variables, X_i are ownership structure indicator variables for ownership by family, government, and other opaque, omitting widely held firms as the baseline category, Y_i is a set of firm-level controls, and Λ are individual country, industry, and year fixed effects. All right-hand variables are lagged by one year. We use logs for our dependent variables to obtain better distributional properties and to reduce the impact of outliers.

We report results using both emissions intensity (scaled emissions) and raw emissions because there is no consensus as to how to appropriately measure emissions. Aswani et al. (2024) argue in favor of emissions intensity as the appropriate metric,[20] and further suggest that revenue-scaled carbon emissions is preferred to asset-scaled emissions given the stronger correlation between emissions and revenue.[21] When unscaled emissions are the dependent variable, we also

[20] They note that using unscaled emissions 'is analogous to using net income rather than ratios such as return on assets (ROAs) to measure financial performance,' that 'the ratio of emissions to net sales the most commonly used metric in practice', and that it 'better captures a firm's emissions performance by avoiding mechanical correlations with firm size' (p. 78).

[21] Aswani et al. (2024), using a sample of U.S. firms, report a correlation of log Scope 1 emissions and log sales of 0.70 (compared to 0.46 for log assets). Focusing on Scope 2 emissions the correlations are 0.847 for log sales and 0.548 for log assets.

control for firms' revenue as emissions likely depend heavily on the intensity of firm activity which is reflected in sales. We report results for total Scope 1 and 2 emissions in all specifications, as these are the focus of the ISSB's sustainability standards. For our baseline tests we also report results using Scope 1 emissions alone, as these are directly controlled by owners and have been the primary focus of some of the prior carbon emissions research.[22]

In Table 3 we report our first results. We find that no category of controlled firm has statistically significantly lower carbon emissions than the widely held firm, conditional on controls and fixed effects. In family-controlled firms, where theory suggests the greatest likelihood of strong preferences for *clean,* we find the opposite. There is a positive and significant coefficient estimate on family in all specifications. These results indicate that family firms emit more carbon than widely held firms, and imply that allowing shareholders to pursue welfare maximization, as in Hart and Zingales (2017), is unlikely to mitigate unpriced environmental externalities.

For other-controlled firms, we also find a positive coefficient, that is significant when considering Scope 1 and 2 emissions. This positive association is consistent with our priors that these firms lack environmental preferences. Given that such firms represent a small portion of our sample, and are heterogeneous, we hesitate to draw conclusive inferences from these coefficients here and in the rest of our tables. For government-controlled firms, we find an insignificant coefficient, also consistent with our priors.

Regarding control variables, we find that emissions are greater when firms are larger, when they have more tangible assets, and in specifications looking at absolute emissions, when they have a greater level of activity measured by revenues. Firms have lower emissions when there is

[22] See, for example, Shive and Forster (2020) and Bolton and Kacperczyk (2023), who use scaled and raw Scope 1 emissions.

more financial slack, with negative coefficients on both cash and profitability. The coefficient estimates on dual class shares are positive, as expected, but insignificant.

In Table 4, we test whether the environmental preferences of controlling shareholders are related to carbon emissions, a key component of our analysis. Using the LLM-identified environmental preferences, we split family firms into the previously described three groups, firms controlled by families with low environmental preferences (low E family), firms controlled by families with high environmental preferences (high E family), and firms controlled by families with unknown environmental preferences (inconclusive E family).

The primary takeaway from Table 4 is that the Hart and Zingales (2017) idea that a change in focus from financial value maximization to shareholder welfare maximization results in different carbon outcomes appears to be true. We find that the environmental preferences of controlling shareholders are indeed associated with carbon emissions, but mostly in ways that indicate worse outcomes. In firms controlled by families with low environmental preferences, total emissions are 23%-25% higher than those of widely held firms.[23] Similarly, family firms labeled inconclusive have emissions 19-20% higher than widely held firms. In contrast, emissions are no different from those of widely held firms when firms are controlled by families with high environmental preferences. Results are similar in our column 3 and 4 tests that use Scope 1 carbon emissions. As reported at the bottom of all four columns of Table 4, low environmental preference families have significantly higher emissions than high preference families.

[23] All economic magnitudes are calculated as $e^{\text{coefficient}} - 1$.

C. Expected Marginal Costs, Preferences, and Emissions

In this section, we extend our analysis to consider the expected marginal costs of becoming *clean*. The evidence we present is suggestive in nature given the difficulty of measuring marginal costs.

Recall that all controlled firms, and particularly family and other controlled firms, face a potential private benefit drift towards *dirty.* Strong environmental preferences are a factor that could shift controlled firms to be *clean,* but the willingness of a controlling owner to act on their environmental preferences should depend on how expensive it is to make their firm *cleaner*. When the costs of becoming *cleaner* are high, preferences must be extremely strong to override the drift towards *dirty*.

We begin our tests at the country level (Table 5, Panel A) and use the CCPI index that captures differences in climate protection efforts relative to international standards. In high CCPI countries, we expect that environmental preferences will have minimal impacts on carbon emissions given the high cost of becoming incrementally *cleaner*. Further, we expect no significant differences in carbon emissions performance when comparing owners with different environmental preferences. In contrast, in low CCPI countries, there is room for preferences to have an impact, given the lower cost of improving emissions. Hence, we expect differences in the impact of relative environmental preferences to be greater in low CCPI countries.

We then conduct tests at the industry level (Table 5, Panel B), splitting firms into two industry subsamples based on the importance of direct (Scope 1) emissions as a fraction of total industry emissions. If direct emissions are relatively unimportant, regulator/stakeholder pressures to reduce emissions will likely be muted, and low-cost options to improve emissions likely remain (if direct emissions are relatively important, substantial reduction measures likely have been already taken, making further reductions relatively expensive).

Across both panels, the results reveal a consistent pattern. With high expected marginal costs of reducing emissions, either because firms operate in high CCPI countries or in industries where direct emissions are relatively important, family firms do not have lower emissions than widely held firms, even when controlling families have high environmental preferences. To illustrate, we focus on column 1. The coefficient for low E families is positive and implies 24% (Panel A) to 46% (Panel B) higher emissions than widely held firms. The coefficient for high E families is also positive and implies 14% (Panel A) to 29% (Panel B) higher emissions than widely held firms. In three of the four cases the coefficient is statistically significant.

In contrast, when the marginal costs of becoming *cleaner* are expected to be low, the results show that family firms with high environmental preferences have *lower* emissions, while those with low environmental preferences are generally indistinguishable from widely held firms. Column 3 illustrates these findings. The coefficient for high E families is negative and implies 16% (Panel A) to 20% (Panel B) lower emissions than widely held firms, with the Panel B coefficient being statistically significant. The coefficient for low E families is positive, close to zero, and insignificant.

Taken together, our results provide suggestive evidence consistent with controlling owners' willingness to act on their environmental preferences depending on the expected costs to make their firm *cleaner*. When expected costs to be even more *clean* are high, even strong pro-environmental preferences are not associated with lower carbon emissions. Only when the expected costs of becoming *cleaner* are low, our results show a positive association between high environmental preferences and better carbon performance.

V. Robustness

In this section, we provide robustness of our main findings: i) we discuss possible limitations of LLM-based environmental preference scores and how we address them, and ii) we offer additional results in measuring expected marginal costs in reducing carbon emissions.

A. Environmental Preferences

Our tests use LLM-based environmental preference scores that are not based on standardized data sources, such as earnings calls or mandatory disclosures. A potential concern is that the LLM aggregates information from diverse sources in ways that may introduce systematic biases in how it accesses information. For example, the LLM may exhibit language-based biases if training content or web retrieval capabilities are predominantly in English, potentially compromising the quality of assessments for families operating in non-English contexts. The LLM could also disproportionately weight sources from wealthier countries or based on wealthier families or larger firms. For example, wealthier families could attract more (positive) media attention, and the LLM might assign such families a higher score—not because of their stronger environmental preferences. While we cannot rule out all types of bias from the use of a LLM-based measure, in this section we construct an alternative family environmental preference score that controls for potential biases such as language and test whether our results hold using this alternative environmental preference score.

In Panel A of Table 6, we regress environmental preference scores on country, family, and firm attributes. We consider whether English is the native language in the firm's headquarter country, country GDP per capita, the wealth status of the family (using the 2022 Forbes list of billionaires), and firm size. The results show that larger firms exhibit higher preference scores while wealthier countries are, contrary to our concerns, associated with lower scores. Families on the Forbes billionaire list also receive higher scores, consistent with wealthier families leaving a larger public footprint. The weak relationship between English speaking countries and preference scores suggests adequate multilingual performance, though we cannot rule out subtle biases.[24]

To address concerns about language and other potential determinants of the preference scores we have employed so far, we create alternative, orthogonalized environmental preference scores using the residuals of the regressions of Panel A of Table 6. Using these orthogonalized scores we then, again, sort family firms into low E (negative residuals) and high E (positive residuals) families. In Panel B of Table 6, we replicate the Table 4 regressions using these alternative, orthogonalized environmental preference scores based on the column 2 estimates from Panel A (results are similar when using column 1 residuals, unreported for brevity). The results confirm our main findings—family firms never have lower emissions performance than widely held firms, and low family environmental preferences are associated with higher carbon emissions relative to widely held firms. Thus, the negative relation between preferences and emissions persists after stripping out the effects of English language, country income, family wealth, and firm size.

B. Alternative CCPI Score

We also test an alternative CCPI score as a country level indicator for expected marginal abatement costs. Specifically, we use only the "emissions" category from the CCPI score to sort firms into low and high CCPI groups, by country. Table A2 in Appendix A replicates the Table 5, Panel A specifications, and the results are substantially similar. Again, family firms have greater emissions in high CCPI emissions countries, regardless of the environmental preferences of the

[24] We note that the LLM provider indicates that their model does not focus exclusively on English-language content, and its web retrieval capabilities are not restricted by language.

family owner. In countries with lower CCPI emissions scores, high E families exhibit approximately 24% lower emissions than widely held firms (significant at the 10% level).

VI. Conclusion

A key takeaway of our paper is that the Hart and Zingales (2017) hypothesis—carbon emissions are different if one allows for shareholder utility maximization—is true. When owners have the freedom to direct firms to satisfy their non-pecuniary preferences, this does influence environmental outcomes. Families having strong environmental preferences are associated with meaningful improvements in carbon emissions when the expected marginal cost of reducing emissions is relatively low. But the aggregate effect of allowing for shareholder utility maximization is associated with worse carbon outcomes as in all other scenarios controlled firms' carbon emissions performance is no better—and often worse—than that of widely held firms. Our evidence suggests that there simply aren't enough families with strong environmental preferences and/or the strength of their preferences are not high enough to overcome either the high marginal cost of further carbon emission reduction or the pecuniary private benefit drift that affects controlled firms.

Our results suggest that shifting from strict financial value maximization to broader shareholder welfare maximization is unlikely to deliver significant environmental gains without stronger regulatory support. The few exceptions we observe—where family-controlled firms perform better in low-cost contexts—underscore that voluntary improvements tend to arise only under narrow conditions. Taken together, these insights caution against relying solely on controlling shareholders to address negative environmental externalities. Instead, they highlight the need for robust policies and institutions to encourage *cleaner* corporate behavior.

Our findings are salient given the widespread prevalence of controlled firms globally and the fact that global emissions growth now comes primarily from developing countries (Copeland, Shapiro, and Taylor, 2022) where controlled firms are the most common (Lins, 2003).

References

Alpízar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman, 2008, Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica, *Journal of Public Economics* 92, 104-1060.

Aminadav, Gur, and Elias Papaioannou, 2020, Corporate control around the world, *Journal of Finance* 75, 1191-1246.

Anderson, Ronald C., and David M. Reeb, 2003, Founding-family ownership and firm performance: evidence from the S&P 500, *Journal of Finance* 58, 1301-1328.

Aswani, Jitendra, Aneesh Raghunandan, and Shiva Rajgopal, 2024, Are carbon emissions associated with stock returns?, *Review of Finance* 28, 75-106.

Bennedsen, Morten, and Joseph P. H. Fan, 2014, *The Family Business Map: Assets and Roadblocks in Long Term Planning*, INSEAD Business Press, Palgrave Macmillan, London.

Bolton, Patrick, and Marcin Kacperczyk, 2023, Global pricing of carbon-transition risk, *Journal of Finance* 78, 3677-3754.

Borsuk, Marcin, Nicolas Eugster, Paul-Olivier Klein, and Oskar Kowalewski, 2024, Family firms and carbon emissions, *Journal of Corporate Finance* 89, 102672.

Busch, Timo, Matthew Johnson, Thomas Pioch, and Matthias Kopp, 2018, Consistency of corporate carbon emission data, *Working Paper*, University of Hamburg Report WWF Deutschland, Hamburg.

Cai, Ye, Jin Xu, and Jun Yang, 2021, Paying by donating: Corporate donations affiliated with independent directors, *Review of Financial Studies* 34, 618-660.

Choi, Chanyeol, Jihoon Kwon, Jaeseon Ha, Hojun Choi, Chaewoon Kim, Yongjae Lee, Jy-yong Sohn, and Alejandro Lopez-Lira, 2025, FinDER: Financial dataset for question answering and evaluating retrieval-augmented generation, *Proceedings of the 6th ACM International Conference on AI in* Finance, 638-646.

Copeland, Brian R., Joseph S. Shapiro, and M. Scott Taylor, 2022, Globalization and the environment, *Handbook of International Economics* 5, 61-146.

Daubanes, Julien, and Jean-Charles Rochet, 2019, The rise of NGO activism, *American Economic Journal: Economic Policy* 11, 183-212.

de Kok, Ties, 2025, ChatGPT for textual analysis? How to use generative LLMs in accounting research, *Management Science* 71, 7888-7906.

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

Duchin, Ran, Janet Gao, and Qiping Xu, 2025, Sustainability or greenwashing: Evidence from the asset market for industrial pollution, *Journal of Finance* 80, 699-754.

Dyck, Alexander, Karl V. Lins, Lukas Roth, Mitch Towner, and Hannes F. Wagner, 2023, Renewable governance: good for the environment?, *Journal of Accounting Research* 61, 279-327.

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

Dyck, Alexander, and Luigi Zingales, 2004, Private benefits of control: An international comparison, *Journal of Finance* 59, 537-600.

Eskildsen, Marc, Markus Ibert, Theis Ingerslev Jensen, and Lasse Heje Pedersen, 2025, In search of the true greenium, *Journal of Financial Economics*, forthcoming.

Franks, Julian, Colin Mayer, Paolo Volpin, and Hannes. F. Wagner, 2012, The life cycle of family ownership: International evidence, *Review of Financial Studies* 25, 1675-1712.

Hart, Oliver, and Luigi Zingales, 2017, Companies should maximize shareholder welfare not market value, *Journal of Law, Finance, and Accounting* 2, 247-275.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, 2019, Firm-level political risk: Measurement and effects, *Quarterly Journal of Economics* 134, 2135-2202.

Hong, Harrison, Jeffrey D. Kubik, and Jose Scheinkman, 2012, Financial constraints on corporate goodness, *Working Paper*, Columbia University.

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

Iliev, Peter, and Lukas Roth, 2023, Director expertise and corporate sustainability, *Review of Finance* 27, 2085-2123.

International Organization of Securities Commissions (IOSCO), 2025, Artificial intelligence in capital markets: Use cases, risks, and challenges, IOSCO report CR/01/2025.

Kim, Sehoon, Bernadette A. Minton, Rohan Williamson, 2025, Climate boards: Do natural disaster experiences make directors more prosocial?, *Working Paper*, University of Florida.

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

La Porta, Rafael, Florencio Lopez-De-Silanes, and Andrei Shleifer, 1999, Corporate ownership around the world, *Journal of Finance* 54, 471-517.

Li, Yinheng, Shaofei Wang, Han Ding, and Hang Chen, 2024, Large language models in finance: A survey, *Proceedings of the Fourth ACM International Conference on AI in Finance*, 374-382.

Lins, Karl V., 2003, Equity ownership and firm value in emerging markets, *Journal of Financial and Quantitative Analysis* 38, 159-184.

Lins, Karl V., Lukas Roth, Henri Servaes, and Ane Tamayo, 2024, Sexism, culture, and firm value: evidence from the Harvey Weinstein scandal and the #MeToo movement, *Journal of Accounting Research* 62, 1989-2035.

Lowry, Robert C., 1998, Religion and the demand for membership in environmental citizen groups, *Public Choice* 94, 223-240.

Masulis, Ronald W., and Syed W. Reza, 2015, Agency problems of corporate philanthropy, *Review of Financial Studies* 28, 592-636.

Masulis, Ronald W., Peter Kien Pham, and Jason Zein, 2011, Family business groups around the world: Financing advantages, control motivations, and organizational choices, *Review of Financial Studies* 24, 3556-3600.

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

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2022, Dissecting green returns, *Journal of Financial Economics* 146, 403-424.

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

Shive, Sophie, and Margaret Forster, 2020, Corporate governance and pollution externalities of public and private firms, *Review of Financial Studies* 33, 1296-1330.

Shleifer, Andrei, and Robert W. Vishny, 1997, A survey of corporate governance, *The Journal of Finance* 52, 737-783.

Siano, Federico, 2025, The news in earnings announcement disclosures: Capturing word context using LLM methods, *Management Science* 71, 9831-9855.

Wang, Shuting, Jiejun Tan, Zhicheng Dou, and Ji-Rong Wen, 2024, OmniEval: An omnidirectional and automatic RAG evaluation benchmark in financial domain, *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, 5737-5762.

Zhao, Yiyun, Prateek Singh, Hanoz Bhathena, Bernardo Ramos, Aviral Joshi, Swaroop Gadiyaram, and Saket Sharma, 2024, Optimizing LLM based retrieval augmented generation pipelines in the financial domain, *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 6: Industry Track),* 279-294.

Figure 1
Percent of Firms Reporting Carbon Emissions

This figure reports the percentage of firms that report CO2-equivalent emissions from 2010 to 2024 for i) the population of LSEG firms, and ii) our sample of firms. The population of LSEG firms includes all non-financial firms with a minimum market capitalization of $100 million, incorporated in one of our sample countries, and includes between 3,086 (in 2010) and 4,510 firms (in 2021). All data are from LSEG.

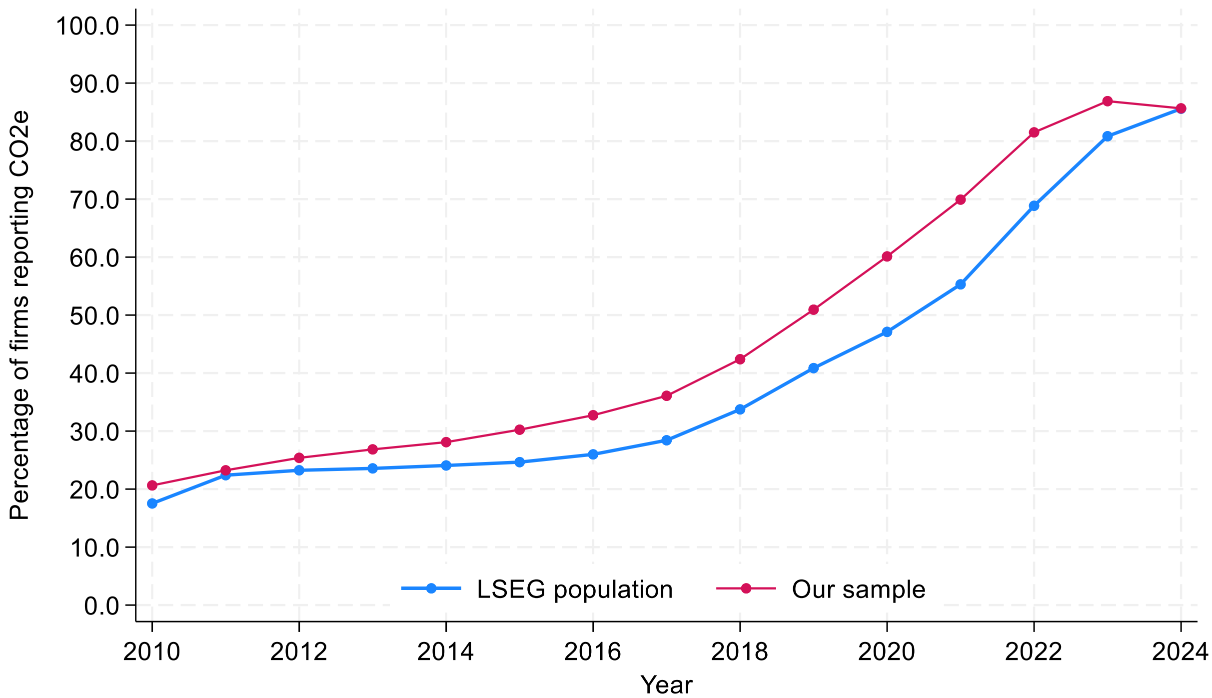


Figure 2
Family Control Around the World

This figure reports the incidence of family control for the 35 countries in our sample. Family control is manually verified for each firm and defined as follows: we classify a firm as family-controlled if the sum of the shares owned by family members is greater than 20%, family members own at least 10% of the shares and have a position of CEO/Chair, or family members own at least 10% of the shares and the company has multiple voting share classes. We also require that family members own more shares than any other shareholder.

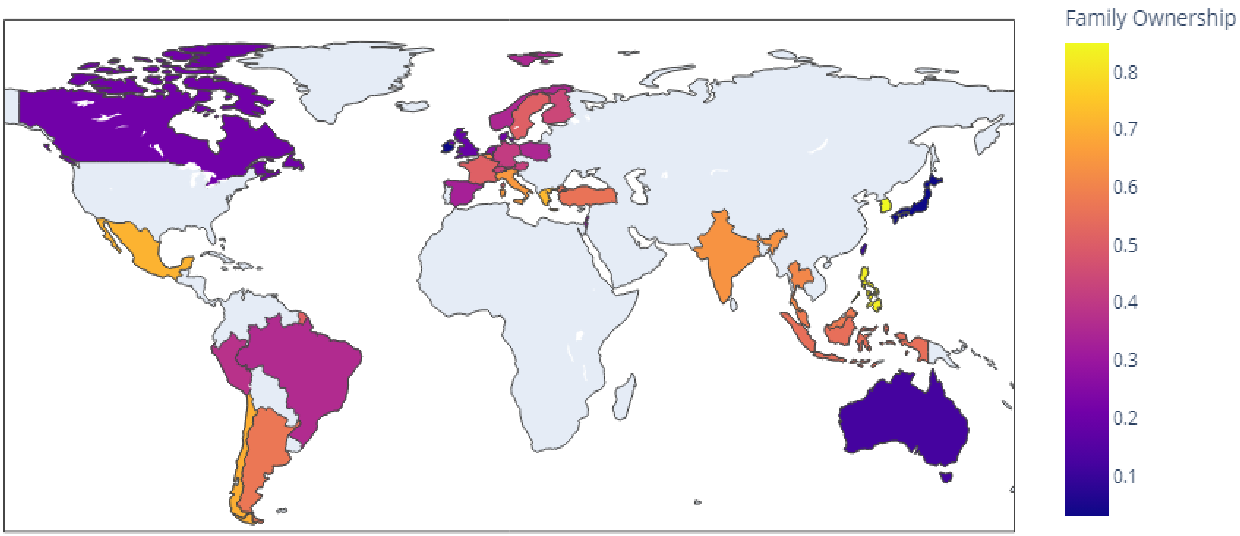


Table 1
Summary Statistics

This table shows summary statistics. Firms are classified into low and high CCPI groups based on their country's overall Climate Change Performance Index (CCPI) score, with the median score used as the cutoff. Similarly, firms are grouped into low and high Scope 1 industries depending on whether the industry's total Scope 1 emissions account for less (greater) than 51% of the industry's total Scope 1 and 2 CO2-equivalent emissions (based on SASB Industry Classifications). The sample covers the years 2023 and 2024. Ownership categories do not add to 100% due to rounding. All variables are described in Appendix Table A1.

Panel A: Summary Statistics

	Mean	Median	SD	N
A. Environmental Performance				
Firm Reports CO2e	0.89	1.00	0.32	7,232
Scope 1 and 2 Emissions				
CO2e	1,699,683	83,950	7,377,957	6,444
Log (CO2e)	11.28	11.35	2.80	6,434
Log (CO2e / Revenue)	3.73	3.61	2.15	6,434
Log (CO2e), Low CCPI Countries	11.88	11.99	2.57	3,064
Log (CO2e), High CCPI Countries	10.72	10.71	2.88	3,205
Log (CO2e), Low Scope 1 Industries	10.38	10.50	2.36	3,155
Log (CO2e), High Scope 1 Industries	12.14	12.29	2.92	3,279
B. Ownership				
Family-controlled	0.42	0.00	0.49	7,232
Government-controlled	0.07	0.00	0.25	7,232
Other-controlled	0.06	0.00	0.24	7,232
Widely Held	0.46	0.00	0.50	7,232
C. Family Owners' Environmental Preferences				
Family E Preference	3.73	2.00	5.24	1,717
Within Low E Family	0.81	0.00	0.90	968
Within High E Family	7.51	5.00	6.06	749
Within Inconclusive E Family	n/a	n/a	n/a	1,282
Low E Family	0.32	0.00	0.47	2,999
High E Family	0.25	0.00	0.43	2,999
Inconclusive E Family	0.43	0.00	0.49	2,999
D. Other Firm Characteristics				
Total Assets (in $ million)	9,642	2,379	26,353	7,232
Log (Total Assets)	21.64	21.59	1.64	7,232
Log (Revenue)	7.37	7.38	1.77	7,232
Cash	0.15	0.11	0.14	7,232
Tangibility	0.31	0.27	0.22	7,232
Leverage	0.26	0.24	0.39	7,232
Profitability	0.04	0.04	0.15	7,232
Dual Class Shares	0.16	0.00	0.36	7,232

Panel B: Summary Statistics by Country

Country	Ownership				Family Owners' Environmental Preferences			CCPI	Average Total Assets (in $ million)	N
	Family	Government	Other	Widely Held	Low	High	Unknown			
Argentina	0.67	0.07	0.07	0.20	0.50	0.50	0.00	45.4	4,807	30
Australia	0.18	0.00	0.04	0.78	0.20	0.10	0.70	45.7	4,337	442
Austria	0.42	0.26	0.11	0.21	0.50	0.25	0.25	58.2	7,859	38
Belgium	0.65	0.06	0.08	0.21	0.04	0.09	0.87	55.0	11,013	71
Brazil	0.42	0.07	0.15	0.37	0.24	0.12	0.64	61.7	10,302	202
Canada	0.26	0.02	0.10	0.62	0.41	0.39	0.20	31.5	7,748	502
Chile	0.71	0.14	0.00	0.14	0.10	0.05	0.85	68.7	8,919	56
Denmark	0.22	0.07	0.25	0.47	0.60	0.10	0.30	75.6	7,525	92
Finland	0.38	0.11	0.11	0.39	0.24	0.34	0.41	61.1	4,003	107
France	0.56	0.08	0.07	0.29	0.44	0.39	0.17	57.1	25,468	272
Germany	0.48	0.07	0.09	0.36	0.12	0.05	0.83	65.8	19,024	346
Greece	0.68	0.27	0.00	0.05	0.40	0.60	0.00	60.3	5,421	37
India	0.66	0.12	0.02	0.20	0.34	0.35	0.30	70.3	3,520	918
Indonesia	0.59	0.18	0.08	0.16	0.23	0.37	0.40	57.2	4,555	102
Ireland	0.15	0.00	0.00	0.85	0.33	0.00	0.67	51.4	14,776	82
Israel	0.44	0.07	0.10	0.39	0.15	0.00	0.85	n/a	4,239	61
Italy	0.69	0.21	0.03	0.07	0.08	0.17	0.75	50.6	10,146	174
Japan	0.09	0.01	0.03	0.87	0.50	0.16	0.34	42.1	17,487	825
Luxembourg	0.50	0.00	0.08	0.42	0.00	0.17	0.83	65.1	9,103	48
Malaysia	0.65	0.17	0.06	0.13	0.68	0.13	0.19	38.6	3,227	240
Mexico	0.68	0.00	0.00	0.32	0.10	0.05	0.85	55.8	7,366	114
Netherlands	0.25	0.00	0.17	0.58	0.54	0.31	0.15	70.0	16,428	106
Norway	0.38	0.09	0.13	0.40	0.38	0.38	0.24	67.5	6,476	110
Peru	0.33	0.22	0.00	0.44	1.00	0.00	0.00	n/a	2,783	36
Philippines	0.89	0.04	0.00	0.08	0.26	0.64	0.11	70.7	8,932	53
Poland	0.39	0.39	0.09	0.13	0.78	0.22	0.00	44.4	8,913	46
Singapore	0.50	0.28	0.00	0.22	0.59	0.26	0.15	n/a	8,310	108
South Korea	0.83	0.06	0.01	0.09	0.23	0.27	0.50	30.0	17,560	280
Spain	0.43	0.09	0.13	0.35	0.13	0.26	0.61	63.4	15,680	108
Sweden	0.48	0.02	0.07	0.44	0.45	0.23	0.31	69.4	3,434	362
Switzerland	0.47	0.06	0.07	0.41	0.38	0.25	0.37	61.9	9,991	236
Taiwan	0.16	0.05	0.04	0.76	0.43	0.04	0.52	36.9	8,028	296
Thailand	0.64	0.17	0.02	0.17	0.41	0.27	0.32	61.4	4,859	222
Turkey	0.66	0.20	0.09	0.05	0.16	0.62	0.22	43.8	7,670	112
U.K.	0.19	0.00	0.10	0.71	0.06	0.02	0.92	62.4	8,950	669

Table 2
Controlled Firms and GHG Emissions Disclosure

This table shows regression estimates of firms' GHG reporting on ownership variables, control variables, and fixed effects. The dependent variable is Firm Reports CO2e, a dummy variable that equals one if a firm reports CO2-equivalent emissions, and zero otherwise. Industry fixed effects are based on SASB Industry Classifications. The sample covers the years 2023 and 2024. All variables are described in Appendix Tables A1. Standard errors are clustered at the country level and *t*-statistics are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

	Firm Reports CO2e
Family-controlled	-0.035
	(-1.56)
Government-controlled	-0.049
	(-1.65)
Other-controlled	0.014
	(0.66)
Log (Total Assets)	0.004
	(0.42)
Log (Revenue)	0.049^{***}
	(6.15)
Cash	-0.114^{**}
	(-2.23)
Tangibility	0.100^{**}
	(2.10)
Leverage	0.005
	(0.53)
Profitability	0.114^{***}
	(3.16)
Dual Class Shares	0.026
	(1.24)
Country FE	Yes
SASB Industry FE	Yes
Year FE	Yes
N	7,232
Adjusted R^2	0.188

Table 3
Controlled Firms and GHG Emissions

This table shows regression estimates of measures of firms' GHG emissions on ownership variables, control variables, and fixed effects. The dependent variables are the log of CO2-equivalent emissions (scaled by revenue and raw emissions). Industry fixed effects are based on SASB Industry Classifications. The sample covers the years 2023 and 2024. All variables are described in Appendix Tables A1. Standard errors are clustered at the country level and *t*-statistics are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

	Scope 1 and 2 Emissions		Scope 1 Emissions	
	Log (CO2e / Revenue)	Log (CO2e)	Log (CO2e / Revenue)	Log (CO2e)
	(1)	(2)	(3)	(4)
Family-controlled	0.158**	0.148**	0.180*	0.170*
	(2.20)	(2.18)	(1.82)	(1.77)
Government-controlled	-0.041	-0.050	-0.082	-0.091
	(-0.31)	(-0.38)	(-0.54)	(-0.59)
Other-controlled	0.309**	0.287**	0.130	0.111
	(2.31)	(2.21)	(0.69)	(0.59)
Log (Total Assets)	0.145***	0.393***	0.207***	0.378***
	(7.60)	(5.74)	(7.35)	(4.21)
Log (Revenue)		0.728***		0.813***
		(9.74)		(9.49)
Cash	-0.212	-0.385	-0.694**	-0.806***
	(-0.77)	(-1.59)	(-2.43)	(-2.90)
Tangibility	2.912***	2.819***	3.113***	3.058***
	(7.86)	(7.76)	(7.95)	(7.97)
Leverage	-0.062	-0.047	-0.184***	-0.173***
	(-1.34)	(-1.04)	(-3.06)	(-2.92)
Profitability	-1.530***	-1.115***	-1.853***	-1.592***
	(-4.17)	(-3.38)	(-3.18)	(-3.07)
Dual Class Shares	0.060	0.070	0.101	0.108
	(0.65)	(0.75)	(0.74)	(0.82)
Country FE	Yes	Yes	Yes	Yes
SASB Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	6,434	6,434	6,236	6,236
Adjusted R^2	0.608	0.773	0.612	0.745

Table 4
Controlled Firms, GHG Emissions, and Environmental Preferences

This table shows regression estimates of measures of firms' GHG emissions on family owners' environmental preferences, ownership variables, control variables, and fixed effects. The dependent variables are the log of CO2-equivalent emissions (scaled by revenue and raw emissions). Industry fixed effects are based on SASB Industry Classifications. The sample covers the years 2023 and 2024. All variables are described in Appendix Tables A1. Standard errors are clustered at the country level and *t*-statistics are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

	Scope 1 and 2 Emissions		Scope 1 Emissions	
	Log (CO2e / Revenue)	Log (CO2e)	Log (CO2e / Revenue)	Log (CO2e)
	(1)	(2)	(3)	(4)
Low E Family	0.224**	0.211**	0.216*	0.203*
	(2.56)	(2.49)	(1.91)	(1.84)
High E Family	0.022	0.018	0.001	-0.005
	(0.25)	(0.22)	(0.01)	(-0.04)
Inconclusive E Family	0.182**	0.173**	0.254**	0.245**
	(2.14)	(2.14)	(2.09)	(2.07)
Government-controlled	-0.052	-0.061	-0.099	-0.107
	(-0.40)	(-0.46)	(-0.65)	(-0.70)
Other-controlled	0.310**	0.288**	0.131	0.112
	(2.31)	(2.21)	(0.70)	(0.59)
Log (Total Assets)	0.150***	0.397***	0.215***	0.384***
	(7.31)	(5.73)	(7.00)	(4.23)
Log (Revenue)		0.729***		0.814***
		(9.75)		(9.55)
Cash	-0.208	-0.381	-0.690**	-0.800***
	(-0.76)	(-1.58)	(-2.41)	(-2.88)
Tangibility	2.900***	2.808***	3.098***	3.044***
	(7.91)	(7.81)	(8.00)	(8.02)
Leverage	-0.054	-0.040	-0.174***	-0.164***
	(-1.20)	(-0.89)	(-2.94)	(-2.79)
Profitability	-1.533***	-1.119***	-1.852***	-1.593***
	(-4.16)	(-3.39)	(-3.15)	(-3.05)
Dual Class Shares	0.057	0.068	0.100	0.107
	(0.62)	(0.72)	(0.73)	(0.82)
Country FE	Yes	Yes	Yes	Yes
SASB Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	6,434	6,434	6,236	6,236
Adjusted R^2	0.608	0.773	0.612	0.745
Low vs. High E Family (*p*-value)	(0.023)	(0.027)	(0.041)	(0.046)

Table 5
Controlled Firms, GHG Emissions, and Expected Marginal Costs

This table shows regression estimates of measures of firms' GHG emissions on ownership variables, family owners' environmental preferences, control variables, and fixed effects. We form subsamples based on i) the median Climate Change Performance Index (CCPI) score, with countries below the median classified as low CCPI (Panel A), and ii) a 50% cutoff for the share of Scope 1 emissions in an industry's total Scope 1 and 2 emissions (based on SASB Industry Classifications), with industries below 50% classified as low Scope 1 industries (Panel B). The dependent variables are the log of total CO2-equivalent emissions (scaled by revenue and raw emissions). The sample covers the years 2023 and 2024. All variables are described in Appendix Tables A1. Standard errors are clustered at the country level and *t*-statistics are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Sample Splits by the Climate Change Performance Index (CCPI)

	Scope 1 and 2 Emissions			
	High CCPI		Low CCPI	
	Log (CO2e / Revenue)	Log (CO2e)	Log (CO2e / Revenue)	Log (CO2e)
	(1)	(2)	(3)	(4)
Low E Family	0.213*	0.184	0.046	0.048
	(1.88)	(1.73)	(0.39)	(0.42)
High E Family	0.129	0.115	-0.176	-0.177*
	(1.37)	(1.32)	(-1.68)	(-1.78)
Inconclusive E Family	0.218*	0.209*	0.016	-0.011
	(1.79)	(1.80)	(0.18)	(-0.16)
Government-controlled	0.128	0.117	-0.167	-0.189
	(0.85)	(0.74)	(-1.12)	(-1.33)
Other-controlled	0.494***	0.455***	0.031	0.031
	(3.52)	(3.41)	(0.23)	(0.24)
Log (Total Assets)	0.186***	0.440***	0.102***	0.398***
	(8.70)	(5.51)	(3.14)	(3.49)
Log (Revenue)		0.717***		0.681***
		(7.49)		(6.54)
Cash	-0.645*	-0.769**	0.286	-0.008
	(-1.91)	(-2.85)	(0.94)	(-0.03)
Tangibility	3.258***	3.140***	2.834***	2.788***
	(6.05)	(6.03)	(6.12)	(5.88)
Leverage	-0.004	0.011	-0.365	-0.417
	(-0.10)	(0.30)	(-1.20)	(-1.34)
Profitability	-1.013***	-0.689*	-2.355***	-1.712***
	(-3.06)	(-1.88)	(-5.08)	(-4.24)
Dual Class Shares	-0.022	0.017	0.135	0.137
	(-0.10)	(0.08)	(1.21)	(1.17)
Country FE	Yes	Yes	Yes	Yes
SASB Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	3,205	3,205	3,064	3,064
Adjusted R^2	0.642	0.795	0.597	0.753
Low vs. High E Family (*p*-value)	(0.309)	(0.393)	(0.197)	(0.181)

Panel B: Sample Splits by Scope 1 Industries

	Scope 1 and 2 Emissions			
	High Scope 1 Industries		Low Scope 1 Industries	
	Log (CO2e / Revenue)	Log (CO2e)	Log (CO2e / Revenue)	Log (CO2e)
	(1)	(2)	(5)	(6)
Low E Family	0.381***	0.377***	0.082	0.066
	(3.05)	(3.18)	(0.89)	(0.67)
High E Family	0.257*	0.243	-0.219*	-0.218*
	(1.72)	(1.66)	(-1.99)	(-1.99)
Inconclusive E Family	0.434***	0.392***	-0.040	-0.032
	(3.95)	(4.00)	(-0.46)	(-0.35)
Government-controlled	0.035	0.006	-0.214	-0.214
	(0.22)	(0.04)	(-0.88)	(-0.91)
Other-controlled	0.334	0.287	0.272*	0.269*
	(1.67)	(1.55)	(1.82)	(1.82)
Log (Total Assets)	0.196***	0.507***	0.101***	0.275***
	(6.91)	(4.77)	(3.44)	(2.98)
Log (Revenue)		0.665***		0.807***
		(6.77)		(7.62)
Cash	-0.281	-0.421	-0.195	-0.360
	(-0.72)	(-1.36)	(-0.66)	(-1.24)
Tangibility	2.873***	2.789***	2.839***	2.770***
	(6.42)	(6.40)	(5.34)	(5.30)
Leverage	0.051	-0.041	-0.068*	-0.055
	(0.16)	(-0.13)	(-1.81)	(-1.47)
Profitability	-1.909***	-1.423***	-1.324***	-1.038***
	(-3.41)	(-2.83)	(-3.83)	(-2.93)
Dual Class Shares	-0.006	0.009	0.139	0.144
	(-0.07)	(0.10)	(1.23)	(1.24)
Country FE	Yes	Yes	Yes	Yes
SASB Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	3,279	3,279	3,155	3,155
Adjusted R^2	0.623	0.779	0.454	0.712
Low vs. High E Family (*p*-value)	(0.265)	(0.242)	(0.007)	(0.012)

Table 6
Orthogonalized Family Owners' Environmental Preferences and GHG Emissions

This table shows regression estimates of family owner's environmental preferences on country, family and firm characteristics (Panel A), and regression estimates of the residuals from Panel A on measures of GHG emissions (Panel B). In Panel A, the dependent variable is the family owners' environmental preferences. In Panel B, the dependent variables are the log of CO2-equivalent emissions (scaled by revenue and raw emissions), the sample covers the years 2023 and 2024. Low (High) E Family, $\varepsilon_{Family\ (2)}$ indicate families with low (high) environmental preferences based on the residuals of column 2 of Panel A. Industry fixed effects are based on SASB Industry Classifications. All variables are described in Appendix Tables A1. Standard errors are clustered at the country level and *t*-statistics are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Explaining Family Owners' Environmental Preferences

	Family Owner's Environmental Preferences (y_{Family})	
	(1)	(2)
English Narrow	0.487	0.496
	(0.89)	(0.90)
English Extensive		0.810
		(1.38)
Log (GDP/Capita)	-0.583***	-0.374*
	(-2.99)	(-1.92)
Forbes List	2.215***	2.198***
	(5.17)	(5.04)
Log (Total Assets)	0.777***	0.802***
	(5.04)	(5.15)
SASB Industry FE	Yes	Yes
N	887	887
Adjusted R^2	0.115	0.117

Panel B: Family Owners' Environmental Preferences and GHG Emissions

	Scope 1 and 2 Emissions	
	Log (CO2e / Revenue)	Log (CO2e)
	(1)	(2)
Low E Family, $\varepsilon_{\text{Family (2)}}$	0.211**	0.196**
	(2.43)	(2.29)
High E Family, $\varepsilon_{\text{Family (2)}}$	0.008	0.014
	(0.08)	(0.16)
Inconclusive E Family	0.181**	0.173**
	(2.13)	(2.14)
Government-controlled	-0.046	-0.055
	(-0.35)	(-0.42)
Other-controlled	0.309**	0.288**
	(2.31)	(2.22)
Log (Total Assets)	0.145***	0.391***
	(7.47)	(5.62)
Log (Revenue)		0.730***
		(9.69)
Cash	-0.203	-0.376
	(-0.74)	(-1.55)
Tangibility	2.903***	2.811***
	(7.94)	(7.85)
Leverage	-0.054	-0.040
	(-1.17)	(-0.88)
Profitability	-1.528***	-1.116***
	(-4.16)	(-3.38)
Dual Class Shares	0.059	0.069
	(0.65)	(0.75)
Country FE	Yes	Yes
SASB Industry FE	Yes	Yes
Year FE	Yes	Yes
N	6,434	6,434
Adjusted R^2	0.608	0.773
Low vs. High E Family (*p*-value)	(0.031)	(0.054)

Appendix A

Table A1
Variable Descriptions and Data Sources

This table reports variable definitions and data sources. Unless otherwise stated, all data are as of fiscal year-end 2023 and 2024.

Variable	Description	Source
A. Environmental Performance		
Firm Reports CO2e	A dummy variable equal to one if a firm reports CO2-equivalent emissions.	LSEG[25]
CO2e	Total Scope 1 and Scope 2 CO2-equivalent emissions (in tonnes). CO2-equivalent emissions account for CO2 as well as other greenhouse gases (CH4, N2O, HFCs, PFCs, SF6, NF3) and are calculated in accordance with the Greenhouse Gas Protocol.	LSEG
Log (CO2e)	Log of total CO2e.	LSEG
Log (CO2e / Revenue)	Log of total CO2e scaled by revenue in millions of $.	LSEG, Worldscope
B. Ownership		
Family-controlled	A dummy variable equal to one if a firm is classified as controlled by a family (as of December 2022). Control requires that the sum of the shares owned by family members is greater than 20% or that family members own at least 10% of the shares and the company has multiple voting class shares, and the sum is greater than any other shareholder.	Manual classification
Government-controlled	A dummy variable equal to one if the largest shareholder of the firm owns at least 20% of the firm and is the government or a sovereign wealth fund (as of December 2022).	Manual classification
Other-controlled	A dummy variable equal to one if the largest shareholder owns at least 20% of the firm and is a private equity fund, hedge fund, venture capital fund, other type of blockholder, or if ownership cannot be established (as of December 2022).	Manual classification
Widely Held	A dummy variable equal to one if a firm is not classified as Family, Government, or Other Opaque (as of December 2022).	Manual classification
C. Family Owners' Characteristics		
Family E Preferences	Family owner's environmental preferences are measured with the sum of five LLM generated measures ranking from zero to 50 (each component has a range of 0 to 10). Larger numbers are associated with greater environmental preferences. The five measures are: 1) personal philanthropy and charitable giving towards environmental causes; 2) public statements and advocacy for environmental issues; 3) participation in environmental NGOs; 4) green investments in the family's personal portfolio outside the firm; 5) policy support and political contributions for environmental policies.	GPT, Perplexity
Low E Family	A dummy variable equal to one if the sum of the LLM generated family owner's ownership preferences is less than or equal to the median, and zero otherwise.	GPT, Perplexity
High E Family	A dummy variable equal to one if the sum of the LLM generated family owner's ownership preferences is greater than the median, and zero otherwise.	GPT, Perplexity

[25] At the time of data extraction, the database was called Refinitiv. Refinitiv was acquired by LSEG in 2021, and rebranding Refinitiv products began in late 2023.

Inconclusive E Family	A dummy variable equal to one if the sum of the LLM generated family owner's ownership preferences is unknown, and zero otherwise.	GPT, Perplexity
Forbes List	Indicates whether the controlling family appears in Forbes' the World's Billionaires 2022.	Forbes
D. Other Firm Characteristics		
Total Assets	Total assets in $ million.	Worldscope
Revenue	Revenue in $ million.	Worldscope
Cash	Cash over total assets.	Worldscope
Tangibility	PP&E over total assets.	Worldscope
Leverage	Long-term debt over total assets.	Worldscope
Profitability	Net income over total assets.	Worldscope
Dual Class Shares	A dummy variable that is equal to one if a firm has a dual class share ownership structure, and zero otherwise.	LSEG, Worldscope
E. Country Characteristics		
CCPI	The Climate Change Performance Index (CCPI) is a standardized framework used to compare the climate performance of 63 countries and the EU. The overall CCPI score is based on four categories: GHG Emissions, Renewable Energy, Energy Use and Climate Policy.	Germanwatch
English Narrow	English Narrow indicates native English-speaking countries (Australia, Canada, Ireland, UK).	Wikipedia
English Extensive	English Extensive indicates countries with widespread English use (India, Malaysia, the Philippines, Singapore).	Wikipedia
Log (GDP/Capita)	Log of GDP per capita in $.	World Bank

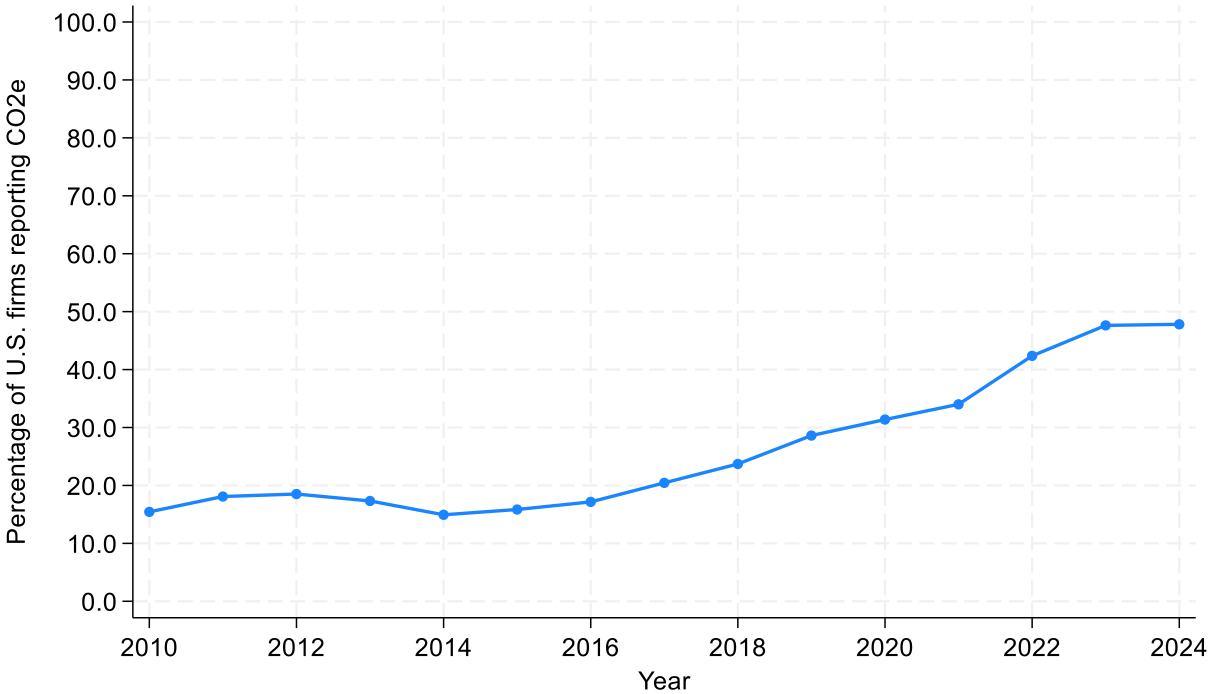
Table A2
Controlled Firms and GHG Emissions: Splits by the GHG Emissions Category of the Climate Change Performance Index

This table shows subsample regression estimates of measures of firms' GHG emissions on ownership variables, family owners' environmental preferences, control variables, and fixed effects. The subsamples are based on a country's score on the GHG Emissions Category of the Climate Change Performance Index (CCPI). The dependent variables are the log of CO2-equivalent emissions (scaled by revenue and raw emissions). Industry fixed effects are based on SASB Industry Classifications. The sample covers the years 2023 and 2024. All variables are described in Appendix Tables A1. Standard errors are clustered at the country level and *t*-statistics are reported in parentheses. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

	Scope 1 and 2 Emissions			
	High CCPI (GHG Emissions Category)		Low CCPI (GHG Emissions Category)	
	Log (CO2e / Revenue)	Log (CO2e)	Log (CO2e / Revenue)	Log (CO2e)
	(1)	(2)	(5)	(6)
Low E Family	0.202*	0.179*	0.081	0.084
	(2.05)	(2.07)	(0.54)	(0.56)
High E Family	0.115	0.106	-0.235*	-0.237*
	(1.30)	(1.32)	(-1.85)	(-1.87)
Inconclusive E Family	0.204	0.206*	0.098	0.054
	(1.73)	(1.82)	(0.70)	(0.41)
Government-controlled	0.061	0.047	-0.093	-0.089
	(0.36)	(0.27)	(-0.60)	(-0.56)
Other-controlled	0.393**	0.369**	0.167	0.151
	(2.35)	(2.30)	(0.74)	(0.72)
Log (Total Assets)	0.186***	0.370***	0.114***	0.427***
	(9.34)	(3.75)	(3.21)	(4.14)
Log (Revenue)		0.795***		0.660***
		(6.89)		(7.82)
Cash	-0.819**	-0.908***	0.203	-0.148
	(-2.39)	(-3.05)	(0.71)	(-0.51)
Tangibility	3.314***	3.225***	2.662***	2.584***
	(5.96)	(6.10)	(6.67)	(6.43)
Leverage	-0.006	0.006	-0.334	-0.390
	(-0.14)	(0.17)	(-1.27)	(-1.52)
Profitability	-1.067**	-0.820*	-2.184***	-1.597***
	(-2.90)	(-2.02)	(-5.15)	(-4.62)
Dual Class Shares	0.083	0.107	0.052	0.041
	(0.53)	(0.68)	(0.38)	(0.31)
Country FE	Yes	Yes	Yes	Yes
SASB Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	3,060	3,060	2,931	2,931
Adjusted R^2	0.656	0.809	0.581	0.748
Low vs. High E Family (*p*-value)	(0.149)	(0.195)	(0.085)	(0.072)

Figure A1
Percent of U.S. Firms Reporting Carbon Emissions

This figure reports the percentage of U.S. firms that report CO2-equivalent emissions from 2010 to 2024 for the population of U.S. LSEG firms, which includes all non-financial firms with a minimum market capitalization of $100 million, and incorporated in the U.S. This sample includes between 1,166 (in 2010) and 2,238 firms (in 2021). All data are from LSEG.



Appendix B
Ownership Classification

B.1. Establishing Controlling Owners

We identify firms' ownership structures as of the year-end 2022, starting with commercially available data sources such as Bureau van Dijk's (BvD) Orbis, LSEG, Datastream, and Worldscope, and then manually researching, categorizing, and verifying each firm's ultimate ownership with data from a variety of additional sources, including annual reports, internet searches, and country guides. The process is described in detail in Section III.D.

B.2. Ownership Stability

To address the question of whether our ownership classification—coded once as of year-end 2022—remains stable over the full sample period, we re-examine each firm's ownership type as of year-end 2024. For each of the 3,769 firms we supply a large language model (LLM) with the firm's identity and its known 2022 ownership type and instruct it to search public sources for evidence that the firm's controlling-owner category changed between year-end 2022 and year-end 2024. A first pass uses GPT-4.1 with live web search over all firms; firms flagged as potential switchers are escalated to a second, higher-cost pass using GPT-5. Because the model returns the specific primary sources—company ownership disclosures, regulatory filings, and news reports—underlying each flagged change, every reclassification is independently checkable by hand. The classification is highly stable: 3,709 of the 3,769 firms (98.4%) retain their 2022 ownership type through year-end 2024, and reclassifications are concentrated in just 60 firms (1.6%). Stability is high within every category—98.9% of family firms, 97.8% of widely held firms, 98.9% of government-controlled firms, and 99.6% of financial/other firms remain in their original type.

Appendix C
LLM Environmental Preference Measure

C.1. Identifying Controlling Families (OpenAI GPT-4o)

We first identify firms classified as family-controlled and use the LLM to determine the controlling family's name, board members, their gender, and the founder. This is achieved through the prompt below.

> PROMPT C.1 Family assessment. You are a financial data expert. You are assessing families that control publicly traded firms. Control is defined as follows. i) the sum of the shares owned by the family members exceeds those of any other shareholder and is greater than 20% OR ii) the sum of family stakes exceeds those of any other shareholder, is greater than 10%, and family members hold the CEO or chair position OR iii) the sum of family stakes exceeds those of any other shareholder, is greater than 10%, and (the firm has multiple voting class shares) OR (the CEO or other board member is a family member).
>
> You are using publicly available data as of December 2022 to make your assessment. For each observation you have several variables, as follows:
> companyname - name of the firm
> familyname - the name of the family that I have assessed as the one controlling the firm (if available)
> country- The country of incorporation of the firm specified in companyname
> dscode - The Datastream identifier of the firm specified in companyname
>
> q1: Assume that each firm is indeed controlled by a family, according to the above given definition of control. Assume also that my assessment of the family name--if available--is correct. Identify the family that in your opinion is most likely the one controlling the firm. Note that the family might represent one or multiple family members. Provide one family name, choosing the most visible one if there are multiple, of the controlling family.
>
> q2: Assess how certain you are about your ability to identify the family in q1 with a posterior confidence score (PCS 0–100).
>
> q3: List up to five members of the controlling family, stating the first name, last name, gender, that meet the condition of having a position on the board of directors in 2022.
>
> q4: Give a summary of the controlling family's involvement in the firm, with historical context, in 80 words. In case you cannot identify the family from q1, give a summary of why you cannot.
>
> In all answers, you are using publicly available data as of December 2022 to make your assessment. Avoid hypothetical or example-based explanations; directly provide insights, conclusions, or statistical trends if applicable. If there is insufficient data and you are not at least 80% sure that you know the answer, reply "NA".
>
> Structure your answer as follows.
> variable q1 as a string (family name(s))
> variable q2 as a number (0 to 100)
> variable q3 as a string (first last (gender), first last (gender), ...)
> variable q4 as a string (80 words)

To validate the family-identification step, we run a controlled recovery exercise on Canadian firms. We inform the model that each firm is family-controlled but withhold the controlling family's name, and ask it to recover the name; we then compare its output to our hand-collected family names. Across 67 Canadian family firms, the model achieved 100% precision in family-name recovery. We note that Canada is an English-language, high-disclosure setting, so this figure is best read as an upper bound on identification accuracy across our full cross-country sample.

C.2. Assessing Family's Environmental Preferences (Perplexity Sonar-pro)

The family name identified through GPT is input into Perplexity Sonar-pro, a search-augmented LLM, to assess the family's personal environmental preferences. We leverage the model's retrieval capability to provide online references, facilitating our (human) review of the assessments. The environmental preferences are evaluated across five dimensions on a scale from 0 to 10: a) Personal philanthropy towards environmental causes; b) Public statements and advocacy for environmental issues; c) Participation in environmental NGOs; d) Green investments in personal portfolios outside the firm; e) Policy support and political contributions for environmental policies. The LLM is explicitly instructed to exclude any consideration of the firm's corporate environmental performance.

The assessments of environmental preferences were generated in July 2026 through the vendor's API, with one query per controlling family (899 families; scores are inherited by all sample firms controlled by the same family). Three features of the model are important: (i) the December 31, 2023 information cutoff is imposed through the vendor's search-timestamp filter, which restricts retrieved web content to material dated on or before the cutoff; (ii) the output is bound to a structured response format, so all 899 assessments parse without manual intervention; and (iii) the model reports a posterior confidence score (PCS, 0–100) for each assessment. Families with confidence below 40 are classified as inconclusive; the remaining families are split at the sample median of the total score into the low- and high-preference groups used in our tests.

A fixed system prompt states the assessment rules; a per-family user prompt supplies the identification context and the scoring rubric. Both are reproduced below, with bracketed placeholders customized for each observation.

PROMPT C2. Preference assessment.
SYSTEM PROMPT (fixed). You are a research assistant specializing in corporate governance and ESG analysis. Your task is to assess the personal environmental preferences of controlling families of publicly listed firms, using web search to find evidence. You must be careful to:
1. Identify the correct family — use the reference individuals and firm context to disambiguate. In case the family name is a frequent name, be especially careful to not confuse the family with another family of the same name.
2. Assess only the family members' PERSONAL environmental preferences — you must ignore information and activities of the firm itself, which the family controls.
3. If the provided family name is a business-group or holding-company label rather than a surname (e.g. an "... group" or an investment vehicle), assess the family behind that group.
4. If you can find information about the family but no evidence of environmental preferences, this is consistent with a low or very low score — not a reason to refuse to assign a score.
5. Ignore the preferences of deceased family members.
6. Evidence may be in the firm's home-country language — search in English and in the local language where useful.
7. Cite only sources that are directly related to your assessment of the family's environmental preferences.

You MUST return your answer as a valid JSON object and nothing else.

USER PROMPT (per family). Your task is to assess the **[family name] family's** environmental preferences as of December 2023. Search the web for evidence.

Context to identify the correct family:
- To avoid confusion, the family is linked to [country] and to [company name] in that country.
- Firm identifier (Datastream code): [Datastream code]
- We assess that the family controls this firm — treat family control as given.
[reference individuals: family board members, if available]

Scoring dimensions (each scored 0 to 10, where 10 indicates strong environmental preferences):
- **a) philanthropy**: Personal philanthropy and charitable giving towards environmental causes
- **b) public_statements**: Public statements and advocacy for environmental issues
- **c) ngo_participation**: Participation in environmental NGOs
- **d) green_investments**: 'Green' investments in the family's personal portfolio outside the firm
- **e) policy_support**: Policy support and political contributions for environmental policies

```
- **f) total_score**: The sum of these five dimensions as a score between 0 and 50

**Additional assessments:**
- **g) confidence**: Assess how certain you are about your ability to identify the family preferences with a posterior confidence score (PCS 0-100)
- **h) summary**: A summary narrative of your assessment, not exceeding 100 words

Remember: ignore the preferences of deceased family members; assess only the family members' personal environmental preferences and ignore information and activities of the firm ([company name]), which the family controls; finding information about the family but no evidence of environmental preferences is consistent with a low or very low score and not a reason to refuse to assign a score.

**Output valid JSON only:**
{
  "a_philanthropy": <int 0-10>,
  "b_public_statements": <int 0-10>,
  "c_ngo_participation": <int 0-10>,
  "d_green_investments": <int 0-10>,
  "e_policy_support": <int 0-10>,
  "f_total_score": <int 0-50>,
  "g_confidence": <int 0-100>,
  "h_summary": "<string, max 100 words>"
}
```

C.3. What the Measure Captures

The preference assessment procedure is best understood by analogy to a large-scale research assistant exercise. The LLM functions as a highly trained assistant that searches the open web on behalf of the research team and scores what it finds according to a structured rubric — the five environmental preference dimensions described in Section II. The key advantage relative to a conventional RA-based approach is scale. Obtaining independent assessments from multiple human coders for 899 controlling families (1,588 family firms) across 35 countries and multiple languages would have been prohibitively costly, if not impossible. The key concerns are about reliability and replicability.

The model we use, Perplexity Sonar-pro, is a retrieval-augmented system that draws on two distinct inputs whose relative contributions to any individual score are not separately identifiable. The first is retrieved web content: at query time the model fetches pages from the open web, and the URLs of retrieved documents are logged as citations. The second is parametric knowledge embedded in the model's weights from pre-training on a large corpus of text. For well-documented families — those with substantial media coverage, foundation websites, or philanthropic records — the model may draw substantially on pre-training knowledge. For less prominent families, retrieved content likely dominates. The citations logged for each firm therefore represent documents the model had access to during generation, not verified claim-level attributions: we cannot establish which retrieved document drove which component of the score, nor can we fully separate the retrieval component from the parametric component.

Our assessments use information available as of December 2023. For retrieved content this cutoff is enforced mechanically: the vendor’s search-timestamp filter restricts retrieval to material dated on or before December 31, 2023. For parametric knowledge no hard cutoff exists, so for families where pre-training knowledge is a significant input the December 2023 anchor is an instruction rather than a guarantee. We note that this is a property of all current retrieval-augmented language models rather than a feature specific to our implementation, and that it implies our scores may in some cases reflect information that postdates our stated baseline. Given that environmental preferences are slow-moving, we do not believe this introduces systematic bias into the cross-sectional comparisons that underpin our analysis. Section C.6 quantifies this directly: re-estimating the measure with information through December 2022 leaves the classification largely unchanged but reduces the number of families classified.

C.4. Source Types

Alongside each assessment, the model logs the URLs of the web documents it draws on. Every one of the 899 family assessments returns at least one source; in total the model cites 8,994 documents, a mean of ten per family, drawn from roughly 3,600 distinct internet domains across 129 top-level domains. This breadth is one of the central advantages of the approach: rather than relying on a single commercial database or disclosure standard, the model assembles evidence from the full range of publicly available sources relevant to each family.

The sources are international in scope. Table C.1 organizes the cited domains by top-level domain and gives examples of the sources retrieved within each. The evidence base spans national and international institutions, family and corporate foundations, company investor-relations and sustainability pages, and business and financial media across the sample's many jurisdictions. The country coverage of the cited domains tracks the geography of our sample — Indian, Korean, Thai, Malaysian, Turkish, Nordic, Italian, French, Japanese, and other national sources all appear — indicating that the model retrieves locally relevant material rather than relying solely on English-language or Western sources.

The composition of sources varies across the preference distribution in an interpretable way: citations to family- and corporate-foundation and philanthropy domains are roughly three times more prevalent in the assessments of high-preference families than of low-preference families (2.5% versus 0.9% of citations). This is the pattern one would expect if the model locates such sources for families that are in fact more engaged, and falls back on general sources when little family-specific environmental material exists.

The relative weight of source types also varies systematically across countries, reflecting cross-country differences in the availability and indexing of public, and especially English-language, content. This variation is one motivation for the orthogonalized preference score we report in Table 7, which removes variation in the measure correlated with English-language exposure, GDP per capita, family wealth, and firm size, and which leaves our conclusions unchanged. We note, as discussed in Section C.3, that citation logs record the documents the model retrieved and not, in any directly measurable way, the weight each document received in the final assessment.

C.5. Illustrative Cases

We provide detailed case studies of six families in Table C2: two classified as high environmental preference, two as low, and two as inconclusive, chosen to span income levels and regions. The high and low cases include the two families discussed in Section III.E of the paper. The two inconclusive cases illustrate the confidence screen at work: in each, the model locates only general or firm-level material rather than family-specific evidence, reports low confidence, and the family is accordingly left unclassified rather than assigned a noisy label.

C.6. Reliability, Stability and Robustness

Retrieval-augmented assessments are time-indexed: model outputs reflect both retrieved content available at query time and knowledge accumulated through pre-training, and both components evolve as models are updated and web indices grow. A useful analogy is re-administering a survey to a population using different and likely better-trained interviewers who can consult additional material: individual answers shift, while the cross-sectional ranking the survey is designed to capture should persist. Because our tests rely on the cross-sectional ranking of families—not on the cardinal value of any individual score—classical measurement error in the scores attenuates regression coefficients toward zero, making the associations we document conservative.

To assess the reliability, stability and robustness of our results we conducted a number of tests. First, we assess the test-retest reliability. Running the identical prompt through the model three independent times on a stratified 150-family subsample yields pairwise score correlations of 0.79-0.84, an intraclass correlation of 0.79 for a single run (0.92 for a three-run average), and 80-84% agreement on the low/high classification among families scored in both runs. All runs use the same stratified subsample of 150 families (50 each drawn from the low, high, and inconclusive categories). Second, we assess the correlations based on two semantically equivalent rewordings of the prompt—one of which also lists the five scoring dimensions in reverse order. Again, these runs use the same stratified subsample of 150 families (50 each drawn from the low, high, and inconclusive categories). These runs produce raw scores that correlate 0.57-0.74 with the original wording, below the 0.79-0.84 test–retest range: because the retrieval queries derive from the prompt itself, rewording changes not only the scoring but part of the evidence retrieved. The classification that enters our regressions is considerably more robust: among families confidently scored under both wordings, 71-76% receive the same low/high label, close to the 80-84% agreement across identical-prompt reruns. We treat wording sensitivity of the raw scores as an additional source of classical measurement error, which attenuates coefficients toward zero and thus makes our tests conservative.

The remaining tests, reported in Table C3, assess the stability and robustness of four alternative preference pipelines where there is both variation across timestamps and across a different vendor (OpenAI GPT-5). Panel B reports the correlations of preference scores across vintages and vendors; we first discuss stability over the information window. Re-estimating all 899 assessments using information through December 31, 2022 yields family scores that correlate 0.72 with the baseline vintage (75% of families scored in both retain their low/high label, n = 334). Further analysis shows that transitions occur disproportionately through the inconclusive category. Notably, coverage grows with the information set: 672 of 899 families clear the confidence screen under the 2023 vintage, versus roughly 380 under the 2022 vintage. The Table also reports agreement across independently generated environmental preference measures. Our scores correlate 0.51 with a re-estimation by a different vendor and architecture (OpenAI GPT-5 with live web search) under the identical prompt; among families scored by two measures, roughly two thirds (67%) receive the same low/high classification.

Panel C reports the coefficients for the three preference categories for family firms. These coefficients show a robust relationship between preferences and carbon emissions across the baseline preference pipeline and three other pipelines (an earlier cut-off vintage, a different vendor (OpenAI GPT-5) using as of Dec 2023 vintage, and a different vendor (OpenAI GPT-5) using as of Dec 2022 vintage). The low-preference coefficient is positive in every estimation, and the low-versus-high difference is statistically significant in three of the four specifications. The Table also reports summary statistics for the alternative measures.

Table C1
Provenance of Sources Cited in the Preference Assessments

This table summarizes the web sources cited by Perplexity sonar-pro across the 899 family assessments (all return at least one source). There are 8,994 citations in total, spanning roughly 3,600 domains, and 129 top-level domains (TLDs). Rows correspond to the most-cited TLDs. Times cited reports the total number of citations to domains under each TLD. Example sources are illustrative of the material retrieved under each TLD. Data are shown for the top 20 TLDs. As discussed in Section C.3, citation logs record documents retrieved at query time and do not quantify any document's contribution to a given score.

TLD	Times cited	Examples of sources
.com	4,285	youtube.com, linkedin.com, timesofindia.indiatimes.com
.org	1,482	en.wikipedia.org, oecd.org, ideas.repec.org
.gov	500	sec.gov, pmc.ncbi.nlm.nih.gov, pubmed.ncbi.nlm.nih.gov
.se	191	lusem.lu.se, di.se, lunduniversity.lu.se
.kr	181	kei.re.kr, koreatimes.co.kr, english.hani.co.kr
.in	171	pwc.in, pib.gov.in, businesstoday.in
.ca	136	newswire.ca, statcan.gc.ca, impact.canada.ca
.uk	129	ons.gov.uk, company-information.service.gov.uk, research.manchester.ac.uk
.jp	124	jstage.jst.go.jp, env.go.jp, crs.or.jp
.fr	112	lemonde.fr, notre-environnement.gouv.fr, lefigaro.fr
.ch	106	blick.ch, cph.ch, tagblatt.ch
.de	100	oekotest.de, welt.de, hermle.de
.edu	96	academia.edu, downloads.cs.stanford.edu, sites.amherst.edu
.my	93	thestar.com.my, mida.gov.my, insage.com.my
.au	85	morningstar.com.au, acnc.gov.au, reneweconomy.com.au
.net	79	business.inquirer.net, slideshare.net
.fi	78	finland.fi, yle.fi, maaseuduntulevaisuus.fi
.sg	60	businesstimes.com.sg, ntu.edu.sg, nlb.gov.sg
.th	57	ph.mahidol.ac.th, sgs.tu.ac.th, market.sec.or.th
.nl	50	epgroningen.nl, research.rug.nl, cbs.nl

Table C2
Assessment of Environmental Preferences of Families

This table shows case studies of six families, two from each preference classification (high, low, inconclusive), chosen to span income levels and regions; the high and low cases include the two families discussed in the text. Score is the model's 0-50 preference score; Conf. is its self-reported confidence (0-100; scores with confidence below 40 are not classified). LLM assessment reproduces the model's narrative (source citations removed for readability).

Classi-fication	Firm	Country	Family	Score	Confi-dence	LLM assessment
High	Kone Oyj	Finland	Herlin	36	88	The Herlin family, identified here through [FAMILY MEMBER], [FAMILY MEMBER], [FAMILY MEMBER], and [FAMILY MEMBER]in the KONE context, shows strong personal environmental preferences. [FAMILY MEMBER] has repeatedly supported Baltic Sea and climate-related work, including founding and serving in the Elävä Itämeri Foundation and working on carbon-sequestration agriculture at Qvidja; public coverage also portrays him as advocating environmentally responsible business and societal action. [FAMILY MEMBER] and [FAMILY MEMBER]'s foundation has an explicit emissions-cutting environmental program. [FAMILY MEMBER]'s public role also appears tied to sustainability work. Direct evidence on private green investments and policy contributions is limited.
High	Sunway Bhd	Malaysia	Cheah	19	78	[FAMILY MEMBER] appears to have strong personal interest in sustainability and some environmental philanthropy. Public interviews and Sunway-related profiles quote him on sustainability, litter reduction, and building a greener future, but the evidence is mostly about his public advocacy rather than private activism. I found little direct evidence of personal NGO roles, green investments outside Sunway, or political/policy giving for environmental causes.
Low	Grupo México	Mexico	Larrea	0	86	Publicly available evidence as of December 2023 shows little to no direct personal environmental engagement by the Larrea family's controlling individual, [FAMILY MEMBER]. The available material focuses on his secrecy and Grupo México's controversies, not on personal philanthropy, advocacy, NGO roles, green investments, or policy support for environmental causes. I found no credible evidence of the family members' personal environmental preferences, so the score is set very low.
Low	Guerbet SA	France	Guerbet	0	88	I found no reliable web evidence, as of December 2023, that the Guerbet family personally donated to environmental causes, made public environmental statements, joined environmental NGOs, held green personal investments, or supported environmental policy. The available results mainly describe Guerbet SA's own CSR/sustainability actions, which are excluded from this assessment. Based on the absence of family-level evidence, the family scores very low on all dimensions.

Table C2, continued

Classi-fication	Firm	Country	Family	Score	Confi-dence	LLM assessment
Incon-clusive	Korea Zinc	South Korea	Chang and Choi	0	18	I found no direct evidence that the Chang/Choi controlling family of Korea Zinc has made personal environmental donations, public pro-environment statements, NGO participation, green personal investments, or political support for environmental policy. The available results are general South Korea climate-attitude sources, not family-specific evidence. Given the lack of identifiable personal ESG behavior for the family as of December 2023, the score is set at the lowest level.
Incon-clusive	Lotte Chemical Titan	Malaysia	Shin	2	18	I found no direct evidence that the Shin family linked to Lotte Chemical Titan made personal environmental donations, public advocacy, NGO involvement, green investing, or policy support by December 2023. The available results mostly concern the firm or unrelated Shin entities, so the assessment is low-confidence and near-zero.

Table C3
Alternative Estimates of the Preference Measure

This table reports summary statistics and regression estimates replicating Table 4 for four alternative LLM-generated environmental preference measures. Panel A reports summary statistics for environmental preferences assessed by i) Perplexity Sonar-pro with a December 2023 cutoff, our baseline from Table 4, ii) Perplexity Sonar-pro with a December 2022 cutoff, iii) OpenAI GPT-5 with a December 2023 cutoff, and iv) OpenAI GPT-5 with a December 2022 cutoff. Panel B reports the pairwise correlations of the four measures below the diagonal and the label agreement (families with high, low, inconclusive E preferences) above the diagonal. Panel C reports regression estimates, where each column re-estimates the Table 4 specification with low/high/inconclusive classifications built from the respective measure, using a median split. The dependent variable is the log of total Scope 1 and Scope 2 CO2-equivalent emissions scaled by revenue column, and all specifications include the Table 4 controls and industry, country, and year fixed effects, with standard errors clustered by country. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Summary Statistics for Four Alternative Preference Measures

	Families assessed	% of 899	Mean	Median	Max	Share at 0
Perplexity, as of Dec 2023 (baseline)	664	74%	3.5	2	36	33%
Perplexity, as of Dec 2022	378	42%	5.3	3	40	16%
GPT-5, as of Dec 2023	502	56%	9.3	7	44	16%
GPT-5, as of Dec 2022	469	52%	8.9	7	41	17%

Panel B: Correlations (Below Diagonal) and Label Agreement (Above Diagonal)

	PPLX, 2023	PPLX, 2022	GPT-5, 2023	GPT-5, 2022
PPLX, 2023	1.00	75% (n=334)	67% (n=408)	68% (n=390)
PPLX, 2022	0.72	1.00	68% (n=281)	73% (n=264)
GPT-5, 2023	0.51	0.65	1.00	81% (n=368)
GPT-5, 2022	0.54	0.63	0.87	1.00

Panel C: Table 4 Specifications Using Each Measure

	Scope 1 and 2 Emissions Log (CO2e / Revenue)			
	PPLX, 2023 (baseline)	PPLX, 2022	GPT-5, 2023	GPT-5, 2022
	(1)	(2)	(3)	(4)
Low E family	0.224**	0.260**	0.232*	0.232**
	(2.56)	(2.53)	(1.83)	(2.23)
High E family	0.022	0.045	0.144	0.044
	(0.25)	(0.48)	(2.36)	(0.62)
Inconclusive E family	0.182**	0.151*	0.134*	0.171**
	(2.14)	(1.93)	(1.86)	(2.21)
Low vs. High E Family (*p*-value)	(0.023)	(0.030)	(0.449)	(0.029)

About ECGI

ECGI is an international non-profit membership organisation that provides a platform for debate and dialogue among academics, policymakers, and business leaders, with a focus on major corporate governance, ESG, and stewardship issues.

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

www.ecgi.global