Intermediaries and Emissions Disclosures*

Rongchen Li†

First draft: October 10, 2024

This draft: October 6, 2025

Abstract

Major intermediaries provide investors with estimates of firms' emissions, using emissions models calibrated with disclosing firms' data. If firms disclose selectively, however, intermediaries risk underestimating non-disclosing firms' emissions, thereby deterring firms from disclosing. Using novel data, I document evidence that higher-emitting firms indeed abstain from disclosure and intermediaries' estimates appear downward biased. Exploiting a plausibly exogenous shift in intermediary coverage, I then show that firms are less likely to disclose their emissions once the intermediary's estimates become available. Finally, I quantitatively explore the potential consequences of refining intermediaries' estimates and the roles played by disclosure mandates.

Keywords: Greenhouse Gas (GHG) Emissions, Information Intermediaries, Voluntary Disclosure, Regulation, Selection Bias

JEL Classification: D82, D83, D91, G24, M41, Q54, Q56, Q58

*This paper is based on my dissertation at Columbia Business School. I am deeply indebted to my dissertation committee—Thomas Bourveau (co-chair), Matthias Breuer, Jon Glover (co-chair), Wei Jiang, and Sehwa Kim—for their tremendous support and mentorship. I also thank Sonakshi Agrawal, Tim Baldenius, Khrystyna Bochkay, John Core, Kent Daniel, Paul Fischer, Xingchao Gao, Russell Han, Kalash Jain, Dian Jiao, Bin Ke, Anthony Le, Lisa Liu, Shunsuke Matsuno, Stephen Penman, Joaquin Peris, Shiva Rajgopal, Frank Zhou, workshop participants at Columbia, 2024 AAA/Deloitte Foundation Doctoral Consortium, MSCI, National University of Singapore, London Business School, University of Hong Kong, Hong Kong Polytechnic University, Arizona State University, Stanford University, Carnegie Mellon University, MIT Sloan, University of Pennsylvania, UT-Austin, Duke University, University of Minnesota, and LSE for helpful discussions. I thank Thomas Rauter and Tom Barry for help with the Trucost Environmental data. I thank Irene Tan for help with the Climate TRACE data. I acknowledge financial support provided by Columbia Business Schoool and the Wharton School. All errors are my own.

†The Wharton School, University of Pennsylvania; rl3@wharton.upenn.edu.

1 Introduction

Climate change poses substantial risks to society. Corporations worldwide both contribute to climate change and are exposed to its ensuing risks (e.g., Griffin, 2017; Greenstone et al., 2023; Sautner et al., 2023). Therefore, investors and other stakeholder groups are placing greater emphasis on firms' environmental performance, such as greenhouse gas (GHG) emissions, on top of their financial performance. However, such information is not readily available for most firms. For example, as of 2022, approximately 80% of public firms in the U.S. do not report their GHG emissions to the CDP or in their annual filings. To fill this information gap, intermediaries such as Bloomberg, Trucost, MSCI, and Sustainalytics produce statistical model-based estimates of *all* firms' emissions, where their models are calibrated using *disclosing* firms' data. Their "one-stop" platforms for corporate environmental data are widely used in both practice and academic research.

In this study, I examine the impact of information intermediaries' emissions estimates on firms' emissions disclosures. To this end, I develop an analytical framework to theorize how such estimates interact with firms' disclosures. The framework introduces intermediaries' estimates as an alternative source of information into a strategic disclosure game between firms and investors of varying sophistication. The intuition is that intermediaries' estimates, which are trained on data from voluntarily disclosing firms—often more carbon-efficient ones—tend to underestimate the non-disclosing firms' emissions. These estimates—if taken at face value—skews investors' beliefs in a direction favorable to firms, thereby reducing firms' incentives to disclose their emissions. To validate the framework empirically, I provide evidence on its core building blocks—that investors penalize firms' emissions, higher-emitting firms refrain from disclosure, and intermediaries produce biased estimates. I then test the predicted deterrent effect of intermediaries' emissions estimates on firms' emissions disclosures by exploiting the expansion of a major intermediary as shock to intermediary presence.

Building on seminal disclosure theories (e.g., Grossman, 1981; Milgrom, 1981), I explore how intermediaries' statistical model-based estimates affect firms' voluntary disclosure to investors. In my model, investors value firms' environmental performance (e.g., Hartzmark and Sussman, 2019; Krueger et al., 2020; Li et al., 2024), consistent with recent evidence that investors penalize high-emitting firms (e.g., Chava, 2014; Bolton and Kacperczyk, 2021a; Bolton et al., 2022; Pástor et al., 2021, 2022). To learn about firms' environmental performance, investors look to firms' disclosures. Absent disclosures, rational investors should be skeptical about firms' performance. Recent empirical and experimental evidence, however, suggests that some investors may miss the negative signal implied by firms' non-disclosure (e.g., Zhou and Zhou, 2020; Jin et al., 2021). Accordingly, I allow a fraction of investors to be naive; that is, they do not view non-disclosure as a negative signal. In the presence of naive investors, who fail to punish non-disclosure, the seminal unraveling result—all firms disclose—breaks down. Some firms, especially those with poor environmental performance, will refrain from disclosure. This selective and incomplete disclosure creates scope for information intermediaries to matter. In my model, just as in practice, the intermediary uses statistical modeling to extrapolate disclosing firms' environmental performance to non-disclosing firms using the observed relation between disclosed emissions and firm-level characteristics (e.g., size, location, industry membership).

My model suggests that intermediaries' estimates underestimate non-disclosing firms' emissions and further deter higher-emitting firms from disclosing their emissions. The bias arises because the intermediaries' statistical models are trained on a self-selected sample of voluntarily disclosing firms. In equilibrium, only lower-emitting firms choose to disclose. This selection biases intermediaries' model-based estimates as in Heckman (1976, 1979). Importantly, the bias affects firms' disclosure incentives: naive investors simply rely on the intermediaries' estimates, allowing marginal firms that would otherwise disclose to exploit the favorable bias by switching to non-disclosure. As fewer firms disclose, the estimation bias is magnified, which in turn further reduces firms' disclosure incentives. In sum, in

the presence of naive investors, intermediaries' estimates exacerbate firms' non-disclosure incentives, thereby deterring disclosure.

Guided by my analytical framework, I empirically explore how intermediaries' estimates affect firms' disclosures. I gather firm-level emissions data from two prominent information intermediaries in the emissions-measurement space: Trucost and Bloomberg. Both intermediaries provide a variety of firm-level environmental impact metrics, including scope 1, 2, and 3 GHG emissions.[1] They collect self-reported emissions from sources such as firms' annual filings, CDP responses, and corporate websites. When firms' disclosures are unavailable, they fill the data gap by providing estimated emissions derived from state-of-the-art statistical models. Both intermediaries have unique features that make them particularly relevant for my investigation. Trucost, an early leader in the business, expanded its coverage significantly after being acquired by S&P Dow Jones Indices in 2016. This expansion constitutes a shock to the coverage of intermediaries, allowing me to test the deterrent effect of intermediary estimates on firms' disclosure. Bloomberg, another major intermediary that launched its GHG emissions estimates in 2019, also has a unique advantage: it allows users to retrieve both its emissions estimates and firms' disclosed emissions, if they are available.[2] This feature allows me to compare the properties of disclosed and estimated emissions, which is key for validating the main building blocks of my analytical framework. The Trucost sample consists of 14,144 firms across 42 markets from 2014 to 2017, and the Bloomberg sample consists of 3,210 U.S. firms from 2020 to 2022. I supplement these emissions data with firm-level fundamental data and price data from Compustat and CRSP.

I first provide evidence that investors penalize firms with high emissions. Consistent with prior studies (e.g., Matsumura et al., 2014; Griffin et al., 2017; Homburg et al., 2024), I

[1] Throughout the paper, I focus on scope 1 emissions because they constitute the most direct measures of firms' GHG footprints according to the GHG Protocol (*link*). Scope 1 emissions are also relatively easy to verify, and collectively form an accurate picture of aggregate emissions across the economy (Leuz, 2022).

[2] Trucost, by contrast, provides emissions in a "waterfall" format; that is, they report their estimates *only* if disclosed emissions are not available.

find that firms' emissions are strongly negatively associated with their market values.[3] This negative association is observed both in specifications using estimated emissions and in those using disclosed emissions. Additional pricing tests, however, reveal that investors on average do not appear to see much incremental value in disclosed emissions, which are typically made public after Bloomberg's estimates. These findings suggest that investors frequently rely on intermediaries' estimates, consistent with my modeling assumption.

I next provide evidence that firms engage in strategic emissions disclosure. To test for strategic disclosure, I examine whether firms' decisions to disclose are negatively associated with their emissions performance. To identify a comparable and credible measure of firms' actual emissions for both disclosing and non-disclosing firms, I use remote-sensed emissions data recently released by Climate TRACE. Launched in 2021, Climate TRACE monitors the facilities of firms operating in heavy-emitting industries (e.g., steel production) using satellite data and machine learning techniques.[4] I aggregate Climate TRACE's facility-level emissions up to the firm level to construct noisy but selection-free signals of firms' emissions. Using a sample of 57 unique firms, I find that, after controlling for expected levels of emissions, firms with lower emissions are more likely to disclose. This finding aligns with the notion that firms strategically disclose their emissions, consistent with my model.

Finally, I document that intermediaries provide biased estimates for non-disclosing firms. My model assumes that intermediaries extrapolate model-based predictions to non-disclosing firms *without* adjusting for self-selection bias. This approach should lead to an unbiased in-sample fit (for disclosing firms) and understated out-of-sample estimates (for non-disclosing firms). With respect to the in-sample fit, Bloomberg details the accuracy of its emissions estimates in sample in a series of white papers (e.g., Bloomberg, 2021, 2024). Bloomberg highlights a "recalibration" procedure that maximizes in-sample fit, demonstrating its commitment to rigor in modeling firms' emissions. This focus on in-sample fit, however, can

[3]This result is also broadly consistent with studies documenting positive associations between emissions and market returns (i.e., proxies for discount rates) (e.g., Bolton and Kacperczyk, 2021b; Pástor et al., 2022).

[4]For more information, see https://climatetrace.org/approach.

lead to out-of-sample underestimation bias. To formally test the out-of-sample performance, I exploit a regulatory change in the United Kingdom (U.K.). Since April 2019, large U.K. firms must provide carbon emissions information in accordance with the Streamlined Energy and Carbon Reporting (SECR) framework in their annual filings in a "comply-or-explain" manner. As a result, 76 firms began disclosing in 2020, which represents a 20% increase in the number of disclosing firms in the U.K. I use this significant increase in the share of disclosing firms to examine whether Bloomberg's emissions estimates understate newly disclosing firms' emissions. Indeed, I find that there is, on average, a significantly negative estimation error for newly disclosing firms' emissions but virtually no bias for always-disclosing firms. This finding is consistent with Bloomberg failing to perfectly correct for the selection in its training sample of voluntarily disclosing firms.

After validating the three building blocks of my analytical framework, I test the main prediction of my model: intermediaries' estimates deter firms' disclosures. I exploit a plausibly exogenous shift in intermediary coverage: the expansion of Trucost in 2016 following its acquisition by S&P. The acquisition allowed Trucost to tap the universe of S&P's Compustat and SNL Financial. As a result of the new data and resources, Trucost doubled its coverage, adding around 7,000 firms globally to the Trucost universe. The extent of the expansion, however, varied across markets, depending on the availability of the required inputs for Trucost's estimation procedure. I exploit this variation in the market-level intensity of the coverage expansion in a continuous-treatment difference-in-differences design. Using a sample of 49,690 firm-year observations across 42 markets, I document robust evidence that firms in markets experiencing a larger coverage expansion are significantly less likely to disclose after the expansion than firms in markets experiencing limited coverage expansions. Importantly, the implied economic magnitude of the deterrence effect is non-trivial: doubling the coverage of emissions estimates in a market is associated with a 1.5-percentage-point decrease in the propensity of disclosing scope 1 emissions. This represents an approximately 16% reduction relative to the pre-2016 average disclosure propensity.

Encouraged by the empirical support for my model, I perform counterfactual analyses to investigate the consequences of refinements in intermediaries' estimates and the consequences of mandatory disclosure regulations. To this end, I calibrate my model to the data using the method of moments. My model estimates suggest that a substantial proportion of investors (~65%) appear to naively rely on intermediaries' estimates.[5] There is notable information asymmetry between naive and rational investors. Naive investors suffer especially from the lack of disclosure and intermediaries' biased estimates. Their information loss—defined as the mean squared error of their posterior beliefs about firms' emissions—is six times higher than that of rational investors. If intermediaries were able to correct their estimates for selection bias, disclosure would increase substantially (~40%), and the information asymmetry between naive and rational investors would decrease significantly (~25%). Interestingly, more realistic refinements of the intermediaries' estimates, which increase model precision but do not correct the bias, would exacerbate the information asymmetry between naive and rational investors, as naive investors rely even more on the biased estimates. By contrast, regulatory interventions such as disclosure mandates are helpful, especially to naive investors. Mandates attenuate the bias in intermediaries' estimates by reducing the selection underlying the sample of disclosing firms. More firm disclosures and less biased intermediary estimates together help narrow the information gap between naive and rational investors. For example, a mandate that randomly requires 30% of firms to disclose would reduce the information asymmetry by around 20%. Thus, mandatory disclosure is a potent strategy for enhancing the overall quality of information concerning corporate emissions, counteracting the distortions casued by intermediaries' estimates.

My study contributes to the literature on firms' emissions disclosures. Several studies explore the determinants and consequences of firms' emissions disclosures (e.g., Matsumura

[5]I acknowledge that this quantification might be overstated since I do not include other disclosure frictions (e.g., Verrecchia, 1983; Dye, 1985) in my model. Accordingly, my model interprets any deviation from full disclosure as evidence of investor naivete. Given this limitation, I advise caution in interpreting the model-implied magnitudes as upper bounds.

et al., 2014; Akbaş and Canikli, 2018; Jaggi et al., 2018; Bolton and Kacperczyk, 2021b; Flammer et al., 2021; Bond and Zeng, 2022; Cohen et al., 2023; Ilhan et al., 2023; Bonetti et al., 2024). Despite growing interest in these disclosures, only around 15,000 firms around the world voluntarily report their emissions (CDP, 2023). This disclosure gap has prompted calls for regulatory interventions such as disclosure mandates (e.g., Bolton et al., 2021; Carattini et al., 2022; Leuz, 2022; Rajgopal, 2022; Busch et al., 2023; Greenstone et al., 2023). It has also spurred the entry of information intermediaries offering emissions estimates as a market response to the disclosure gap. My study offers novel analytical and empirical insights on this response, highlighting the limitations and consequences of relying solely on intermediaries' emissions estimates to address the disclosure gap. Notably, my analysis indicates that disclosure regulation can mitigate the bias by purging self-selection in the sample of disclosing firms that intermediaries use to calibrate their models. Thus, intermediaries and disclosure regulation may act as *complements* rather than substitutes in the space of emissions disclosure.

My study also contributes to the literature on the measurement of corporate emissions. Credible measurement of emissions holds firms accountable and fosters effective social planning. One stream of the literature develops and describes methods to estimate facility- or firm-level emissions based on spatial emissions data measured through emissions monitors and satellites (e.g., Esparza et al., 2023; Hobbs et al., 2023; Breuer et al., 2025). Another stream examines the quality of firm-level emissions measurements. Several studies, for example, document bias in firms' disclosed emissions (e.g., Kim and Lyon, 2011, 2015; Kadach et al., 2024; Zhang, 2024; Lepere et al., 2025). More pertinent to my study, recent studies raise questions about the quality of intermediaries' emissions estimates and scores. Aswani et al. (2024) find that emissions estimates are often a simple function of a few firm fundamentals (e.g., sales) and do not sufficiently capture within-industry differences in carbon efficiency. In addition, several studies provide evidence that intermediaries appear to bias their estimates to cater to their stakeholders (e.g., Halper et al., 2022; Tang et al., 2022;

Agrawal et al., 2023). I add to this literature by documenting a novel bias in intermediaries' estimates that arises from firms' selective disclosures. My study highlights the importance of addressing firms' selection into disclosures if self-disclosed data are used for emissions modeling.

My study is also broadly related to various other streams of literature. It adds to a nascent literature exploring how big data technologies (re)shape investors' information processing and firms' responses. Mukherjee et al. (2021) suggest that satellite data may subsume information disclosed through conventional channels. Kang et al. (2021) find that alternative data reduce information asymmetry between firms and investors. Firms also adapt to technological changes by disciplining their investment decisions and changing their disclosure tactics (e.g., Zhu, 2019; Cao et al., 2023). My study analyzes the impacts of a new breed of data providers that leverage advances in statistical modeling (e.g., machine learning) to estimate corporate emissions. I document evidence that firms strategically exploit favorable model-based emissions estimates, which might impede the quality of information concerning corporate emissions. My study is also connected to the literature on how behavioral frictions affect firms' voluntary disclosure and the role of disclosure regulation (e.g., Blankespoor et al., 2019; Zhou and Zhou, 2020; Bertomeu, 2024; Bourveau et al., 2025; Huang et al., 2025). In line with this literature, I allow some investors to deviate from rational expectations (e.g., due to lack of strategic thinking, limited attention, or high processing costs). Lastly, my study relates to the literature on intermediaries and corporate disclosure. This literature documents that intermediaries can compel firms to disclose more through enhanced monitoring (e.g., Belnap, 2023). It also suggests, however, that intermediaries may deter firms' disclosure, if they produce an alternative signal that substitutes for firms' disclosure (e.g., Einhorn, 2018; Frenkel et al., 2020; Breuer et al., 2022). My study adds to this literature by providing yet another reason intermediaries may lead to less disclosure: they may inadvertently bias some investors' assessments of firms' environmental performance, which appears particularly relevant in the context of firms' emissions disclosures.

2 Analytical Framework

2.1 Model Setup

To explore how intermediaries' estimates affect firms' voluntary disclosure, I build on the seminal model of disclosure by Grossman (1981) and Milgrom (1981), and enrich it with features motivated by relevant institutional details and intermediaries' practices.

2.1.1 Investors

In my model, investors care about firms' environmental performance when allocating capital to firms in the economy (e.g., Hartzmark and Sussman, 2019; Krueger et al., 2020). Each firm's environmental performance is summarized by an emissions metric e. The common prior of firms' emissions is given by $e = \mu + \varepsilon^e$ where $\varepsilon^e \sim \mathcal{N}(0, \sigma_e^2)$. Consistent with recent empirical evidence (e.g., Chava, 2014; Bolton et al., 2022; Pástor et al., 2021, 2022), I assume that the higher the emissions, the lower the investors are willing to bid.[6]

Investors need to learn about firms' emissions by processing firms' voluntary disclosures and other relevant information. For firms without emissions disclosures, investors cannot rely on the disclosed information. That said, they can use the choice of non-disclosure as an informative signal about the firms' emissions. Rational investors understand that whether to disclose is a strategic decision. Hence, they expect non-disclosing firms to exhibit high emissions.

Not all investors may act rationally, though. Some investors may miss the negative signal

[6]Prior literature provides mixed explanations for why investors care about corporate emissions. Some studies suggest that certain investors have non-pecuniary preferences for green assets, i.e., a warm glow, (e.g., Barber et al., 2021; Pástor et al., 2021; Bonnefon et al., 2022; Goldstein et al., 2024). Others suggest that investors price cash flow risks associated with emissions (e.g., Bolton and Kacperczyk, 2021a; Ilhan et al., 2021; Derrien et al., 2025), while Pedersen et al. (2021) and Friedman et al. (2024) consider both possibilities. My model is agnostic about the reasons investors care about emissions. What matters is that they punish firms with higher emissions, as later supported by my evidence.

implied by firms' non-disclosure (e.g., Zhou and Zhou, 2020; Jin et al., 2021).[7] Accordingly, I allow a fraction λ of investors to be naive. These investors fail to internalize non-disclosure as a negative signal. In the same vein, these investors neglect any potential selection bias in the sources of information they use to revise their priors. This modeling choice allows for deviations from full rationality, in line with recent work calling for a careful (re)consideration of rationality assumptions in disclosure games (e.g., Blankespoor et al., 2020; Bertomeu, 2024). I denote naive investors' beliefs as $\mathbb{E}^{\nu}[e|\mathcal{I}]$ to differentiate them from rational investors' $\mathbb{E}[e|\mathcal{I}]$, where $\mathcal{I}$ is the set of information available to investors.

2.1.2 Firms

Firms strategically decide whether to disclose their privately observed measurement of their emissions—denoted by $\hat{s}$—or to withhold the information. The measurement is an unbiased signal of the underlying emissions, i.e., $\hat{s} = e + \varepsilon^s$, where $\varepsilon^s \sim \mathcal{N}(0, \sigma_s^2)$ is the measurement noise. The firms' objective is to maximize their valuation—or equivalently, to minimize investors' penalization on their emissions—through their disclosure ($d = \hat{s}$) or non-disclosure ($d = \varnothing$) choice.For parsimony, I assume that firms are penalized based on a weighted average of naive and rational investors' beliefs regarding their emissions: $\mathbb{E}_{\lambda}[e|\mathcal{I}] \equiv \lambda\mathbb{E}^{\nu}[e|\mathcal{I}] + (1-\lambda)\mathbb{E}[e|\mathcal{I}]$.[8] When firms' disclosures are the only source of information, each firm's choice is characterized by

$$\min_{d \in \{\hat{s}, \varnothing\}} \mathbb{E}_{\lambda}[e|d] \tag{1}$$

If all investors are rational, all firms will disclose; that is, disclosure unravels. However,

[7] Relatedly, the behavioral literature studies the role of bounded rationality in the form of level-k thinking in strategic games (e.g., Nagel, 1995; Bosch-Domenech et al., 2002; Brown et al., 2012; Bourveau et al., 2025). More broadly, a growing literature documents that economic agents exhibit a cognitive bias of selection neglect in their belief formation process (e.g., Juslin et al., 2007; Jehiel, 2018; Enke, 2020; Han et al., 2023).

[8] The reduced-form characterization of firms' objectives is consistent with empirical evidence that prices reflecting average beliefs, and is adopted in recent studies of firms' disclosure behaviors facing investors with bounded rationality (e.g., Zhou and Zhou, 2020; Aghamolla and Smith, 2023). The setup also echoes the weighted average structure derived in prior analytical studies featuring heterogeneous (risk-averse) investor groups (e.g., Daniel et al., 2001; Hirshleifer and Teoh, 2003). In Appendix A, I show that the main results continue to hold when firms' valuation is determined by market clearing among risk-averse investors.

in the presence of some naive investors who fail to properly punish non-disclosing firms, the seminal unraveling result breaks down. Firms with poorer performance (i.e., higher emissions) do not find it in their best interest to disclose. Absent disclosure, naive investors evaluate firms using the prior mean μ, i.e., $\mathbb{E}^{\nu}[e|d = \varnothing] = \mu$. The muted non-disclosure penalty among naive investors weakens the unraveling force. As a result, only lower-emitting firms disclose in equilibrium, while higher-emitting firms refrain disclosing, taking advantage of naive investors' valuations based on the prior mean. The equilibrium disclosure strategy follows a threshold structure, which is formally characterized in the following lemma.

Lemma 1. *Without an intermediary, firms disclose if and only if* $\hat{s}-\mu \leq \iota^N$. *The equilibrium disclosure threshold* $\iota^N \in \mathbb{R}$ *solves*

$$\iota^N = (1-\lambda)\mathbb{E}\left[\hat{s}-\mu|\hat{s}-\mu \geq \iota^N\right] \tag{2}$$

for $\lambda \in (0,1)$. *When* $\lambda \to 1$, $\iota^N \to 0$. *When* $\lambda \to 0$, $\iota^N \to +\infty$, *i.e., disclosure unravels.*

2.1.3 The Intermediary

Incomplete disclosure creates scope for information intermediaries to matter, as investors may look to intermediaries for information on non-disclosing firms. In my framework, I consider a representative information intermediary, which adopts an estimation procedure that closely mirrors intermediaries' real-world practice. The intermediary develops a proprietary statistical model of firms' emissions. The intermediary trains the statistical model using the observed relation between disclosed emissions and firm-level characteristics. The intermediary then extrapolates the fitted model to predict all firms' emissions, including non-disclosing firms' emissions. I assume that the intermediary's emissions model follows the form $e = x+\varepsilon^x$, where $x \equiv f(z)$, and z is a vector of observable features that collectively

explain variations in firms' emissions.[9] In practice, intermediaries' models typically incorporate features such as financial data (e.g., sales), geographical information (e.g., location), and industry-specific characteristics (e.g., fuel usage) that are available for both disclosing and non-disclosing firms. Intuitively, x can be interpreted as the expected level of emissions given a firm's observed activities. The error term $\varepsilon^x \sim \mathcal{N}(0, \sigma_x^2)$ captures variations in firms' emissions not explained by the observable features and is assumed to be independent of these features. The smaller the variation in the error term, the closer the intermediary's model is to the "true" data-generating process of firms' emissions.

Investors revise their beliefs upon observing the intermediary's model-based emissions estimates. In an ideal scenario where the true value of x is known, investors' updated beliefs should be $\mathbb{E}[e|x] = x$; in other words, x "overwrites" the common prior μ.[10] In reality, however, x is unknown and must be estimated. Consistent with major intermediaries' practice (e.g., Trucost, 2020; Bloomberg, 2021; MSCI, 2024), I assume that the representative intermediary in my framework understands the functional form of f and estimates its parameters empirically using disclosed emissions $\{\hat{s}\}$. For tractability, I adopt a linear functional form for f.[11] The econometric model the intermediary uses in its estimation is given by $\hat{s} = f(z) + \varepsilon$, where $\varepsilon \equiv \varepsilon^x + \varepsilon^s$ represents the intermediary's specification error, which is a composite of unexplained emissions ε^x and firms' measurement noise ε^s.[12] If disclosures are random, the ordinary least squares (OLS) estimator $\hat{x} \equiv \hat{f}(z)$ is consistent, i.e., the estimator

[9] Since μ is common knowledge, the intermediary's role can be viewed as modeling the unknown emissions ε^e; that is, μ is part of $f(z)$ function (e.g., as part of the constant term). As long as the intermediary's emissions modeling is not completely irrelevant, $\sigma_x^2 < \sigma_e^2$, which means that the intermediary's model is an "upgrade" to investors' common prior with respect to explaining variations in firms' emissions.

[10] Due to the exogeneity assumption, i.e., $\mathbb{E}[\varepsilon^x|z] = 0$ and hence $\mathbb{E}[\varepsilon^x|x] = 0$, the revised belief after knowing x satisfies $\mathbb{E}[e|x] = \mathbb{E}[x + \varepsilon^x|x] = x$. The prior μ becomes redundant as the intermediary's model captures strictly more variations in firms' emissions than the common prior does.

[11] This assumption is less restrictive than it may seem because the regressors can be arbitrarily transformed (e.g., polynomial). In addition, any model nonlinear in its parameters can be piece-wise linearized, allowing for a linear representation of the estimation model.

[12] Following the convention in the literature of assuming mutually independent errors, I assume that ε^x and ε^s are independent such that $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ where $\sigma^2 \equiv \sigma_x^2 + \sigma_s^2$. Here, I also assume that the measurement error ε^s is unrelated to any firm characteristic in z such that the exogeneity condition, i.e., $\mathbb{E}[\varepsilon z] = 0$, is not violated when estimating the model.

$\hat{x}$ approaches its true value x as the number of (randomly) disclosing firms increases.[13]

The intermediary's estimates will be biased if disclosures are correlated with the intermediary's specification error ε. In this case, the estimator $\hat{x}$ asymptotically equals its true value x plus a selection bias $\mathcal{B} \equiv \mathbb{E}[\varepsilon|d = \hat{s}]$.[14] If neither the intermediary nor the investors properly correct for the selection bias, it will distort investors' beliefs about corporate emissions and, in turn, affect firms' disclosure incentives.

Correcting for the selection bias requires knowledge of firms' disclosure strategies. In my framework, the intermediary does not correct for the selection bias. In practice, the intermediary may recognize the selection in firms' disclosures but chooses not to impose the correction for a couple of reasons. First, the intermediary is likely evaluated by outsiders based on the revealed quality of its estimates. For naive investors especially, one plausible metric to gauge the intermediary's performance is the in-sample fit, i.e., the accuracy of the intermediary's estimates in predicting firms' disclosed emissions. Although emissions estimates $\hat{x}$ are subject to the selection bias, the bias does not manifest in the in-sample fit (i.e., $\mathbb{E}[\hat{s} - \hat{x}|d = \hat{s}] = 0$). Conversely, correcting for the self-selection bias and releasing x would result in globally unbiased estimates (i.e., $\mathbb{E}[\hat{s} - x] = 0$) but biased in-sample estimates (i.e., $\mathbb{E}[\hat{s} - x|d = \hat{s}] = \mathcal{B} \neq 0$). Second, the intermediary may be uncertain about the proper underlying economic model to account for the selection in firms' voluntary disclosures.[15] This model uncertainty could make the intermediary refrain from explicitly addressing the selection, opting instead to remain agnostic rather than take risks by imposing definitive structural modeling assumptions.

[13] More generally, consistency is ensured as long as disclosure is unrelated to the intermediary's specification error ε, i.e., $\mathbb{E}[\varepsilon \mathbb{1}_{d=\hat{s}}] = 0$. For example, if disclosure is related to an observable feature (e.g., market capitalization), the OLS estimator $\hat{x}$ will still provide a consistent estimate of x.

[14] In the baseline model, the selection bias $\mathcal{B}$ will always equal the Heckman omitted term $\mathbb{E}[\varepsilon|d = \hat{s}]$. I present scenarios where the selection bias can only be represented by a linear approximation of the omitted term in Online Appendix I.

[15] To isolate the impacts of intermediaries, my model abstracts from other factors that may also affect firms' decisions to voluntarily disclose emissions. Analyses incorporating these factors are provided in Online Appendix II.

2.2 Equilibrium Analysis

In the presence of the intermediary, firms disclose their emissions measurements $\hat{s}$ if and only if disclosure yields favorable outcomes. Compared to the game without the intermediary, the structure of firms' objective changes because investors now have access to an additional piece of information: the intermediary's estimates $\hat{x}$. Figure 1 illustrates the timeline of the voluntary disclosure game with the intermediary, which mirrors the sequence of events in real-world practice. Despite the multi-period presentation of the game in Figure 1, the game is in essence a stage game given independent and identically distributed (i.i.d.) error terms. The problem faced by firms is given by

$$\min_{d \in \{\hat{s}, \varnothing\}} \mathbb{E}_\lambda[e|\hat{x}, d] \tag{3}$$

where $\mathbb{E}_\lambda[e|\hat{x}, d] \equiv \lambda \mathbb{E}^\nu[e|\hat{x}, d] + (1-\lambda)\mathbb{E}[e|\hat{x}, d]$ is the weighted average beliefs across the two types of investors. Naive investors take the intermediary's estimates at face value; that is, $\mathbb{E}^\nu[e|\hat{x}] = \hat{x}$, regardless of any potential bias in $\hat{x}$. In contrast, rational investors adjust for the selection bias using rational expectations such that $\mathbb{E}[e|\hat{x}] = x$.[16] When firms disclose their measurements $\hat{s}$, both types of investors update their beliefs using Bayes' rule. However, if firms withhold their measurements, naive investors' beliefs remain intact, while rational investors revise their beliefs using rational expectations. The weighted average beliefs $\mathbb{E}_\lambda[e|\mathcal{I}]$, where $\mathcal{I} = \{\hat{x}, d\}$, are given by

$$\mathbb{E}_\lambda[e|\hat{x}, d] = \begin{cases} \lambda\left(\frac{\kappa}{\kappa+1}\hat{x} + \frac{1}{\kappa+1}\hat{s}\right) + (1-\lambda)\left(\frac{\kappa}{\kappa+1}x + \frac{1}{\kappa+1}\hat{s}\right) & d = \hat{s} \\ \lambda\hat{x} + (1-\lambda)\mathbb{E}\{\frac{\kappa}{\kappa+1}x + \frac{1}{\kappa+1}\hat{s}|\hat{x}, \varnothing\} & d = \varnothing \end{cases} \tag{4}$$

[16] Even though the intermediary's model f is proprietary information, rational investors are able to reverse-engineer the true x from $\hat{x}$ if the bias is an only function of known parameters in a rational expectation equilibrium, if exists. This applies to the firms as well.

where $\kappa \equiv \frac{\tau_x}{\tau_s} \equiv \frac{\sigma_s^2}{\sigma_x^2}$ is the relative precision of the intermediary's model-based estimates compared to firms' measurements. The equilibrium disclosure strategy is formally characterized by the following proposition.

Proposition 1. *With a representative intermediary who does not correct for the selection bias in its estimates $\hat{x}$, the equilibrium disclosure strategy follows a threshold structure w.r.t. $\varepsilon \leq \iota$ where $\iota \in \mathbb{R}$ such that*

$$\iota = (1-\lambda)\mathbb{E}\left[\varepsilon|\varepsilon \geq \iota\right] + \lambda\mathbb{E}[\varepsilon|\varepsilon \leq \iota] \tag{5}$$

for $\lambda \in (0,1)$. When $\lambda \to 0$, $\iota \to +\infty$, i.e., disclosure unravels; when $\lambda \to 1$, $\iota \to -\infty$, i.e., there is virtually no disclosure.

A comparison of Proposition 1 and Lemma 1 underscores the crucial impact of the intermediary's estimates—the availability of these estimates further *deters* firms' disclosure in the presence of naive investors. The first term on the right-hand side of Equation 5 represents the unraveling force driven by $1-\lambda$ rational investors, which is also seen on the right-hand side of Equation 2, i.e., the equation that characterizes firms' disclosure strategies in the absence of the intermediary's estimates. The second term on the right-hand side of Equation 5 represents an additional pooling force driven by λ naive investors' inability to correct for the equilibrium selection bias $\mathcal{B} = \mathbb{E}[\varepsilon|\varepsilon \leq \iota] < 0$. The negative bias leads to an underestimation of firms' expected emissions, which naive investors fail to recognize. Consequently, compared to the case without the intermediary, more firms would opt for non-disclosure to take advantage of the favorable bias in the intermediary's estimates and, in turn, in naive investors' beliefs. As fewer firms disclose, the selection problem worsens, leading to a magnified bias that further dampens firms' disclosure incentives. In other words, the pooling force, just like the unraveling force, is self-reinforcing. In the extreme case where nearly all investors are naive, i.e., $\lambda \to 1$, the disclosure mass is infinitely thin, i.e., $\iota \to -\infty$.

In this case, only the best-performing firms disclose, leading the intermediary's estimation to treat all firms as though they are of the best possible type. For intermediate values of λ, the unraveling force and the pooling force ultimately balance each other in equilibrium. However, this equilibrium entails a lowered disclosure propensity than in the equilibrium in which investors do not have access to the intermediary's estimates. This constitutes the main prediction of my analytical analysis, which is formulated in the following corollary.

Corollary 1. *Given $\lambda \in (0,1)$, standard normal p.d.f. ϕ, and standard normal c.d.f. Φ, the standardized disclosure thresholds with and without the intermediary, denoted by $\tilde{\iota}$ and $\tilde{\iota}^N$ respectively, solve*

$$\tilde{\iota} = (1-\lambda)\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})} + \lambda\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})} \tag{6a}$$

$$\tilde{\iota}^N = (1-\lambda)\frac{\phi(\tilde{\iota}^N)}{1-\Phi(\tilde{\iota}^N)} \tag{6b}$$

Importantly, the solutions are unique and satisfy $\tilde{\iota} < \tilde{\iota}^N$. Hence, the disclosure mass $\mathbb{E}d = \Phi(\tilde{\iota}) < \Phi(\tilde{\iota}^N) = \mathbb{E}d^N$. Moreover, $\frac{\partial \tilde{\iota}}{\partial \lambda} < 0$, $\frac{\partial \tilde{\iota}^N}{\partial \lambda} < 0$.

Overall, the analytical framework highlights two channels through which naive investors affect firms' disclosure incentives. First, the presence of naive investors—who fail to properly penalize non-disclosures—always prevents full unraveling, regardless of the availability of the intermediary's estimates. Second, in the presence of the intermediary, naive investors' inability to correct for the selection bias in the intermediary's estimates further weakens the unraveling force. Both effects become more pronounced as the share of naive investors increases.

Figure 2 graphically illustrates the major insights of Corollary 1 by plotting the standardized disclosure thresholds and the corresponding disclosure propensities over a range of $\lambda \in (0,1)$. Comparing disclosure propensities in the presence of the intermediary (the solid line) versus in the absence of the intermediary (the dashed line) demonstrates the deterrence effect of the intermediary's emissions estimates. When there is a non-trivial number

of naive investors, the introduction of the intermediary's estimates is expected to cause a sharp decline in firms' disclosure incentives—a testable hypothesis that I examine in detail in Section 3.3.

3 Empirical Analyses

In this section, I empirically validate the main building blocks of my model—investors penalize firms' emissions, higher-emitting firms abstain from disclosure, and intermediaries produce biased estimates. I then test the main prediction of the model: that the intermediaries' estimates deter firms' disclosure.

3.1 Data

I gather data for my empirical analyses from multiple sources. The firm-level emissions data come from two prominent information intermediaries: Trucost and Bloomberg. They collect disclosed emissions from sources such as the CDP repository, firms' annual and/or sustainability reports, and corporate websites. Importantly, both Trucost and Bloomberg offer statistical model-based estimates of firms' emissions, with their models trained on data from disclosing firms, as assumed in my analytical framework.[17] Bloomberg, a major intermediary that launched its GHG emissions estimates in 2019, offers a unique advantage over its peers by providing access to both its own emissions estimates and the emissions disclosed by firms, if they are available. This feature enables me to compare the properties of disclosed and estimated emissions for the *same* firm, which is crucial for validating the main building blocks of my analytical framework. The Bloomberg sample consists of 3,210 U.S. firms from 2020 to 2022. Trucost's environmental dataset is accessible through WRDS. Notably, Trucost's acquisition by S&P Dow Jones Indices in 2016 doubled the coverage of

[17]Emissions data provided by intermediaries are highly correlated. Busch et al. (2022) document that Trucost, MSCI, and Sustainalytics provide roughly the same scope 1 emissions for disclosing firms and highly correlated emissions estimates (with correlations between 0.75 to 0.88) for non-disclosing firms.

its emissions estimates. The Trucost sample consists of 14,144 firms across 42 markets from 2014 to 2017. I supplement these emissions data with firm-level fundamental data and price data from Compustat and CRSP.

Table 1 tabulates summary statistics for the data used in my empirical analyses. All emissions and firm-level characteristics are transformed using the inverse hyperbolic sine function.[18] Emissions disclosures are infrequent: in the Bloomberg sample, 28% of U.S. firms disclose their scope 1 emissions, while globally, only 10% of firms in the Trucost sample report their scope 1 emissions directly to CDP or in their annual filings. Firms with higher emissions tend to disclose more, as these firms likely concentrate in heavily emitting industries and hence face higher demands from their investors and stakeholders for greater transparency regarding their emissions. There is virtually no bias in Bloomberg's emissions estimates in the sample of disclosing firms—the economic magnitude of the in-sample bias among U.S. firms is negligible—consistent with what Bloomberg claims in its white papers (e.g., Bloomberg, 2021, 2024). The expansion of Trucost's coverage in 2016 is significant. Half of the firms in the Trucost sample were newly covered in 2016. The median intensity of country-level expansion reached 178%, nearly tripling the pre-2016 coverage.

3.2 Model Validation

3.2.1 Investors' Emissions Pricing

The model assumes that investors punish firms with high emissions. Following prior studies (e.g., Matsumura et al., 2014; Griffin et al., 2017; Homburg et al., 2024), I examine whether higher-emitting firms receive a discount in their market valuations. Formally, I consider the

[18] The inverse hyperbolic sine behaves similarly to the logarithm but applies to the entire real line $\mathbb{R}$ (e.g., Bellemare and Wichman, 2020). Importantly, the transformation greatly reduces the skewness in the distribution of emissions. The more bell-shaped distribution of transformed emissions is more in accordance with the normality assumptions made throughout Section 2.

following monthly regression, which is akin to a log-log specification:

$$ME_{i,t} = \alpha_j + \alpha_t + \beta \cdot GHG_{i,t} + \gamma \cdot X_{i,t} + \epsilon_{i,t} \tag{7}$$

where $ME_{i,t}$ is the inverse hyperbolic sine of firm i's market value in year-month t; $GHG_{i,t}$ is the inverse hyperbolic sine of firm i's latest scope 1 estimated (GHG^E) or disclosed (GHG^A) emissions available in year-month t;[19] $X_{i,t}$ is a vector of firm characteristics—total assets, sales, operating profits, PP&E, capital expenditure, and debt—all transformed using the inverse hyperbolic sine; α_j and α_t are industry and year-month fixed effects, respectively. Standard errors are clustered at the firm level.

Panel A of Table 2 reports the results for examinations of investors' emissions pricing. Consistent with prior studies, I find that firms' emissions are strongly negatively associated with their market values. Additionally, Bloomberg's data allow me to focus on a distinct aspect of emissions pricing that is particularly relevant to my model—the relative pricing power of estimated versus disclosed emissions. Column 2 offers a first glance at the problem: the pricing power of estimated emissions does not decline when disclosure is also available. Columns 3 to 5, which focus on the subsample of disclosing firms, further reveal that investors on average do not appear to see much incremental value in disclosed emissions, which are by and large made public after Bloomberg's estimates.[20] Furthermore, as shown in Panel B of Table 2, the sensitivity of investors' pricing is higher for firms facing higher climate risk à la

[19] To ensure that the information on firms' scope 1 emissions is available at the time when firms are valued, I lag firm i's emissions in any fiscal year for six months. For example, for firm i whose fiscal year-end is December 31, its scope 1 information of year fiscal year 2020 is matched to months from 2021 July to 2022 June.

[20] According to Bloomberg, its estimates are first made available soon after firms' financial reports are published. In the U.S., firms normally take 45 to 90 days to file their 10-Ks after the fiscal year-end. On the other hand, while emissions metrics are sometimes mentioned in firms' 10-Ks, they are thus far mainly reported to CDP or discussed in firms' sustainability reports—both of which are typically not available until four or six months after the fiscal year-end. For example, Chevron, a major U.S. oil and gas company, filed its 2023 annual report in which only carbon reduction targets are discussed on February 23, 2024. Chevron's exact scope 1, 2, and 3 GHG emissions figures are disclosed in its 2023 performance summary (*link*), which was posted on April 4, 2024.

Sautner et al. (2023).[21] Collectively, these findings seem in line with the idea that investors care about firms' emissions and frequently rely on intermediaries' estimates, consistent with my model.[22]

3.2.2 Firms' Selective Disclosure

Firms in my model strategically disclose their emissions. Hence, I examine whether firms' decisions to disclose are negatively associated with their emissions level after controlling for their expected levels of emissions. To identify a measure of firm-level emissions that is independent of strategic disclosure and hence comparable between disclosing and non-disclosing firms, I use remote-sensed facility-level emissions data recently released by Climate TRACE.[23] I consider three different methods of aggregating facility-level emissions to the firm level. First, I sum over all facility-level emissions with non-zero ownership links. Second, I compute the weighted sum of facility-level emissions based on ownership percentages. Third, I consider "consolidating" facility-level emissions: full consolidation if the firm owns more than 50% of the facility and proportional consolidation otherwise. These three aggregate measures represent a set of noisy but selection-free signals of firm-level emissions. After matching with the Climate TRACE's facility-level emissions and ownership information, the sample consists of 57 unique U.S. firms' data in 2022.[24] The sample spans firms operating in heavily emitting industries, which are the focus of Climate TRACE.

[21] Consistent with the notion that firms' emissions are more closely tied to regulatory risks and transition opportunities than to physical risks, I find greater sensitivity in the pricing of emissions only among firms with higher $Risk^{OP}$ and $Risk^{RG}$, but not higher $Risk^{PH}$. In untabulated analysis, I likewise do not observe incremental pricing power of disclosed emissions among firms with higher climate risk.

[22] Investors may frequently rely on intermediaries' estimates, either because they are genuinely of high quality or because they are perceived as trustworthy. While results in Table 2 do not distinguish between these explanations, the analyses in Section 2.1.3 and Section 3.2.3 suggests that the former is unlikely.

[23] Relatedly, using emissions data reported by firms or estimated by Trucost, Shi et al. (2023) find that voluntarily disclosing firms generate lower scope 1 emissions. Unlike Shi et al. (2023), I abstain from using emissions estimated by intermediaries, as these estimates might not be credible measures of corporate emissions due to the potential selection bias discussed in Section 2.1.3 and Section 3.2.3.

[24] The ownership tree provided by Climate TRACE reports ownership information only for the year 2022 as of the time of my data retrieval. To avoid noise arising from changes in ownership, my analysis focuses solely on the year 2022.

Next, I examine whether lower-emitting firms are more likely to disclose, after controlling for scale of operations. Specifically, I estimate the following linear probability models:

$$\begin{aligned} Disclose_i &= \alpha + \beta \cdot GHG_i^R + \gamma \cdot X_i + \epsilon_i \\ Disclose_i &= \alpha + \beta \cdot (GHG_i^R - GHG_i^E) + \epsilon_i \end{aligned} \tag{8}$$

where $Disclose_i$ equals one if firm i provides scope 1 disclosure in the year 2022; GHG_i^R is the inverse hyperbolic sine of firm i's aggregated facility-level emissions in the year 2022. In these specifications, I consider two ways of controlling for the expected level of emissions. First, I control for the expected level of emissions using a set of firm-level characteristics that collectively capture the scale of firms' operations. Second, I subtract Bloomberg's estimated emissions GHG^E from the aggregated Climate TRACE remote-sensed emissions GHG^R. Figure 3 illustrates the results of estimating linear probability models outlined in Equation 8.[25] The panels consistently display downward-sloping fitted lines, which imply that firms with higher emissions, after controlling for scale of operations, are less likely to disclose. In addition, Table 3 presents the results of estimating the Probit model variants of the linear probability models outlined in Equation 8, which corroborate the graphical evidence. These findings, albeit statistically weak due to the small sample size, closely align with firms' strategic disclosure of emissions implied by my model.

3.2.3 The Intermediary's Bias

The model features a representative intermediary that extrapolates its model-based predictions to non-disclosing firms *without* adjusting for self-selection bias. As discussed in Section 2.1.3, this approach should lead to an unbiased in-sample fit and understated out-of-sample estimates. A look at Bloomberg's white papers (e.g., Bloomberg, 2021, 2024) suggests that Bloomberg does indeed focus on the in-sample fit. In particular, as displayed

[25] I visualize the estimation results using the `binscatter` method for a clearer and more robust view of the relationship between firms' disclosure decisions and their emissions (e.g., Cattaneo et al., 2024).

in Figure 4, Bloomberg uses a "recalibration" procedure that effectively maximizes in-sample fit. This focus on in-sample fit provides prima facie evidence consistent with the idea that Bloomberg does not explicitly correct for the selection bias in its emissions estimates.

Next, I examine the out-of-sample performance of Bloomberg's emissions estimates, which requires comparing estimated and disclosed emissions for firms that *do not* disclose voluntarily. To this end, I exploit a regulatory change in the U.K. that came into effect for fiscal years starting after April 1st, 2019. The regulation requires large U.K. firms to provide information on their GHG emissions in accordance with the SECR framework in their annual filings.[26] Following the regulatory shock, 76 firms started to disclose their GHG emissions in their 2020 annual filings, which represents a 20% increase in the number of disclosing firms in the U.K. capital markets. The model Bloomberg employed to predict firms' emissions for fiscal year 2020 is calibrated using information available prior to 2020. Subsequently, the model is fed with inputs extracted from firms' fiscal year 2020 filings to generate emissions estimates for that year. Therefore, the estimated emissions for firms that started to disclose due to the regulation constitute out-of-sample estimates.[27]

I use the significant increase in the number of disclosing firms to examine whether Bloomberg's emissions estimates understate newly disclosing firms' emissions. Figure 5 plots the comparison of signed estimation errors (i.e., $GHG^E - GHG^A$) for always-disclosing firms versus newly disclosing firms. In the first reporting cycle after the regulation was introduced, the estimated emissions for newly disclosing firms (i.e., out-of-sample firms) are, on average, much lower than their disclosed emissions. Meanwhile, the fit for always-disclosing firms

[26] Companies and LLPs are considered "large" if they meet two or more of the following criteria: i) a turnover of £36 million or more; ii) a balance sheet of £18 million or more, and iii) 250 employees or more. Prior to the 2019 "comply-or-explain" regulation, public U.K. firms listed on the main market of the London Stock Exchange (LSE) had already been required to disclose their scope 1 and scope 2 GHG emissions since 2013, which explains the relatively frequent disclosures made by U.K. firms prior to 2019.

[27] Admittedly, given the "comply-or-explain" nature of the regulation, it is possible that not all newly disclosing firms started to disclose because of the regulation. That said, the non-trivial increase in the number of firms disclosing gives me comfort that there is indeed a sharp increase in the pressure to disclose. Furthermore, any self-selection in the subsample of newly disclosing firms only attenuates the magnitude of any underestimation.

(i.e., in-sample firms) appears accurate, consistent with Bloomberg's claims in its white papers. Moreover, the negative estimation bias among newly disclosing firms diminishes in the following reporting cycles, which is expected given that these out-of-sample firms become in-sample firms when Bloomberg recalibrates its model. To assess the statistical significance of estimation errors in sample and out of sample, I consider the following cross-sectional regression:

$$Error_i = \alpha + \beta \cdot NewlyDisclose_i + \epsilon_i \tag{9}$$

where $Error_i$ is the signed estimation error, $GHG^E - GHG^A$, or the signed percentage estimation error, $GHG^E/GHG^A - 1$; $NewlyDisclose_i$ equals one if firm i did not disclose in 2019 but started to disclose in 2020 and equals zero if firm i always disclosed in both years; α is the constant term. Standard errors are clustered at the emission-tier level.[28]

Table 4 reports the comparison between estimation errors for always-disclosing firms (in-sample fit) and newly disclosing firms (out-of-sample fit). There is a significantly negative bias for newly disclosing firms' fiscal year 2020 emissions. While the in-sample bias is negligible, the out-of-sample underestimation is sizable, amounting to approximately 16% of non-disclosing firms' emissions, or 1.3% of these firms' arcsin-transformed emissions. Consistent with the graphical evidence, the significant negative bias among newly disclosing firms diminishes in reporting cycles subsequent to 2020. Additionally, the trend in R^2 also suggests that whether a firm is a newly disclosing firm in 2020 is a meaningful explanatory variable only in 2020.[29] Collectively, these findings are consistent with the notion that Bloomberg fails to perfectly adjust for the selection bias in its training sample.

[28]Firms are segmented into 50 groups (i.e., emission tiers) based on the level of their disclosed emissions. In untabulated analyses, the estimation yields similar results if robust standard errors are used to account for heteroscedasticity in estimation errors across firms. The results are also robust to including "red flag" disclosures and estimates—those with emissions below one.

[29]Prior empirical studies have documented evidence consistent with firms reduce their emissions following mandatory carbon disclosure rules (e.g., Downar et al., 2021; Jouvenot and Krueger, 2021; Yang et al., 2021; Tomar, 2023). If newly disclosing firms anticipated the regulation and began mitigating their emissions prior to 2020, any underestimation would be attenuated. The fact that I observe significant underestimation among newly disclosing firms, despite the potential anticipation effect working against my estimation, helps establish further robustness of my findings.

3.3 The Deterrence Effect of Intermediaries' Estimates

After validating the major building blocks of my model, I turn to testing my model's main prediction: the availability of intermediaries' emissions estimates deters firms' emissions disclosure. I exploit the expansion of Trucost in 2016—following the acquisition by S&P—as a plausibly exogenous shock to intermediary coverage. The acquisition allowed Trucost to tap the universe of S&P Global's Compustat and SNL Financial. Both databases offer rich firm-level data such as financials and industry classifications that are important inputs for Trucost's emissions estimation model. Thanks to the new resources, Trucost doubled its global coverage, adding new emissions estimates for around 7,000 firms to its environmental data pack. The extent of the expansion varies across countries, depending on the availability of the required inputs for Trucost's estimation procedure and the extent of local firms' emissions disclosures. I consider two ways of defining the treatment: one at the country level and the other at the firm level. The first treatment, *Expansion*, is calculated as the increase in the number of firms newly covered by Trucost in 2016 in a given country, scaled by the number of firms already covered by Trucost in 2015 in that country. The second treatment, $NewCoverage$, equals one if a firm is newly covered by Trucost in 2016 and zero otherwise. The two treatment definitions triangulate the treatment effect inference. The firm-level treatment compares firms within the same regulatory regime, which makes it robust to country-level trends in emissions disclosure. However, identification may be affected by differential trends between firms that were already covered (typically larger) and firms that are newly covered (typically smaller). By contrast, the country-level treatment compares average outcomes across firms within a country, which helps to smooth out size-related differences between larger and smaller firms.

Trucost provides detailed sources of its scope 1 emissions. In calculating the dependent variable, *Disclose*, I only account for disclosure of the *exact* value of scope 1 emissions to

CDP or in firms' annual filings such as 10-Ks and sustainability reports.[30]

Utilizing country- or firm-level variations in the expansion of Trucost's emissions estimates, I test my primary prediction through a difference-in-differences design involving a sample of 14,144 firms from 2014 to 2017.[31] The specification with the tightest set of fixed effects is as follows:

$$Disclose_{i,t} = \alpha_i + \alpha_{j,t} + \beta \cdot Treat_{c(i)} \times Post_t + \epsilon_{i,t} \tag{10}$$

where $Disclose_{i,t}$ equals one if firm i provides scope 1 disclosure in year t; $Treat_c(i)$ is one of the two treatments defined earlier; $Post_t$ equals one if t is greater than or equal to the treatment year, i.e., year 2016; α_i and $\alpha_{j,t}$ are firm fixed effects and industry-year fixed effects, respectively. Standard errors are clustered at the country level or at the firm level, depending on the source of variation in the treatment.

Table 5 reports results of the difference-in-difference model outlined in Equation 10. I find robust evidence of the deterrence effect of Trucost's estimates on firms' scope 1 disclosures across different fixed effects structures and different treatment definitions. The economic magnitude of the deterrence effect is non-trivial: doubling the coverage of emissions estimates in a market is associated with a reduction in the propensity to disclose by 1.1-1.5 percentage points, which represents 12-16% of the average propensity for disclosure in 2015. I acknowledge that the intensity of the expansion might be correlated with other factors driving firms' scope 1 GHG disclosure over time.[32] To address these potential confounding

[30] I do not include partial disclosures (e.g., "Estimate based on partial data disclosure in CDP" or "Estimate used instead of disclosure - data does not cover global operations"). While firms disclose relevant information in these cases, the incompleteness or ambiguity in their disclosures leads Trucost to use estimated values instead. Moreover, firms may selectively disclose information used as inputs in Trucost's estimation (e.g., fuel usage), which, in the scope of my analytical framework, effectively functions as non-disclosure.

[31] I drop U.K. firms and French firms from my sample as both countries introduce mandatory or semi-mandatory *firm-level* emissions reporting rules around the sample period. As previously discussed, public U.K. firms listed on the LSE main market have been required to disclose their GHG emissions since 2013. Meanwhile, the *Grenelle II* law in France has required French organizations with more than 500 employees to produce a carbon emissions report since 2012 (e.g., Bourveau et al., 2024).

[32] In fact, since Trucost relied heavily on firms' voluntary disclosure to compile its environmental dataset,

factors, I evaluate the parallel trends assumption of my difference-in-difference design by examining the dynamic treatment effect of coverage expansion on direct disclosure. The graphical evidence presented in Figure 6 is consistent with the parallel trends assumption. Furthermore, the specifications using the firm-level treatment *NewCoverage* allows for comparisons of firms operating within the same country. As shown in columns 5 and 6 of Table 5, the deterrence effect remains robust when using this alternative firm-level treatment. The economic magnitude of the treatment effect is also similar: coverage by Trucost reduces the propensity to disclose by 1.1 percentage points. Taken together, the empirical results are consistent with the main modeling prediction that the availability of intermediaries' emissions estimates deters firms' voluntary emissions disclosures. Notably, there is still an overall increasing trend in firms' emissions disclosures, while the introduction of Trucost's coverage *slows down* the increasing trend.

4 Further Implications

Collectively, my analytical framework and empirical evidence suggest that intermediaries' emissions estimates are subject to selection bias, and the biased estimates deter firms' emissions disclosures. These findings raise concerns when looking towards intermediaries such as Trucost and Bloomberg to address the lack of disclosure on firms' emissions.[33]

Encouraged by the explanatory power of my analytical framework, I further take the model to real-world data through structural estimation to examine the roles of intermediaries and disclosure mandates in shaping the information environment regarding corporate

the treatment is higher for markets with lower levels of disclosure prior to the 2016 Trucost expansion. These include major developing markets (e.g., greater China, India) and non-EU developed markets (e.g., the U.S., Japan, and South Korea). If firms in these markets are catching up with regard to sustainability reporting, the differential trends between high- and low-treatment markets, if any, should be expected to attenuate the deterrence effect. By contrast, if firms in these markets experience less peer pressure to disclose, the differential trends between high- and low-treatment markets would amplify the estimated deterrence effect.

[33]The SEC's climate-related risk disclosure rule (*link*) drew significant opposition. One argument holds that the availability intermediaries' emissions estimates make the rule unnecessary (Henderson and Parker, 2023).

emissions. Using structurally estimated model parameters, I quantify the impacts of intermediaries' estimates on investors' evaluations of corporate emissions and explore how alternative approaches by intermediaries could affect these impacts. Additionally, I analyze the interactions between intermediaries and disclosure mandates, showing how the two function as complements rather than substitutes in filling the information gap concerning corporate emissions.

4.1 Structural Estimation

I conduct a structural estimation of my model using the sample of U.S. firms from 2020 to 2022.[34] Key parameters of my model can be estimated using the method of moments. The identification of modeling parameters is discussed in Appendix B. Table 6 reports the results of the structural estimation. The estimation suggests that there is a substantial share of naive investors—65% of investors fail to rationally punish non-disclosing firms.

The structurally estimated parameters allow me to take a step further in quantifying the economic consequences of the intermediaries' approach. The results are also reported in Table 6. First, I quantify the magnitude of the overall bias and in particular the bias for non-disclosing firms. The estimation suggests that the overall bias equals -0.60 in the unit of arcsin-transformed emissions, and the bias for non-disclosing firms equals in -0.79 in the unit of arcsin-transformed emissions. The magnitude of the overall bias accounts for 6.7% of an average firm's transformed estimated emissions, while the magnitude of the non-disclosing firms' bias accounts for 9.4% of an average non-disclosing firm's transformed estimated emissions.[35]

[34] To conservatively quantify magnitudes of the specification error ε and the self-selection bias $\mathcal{B}$, I exclude "red flag" disclosures and estimates—those with emissions below one—and further winsorize estimation errors such that the distribution more closely resembles the assumed normal distribution.

[35] The structurally estimated out-of-sample bias is larger than the reduced-form estimate presented in Section 3.2.3. This is expected, as a key distinction between the U.S. and U.K. markets is that there is currently no firm-level emissions disclosure mandate in the U.S., whereas public U.K. firms listed on the LSE main market have been required to disclose their GHG emissions since 2013. Section 4.3 further discusses the impact of disclosure mandates on intermediaries' estimates.

Second, following Bertomeu et al. (2020), I examine investors' information loss in the decision making process (e.g., capital allocation). The information loss, denoted by $\mathcal{L}$, is defined as the mean squared errors in investors' posterior beliefs of firms' emissions. The formal definition and analytical solution of investors' information loss are presented in Appendix C. The first observation is that disclosure helps reduce all investors' information loss, which is intuitive as one more signal of emissions helps investors improve the precision of their posteriors. When firms are equipped with perfect measurement of their emissions (i.e., $\kappa = 0$), all investors' information loss is zero if firms disclose (i.e., $\mathcal{L}^d = 0$), since all weights in investors' posteriors are placed on firms' perfect measurement. The second observation, echoing discussions in Section 2.2, is that naive investors' failure to penalize non-disclosing firms directly harms them and that the bias in intermediaries' estimates exacerbates the loss. The estimation, as shown in the last two columns of Table 6, suggests that naive investors' ex ante information loss is more than six times that of rational investors. When firms' emissions measurement is imperfect (i.e., $\kappa > 0$), the information asymmetry between naive and rational investors becomes more pronounced.[36]

4.2 Refining Intermediaries' Estimates

Equipped with the parameter estimates, I explore how refinements in intermediaries' estimates affect firms' disclosures and investors' information loss.

First, I examine the consequences of improvements in intermediaries' emissions modeling. In my analytical framework, the precision of intermediaries' emissions model is captured by $\tau_x \equiv \sigma_x^{-2}$. While more precise emissions modeling (i.e., higher τ_x and hence higher κ) will not affect the equilibrium disclosure strategy, it does affect investors' information loss by

[36] In fact, firms themselves estimate that their emissions measurements are subject to an average error rate of 25% to 30%, according to a survey by BCG (*link*). Relatedly, García-Vega et al. (2023) find that firms' self-reported emissions to the CDP are error-prone in that breakdowns fail to add up to totals, while Brander et al. (2023) document frequent inconsistencies between the emissions disclosed in annual reports and those reported to the CDP for the same firms.

changing the relative precision κ and the variance of the specification error σ^2. Intriguingly, when firms disclose, naive investors' information loss, relative to rational investors', increases as intermediaries refine their models. The intuition is that intermediaries' biased estimates receive higher weights as intermediaries' emissions models become more precise. When firms do not disclose, however, all investors' information loss decreases due to both the reduced total variance σ^2 and the narrowed distribution of the modeling error ε^x. Overall, as intermediaries refine their model, the ex ante information asymmetry between naive and rational investors widens, as shown in Panel A of Figure 7.

Next, I explore what happens if intermediaries correct for the selection bias in their estimates. De-biasing effectively makes intermediaries release the "true" expected emissions x. The correction mutes firms' non-disclosure incentives arising from naive investors' inability to correct for the selection bias in intermediaries' estimates. Hence, firms' disclosure propensity increases, as suggested by Corollary A.1. The estimation result in Panel A of Table 7 suggests that the correction would lead to a 40% increase in the percentage of firms disclosing.

The selection-free estimates benefit naive investors and rational investors to different extents. Provided with selection-free estimates, naive investors are able to form unbiased prior beliefs of corporate emissions, which eliminates the information asymmetry between naive and rational investors when firms disclose. That said, naive investors still struggle to properly penalize non-disclosing firms. Indeed, the consequence of their failure to penalize non-disclosure is larger, as the non-disclosing firms under the new equilibrium, where intermediaries de-bias their estimates, are of worse types.[37] Additionally, the increased disclosure propensity also shifts more weight toward the information loss conditional on disclosure, affecting the ex ante evaluation of information loss. The estimated information loss in Panel A of Table 7 highlights the overall positive impact of the correction vis-à-vis mitigating the

[37]Relatedly, because only worse-performing firms remain non-disclosing, the variance of the modeling error ε^x conditional on non-disclosure decreases, benefiting both naive and rational investors equally.

information asymmetry between naive and rational investors: the extent of the information asymmetry may decrease by at least 25%.

4.3 The Role of Disclosure Mandates

Having demonstrated the limitations of relying solely on intermediaries to address issues of insufficient disclosure and information asymmetry between naive and rational investors, I now explore the potential roles of disclosure mandates.

Disclosure mandates mitigate self-selection in the sample of disclosing firms, attenuating the selection bias in intermediaries' estimates. Suppose a regulation randomly requires a fraction θ of firms to disclose. Corollary A.2 characterizes the equilibrium disclosure strategy in the presence of the disclosure mandates and underscores how the disclosure mandates interact with intermediaries' estimates. In essence, disclosure mandates prompt a partial correction of the selection bias without requiring intermediaries to initiate the correction. The attenuated bias increases firms' voluntary disclosure incentives, leading to additional disclosures beyond those required by the mandates. The magnitudes of both the shrinkage in the selection bias and the incremental disclosure are significant even for smaller-scale mandates. For example, a mandate requiring 30% of firms to disclose reduces the selection bias by over 60% and prompts approximately 20% more firms to voluntarily disclose. In terms of the eventual disclosure outcome, the mandate is as effective as having the intermediary correct the bias, as seen in Panel B of Table 7.[38] To the extent that intermediaries' emissions models are trained on global data, a regional disclosure mandate may produce spillover effects in regions outside the mandate's scope.

Disclosure mandates help reduce the information asymmetry between naive and rational investors. While all investors' information loss decreases as disclosure mandates apply to

[38] If the mandate is assigned based on firm size—as in the forthcoming EU Corporate Sustainability Reporting Directive (Directive 2022/2464)—rather than randomly, the interaction effect will be attenuated. For detailed comparisons between randomized and size-based disclosure mandates, please refer to Online Appendix II.3.

more firms, naive investors' information loss decreases more rapidly. This reduction in the information asymmetry is expected, as disclosure mandates function similarly to partially debiasing the intermediary's estimates. Panel B of Figure 7 suggests that a mandate requiring 30% of firms to disclose reduces the information asymmetry between naive and rational investors by around 20%. Overall, my analysis highlights the complementarity between the intermediaries' estimates and disclosure mandates, offering a novel perspective to the ongoing debate surrounding the costs and benefits of mandatory climate-related disclosure rules.

5 Conclusion

I study the impact of information intermediaries' emissions estimates on firms' emissions disclosures. First, I analytically model how firms' voluntary disclosure decisions and intermediaries' estimates are co-determined in the presence of naive investors. The model predicts that intermediaries' estimates amplify the existing friction caused by investors' bounded rationality and hence deter firms' voluntary disclosures. Next, I empirically validate the main building blocks of the analytical model and document evidence consistent with the main prediction of the model—that intermediaries' estimates deter firms' voluntary disclosures. Lastly, I calibrate model parameters using real-world data. Equipped with estimated parameters, I quantify the potential consequences of refining intermediaries' estimates. I also analyze the effect of disclosure mandates and consider the policy implications.

This study by no means suggests that intermediaries should stop providing statistical model-based estimates of firms' emissions. If intermediaries properly account for self-selection in firms' voluntary disclosures, their estimates will benefit all investors, particularly those who are less informed about the implications of strategic disclosures. Admittedly, explicitly correcting for the selection bias is challenging, as the the exact quantities of the adjustments depend heavily on modeling choices. That said, if the goal is to improve the overall quality of information on firms' emissions instead of prettifying the in-sample esti-

mates, intermediaries should not be afraid to make mistakes or to make firms feel pressured. Climate TRACE's approach and its welcoming attitude towards public scrutiny exemplify such a mindset, as suggested in the following excerpt from a Science article:

> *"... One thing is for certain: As Climate TRACE grows more comprehensive, and its claims more definitive, it will draw more scrutiny, and opposition. But both opposing countries and industries are welcome to make their case. 'Our numbers are public,' McCormick says. 'And anyone who has different numbers will need to have a reason they're different.' "* —Voosen (2023)

My study has important limitations. First, while I show that my simple stage-game framework featuring naive investors has notable explanatory power, as it aligns well with multiple patterns in the data, I cannot rule out the possibility that these patterns may also be consistent with alternative interpretations. For one, consistent with the notion of satisficing (Simon, 1955), some sustainability-labeled investors only require supportable data to fulfill their investment mandates (e.g., Berg et al., 2024; Kim, 2024; Choi et al., 2025). Such "checking-the-box" investors affect firms' disclosure strategies in a manner similar to the naive investors in the model, but point to a conceptually distinct characterization. For another, the deterrence effect of intermediaries' estimates on firms' disclosures could also be driven by these estimates crowding out firms' voluntary disclosures through economic forces that I have not modeled (e.g., Einhorn, 2018; Frenkel et al., 2020; Breuer et al., 2022; Banerjee et al., 2024). Second, while a distinctive aspect of my study is the use of structural estimations to provide insights into quantities beyond mere directions, this exercise needs to be interpreted with caution. My analytical framework abstracts away from several factors that are likely important in explaining firms' decisions to voluntarily disclose emissions. For example, it might be costly for firms to track their own emissions, which suggests another reason firms abstain from disclosures (e.g., Amiraslani et al., 2024). The omission of such frictions may lead to an over-interpretation of the widespread non-disclosure as being driven

solely by my proposed forces (i.e., investors' naivete and intermediaries' approach). The omission could also lead to an over-estimation of the potential consequences of refinements in intermediaries' estimates and disclosure mandates.

References

Aghamolla, C. and Smith, K. (2023). Strategic complexity in disclosure. *Journal of Accounting and Economics*, 76(2-3):101635.

Agrawal, S., Liu, L. Y., Rajgopal, S., Sridharan, S. A., Yan, Y., and Yohn, T. L. (2023). ESG ratings of ESG index providers. *Working Paper*.

Akbaş, H. E. and Canikli, S. (2018). Determinants of voluntary greenhouse gas emission disclosure: An empirical investigation on Turkish firms. *Sustainability*, 11(1):107.

Amiraslani, H., Chen, X., Ormazabal, G., and Pope, P. F. (2024). Climate disclosure regulation and investor wealth. *Working Paper*.

Aswani, J., Raghunandan, A., and Rajgopal, S. (2024). Are carbon emissions associated with stock returns? *Review of Finance*, 28(1):75–106.

Banerjee, S., Marinovic, I., and Smith, K. (2024). Disclosing to informed traders. *The Journal of Finance*, 79(2):1513–1578.

Barber, B. M., Morse, A., and Yasuda, A. (2021). Impact investing. *Journal of Financial Economics*, 139(1):162–185.

Bellemare, M. F. and Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, 82(1):50–61.

Belnap, A. (2023). The effect of intermediary coverage on disclosure: Evidence from a randomized field experiment. *Journal of Accounting and Economics*, 75(1):101522.

Berg, F., Heeb, F., and Kölbel, J. (2024). The economic impact of esg ratings. *Working Paper*.

Bertomeu, J. (2024). Disclosure paternalism. *Journal of Accounting and Economics*, 77(2-3):101662.

Bertomeu, J., Ma, P., and Marinovic, I. (2020). How often do managers withhold information? *The Accounting Review*, 95(4):73–102.

Blankespoor, E., deHaan, E., and Marinovic, I. (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, 70(2-3):101344.

Blankespoor, E., Dehaan, E., Wertz, J., and Zhu, C. (2019). Why do individual investors disregard accounting information? The roles of information awareness and acquisition costs. *Journal of Accounting Research*, 57(1):53–84.

Bloomberg (2021). Distributional greenhouse gas emissions estimates: Data challenges and modeling solutions. *Working Paper*.

Bloomberg (2024). Bloomberg's greenhouse gas emissions estimates model: A summary of challenges and modeling solutions. *Working Paper*.

Bolton, P., Halem, Z., and Kacperczyk, M. (2022). The financial cost of carbon. *Journal of Applied Corporate Finance*, 34(2):17–29.

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

Bolton, P. and Kacperczyk, M. T. (2021b). Carbon disclosure and the cost of capital. *Working Paper*.

Bolton, P., Reichelstein, S., Kacperczyk, M. T., Leuz, C., Ormazabal, G., and Schoenmaker, D. (2021). Mandatory corporate carbon disclosures and the path to net zero. *Management and Business Review*, 1(3).

Bond, P. and Zeng, Y. (2022). Silence is safest: Information disclosure when the audience's preferences are uncertain. *Journal of Financial Economics*, 145(1):178–193.

Bonetti, P., En, Y. E., Kadach, I., and Ormazabal, G. (2024). Climate disclosures and decarbonization along the supply chain. *Working Paper*.

Bonnefon, J.-F., Landier, A., Sastry, P., and Thesmar, D. (2022). The moral preferences of investors: Experimental evidence. *Working Paper*.

Bosch-Domenech, A., Montalvo, J. G., Nagel, R., and Satorra, A. (2002). One, two, (three), infinity, ...: Newspaper and lab beauty-contest experiments. *American Economic Review*, 92(5):1687–1701.

Bourveau, T., Breuer, M., and Stoumbos, R. (2025). Learning to disclose: Disclosure dynamics in the 1890s streetcar industry. *The Review of Financial Studies*, 38(9):2602–2651.

Bourveau, T., Garel, A., and Romec, A. (2024). Mandatory carbon disclosure: Evidence from France. *Working Paper*.

Brander, M., Hoepner, A. G., Rogelj, J., Saini, T., and Schneider, F. (2023). Corporate carbon disclosure: A critical review. *Working Paper*.

Breuer, M., Hombach, K., and Müller, M. A. (2022). When you talk, I remain silent: Spillover effects of peers' mandatory disclosures on firms' voluntary disclosures. *The Accounting Review*, 97(4):155–186.

Breuer, M., Rauter, T., Muhn, M., Schütt, H., Barry, T., and Li, R. (2025). Corporate footprints. *Working Paper*.

Brown, A. L., Camerer, C. F., and Lovallo, D. (2012). To review or not to review? Limited strategic thinking at the movie box office. *American Economic Journal: Microeconomics*, 4(2):1–26.

Busch, T., Cho, C. H., Hoepner, A. G., Michelon, G., and Rogelj, J. (2023). Corporate greenhouse gas emissions' data and the urgent need for a science-led just transition: Introduction to a thematic symposium. *Journal of Business Ethics*, 182(4):897–901.

Busch, T., Johnson, M., and Pioch, T. (2022). Corporate carbon performance data: Quo vadis? *Journal of Industrial Ecology*, 26(1):350–363.

Cao, S., Jiang, W., Yang, B., and Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of AI. *The Review of Financial Studies*, 36(9):3603–3642.

Carattini, S., Hertwich, E., Melkadze, G., and Shrader, J. G. (2022). Mandatory disclosure is key to address climate risks. *Science*, 378(6618):352–354.

Cattaneo, M. D., Crump, R. K., Farrell, M. H., and Feng, Y. (2024). On binscatter. *American Economic Review*, 114(5):1488–1514.

CDP (2023). CDP 2023 disclosure data factsheet. *CDP (link)*.

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

Choi, S., Ferri, F., and Macciocchi, D. (2025). Do mutual funds respond to mechanical changes in esg ratings? *Working Paper*.

Cohen, S., Kadach, I., and Ormazabal, G. (2023). Institutional investors, climate disclosure, and carbon emissions. *Journal of Accounting and Economics*, 76(2-3):101640.

Daniel, K. D., Hirshleifer, D., and Subrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *The Journal of Finance*, 56(3):921–965.

Derrien, F., Krueger, P., Landier, A., and Yao, T. (2025). ESG news, future cash flows, and firm value. *The Journal of Finance*, n/a(n/a).

Downar, B., Ernstberger, J., Reichelstein, S., Schwenen, S., and Zaklan, A. (2021). The impact of carbon disclosure mandates on emissions and financial operating performance. *Review of Accounting Studies*, 26(3):1137–1175.

Dye, R. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research*, 23(1):123–145.

Einhorn, E. (2018). Competing information sources. *The Accounting Review*, 93(4):151–176.

Enke, B. (2020). What you see is all there is. *The Quarterly Journal of Economics*, 135(3):1363–1398.

Esparza, Á. E., Rowan, G., Newhook, A., Deglint, H. J., Garrison, B., Orth-Lashley, B., Girard, M., and Shaw, W. (2023). Analysis of a tiered top-down approach using satellite and aircraft platforms to monitor oil and gas facilities in the Permian basin. *Renewable and Sustainable Energy Reviews*, 178:113265.

Flammer, C., Toffel, M. W., and Viswanathan, K. (2021). Shareholder activism and firms' voluntary disclosure of climate change risks. *Strategic Management Journal*, 42(10):1850–1879.

Frenkel, S., Guttman, I., and Kremer, I. (2020). The effect of exogenous information on voluntary disclosure and market quality. *Journal of Financial Economics*, 138(1):176–192.

Friedman, H. L., Heinle, M. S., and Luneva, I. (2024). Implications of introducing investor-focused ESG reporting. *Working Paper*.

García-Vega, S., Hoepner, A. G., Rogelj, J., and Schiemann, F. (2023). Abominable greenhouse gas bookkeeping casts serious doubts on climate intentions of oil and gas companies. *Working Paper*.

Goldstein, I., Kopytov, A., Shen, L., and Xiang, H. (2024). On ESG investing: Heterogeneous preferences, information, and asset prices. *Working Paper*.

Greenstone, M., Leuz, C., and Breuer, P. (2023). Mandatory disclosure would reveal corporate carbon damages. *Science*, 381(6660):837–840.

Griffin, P. (2017). CDP carbon majors report 2017. *Carbon Disclosure Project (link)*.

Griffin, P. A., Lont, D. H., and Sun, E. Y. (2017). The relevance to investors of greenhouse gas emission disclosures. *Contemporary Accounting Research*, 34(2):1265–1297.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics*, 24(3):461–483.

Halper, J., Grieve, D., Shriver, T., Cadwalader, W., and Taft, L. (2022). ESG ratings: A call for greater transparency and precision. *Harvard Law School Forum on Corporate Governance (link)*.

Han, B., Hirshleifer, D., and Walden, J. (2023). Visibility bias in the transmission of consumption beliefs and undersaving. *The Journal of Finance*, 78(3):1647–1704.

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

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. 5(4):475–492.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Henderson, T. and Parker, D. (2023). Mandated emissions disclosures: The bad exceeds the good. *RealClearMarkets (link)*.

Hirshleifer, D. and Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of accounting and economics*, 36(1-3):337–386.

Hobbs, M., Kargar, A. R., Couture, H., Freeman, J., Söldner-Rembold, I., Ferreira, A., Jeyaratnam, J., O'Connor, J., Lewis, J., Koenig, H., et al. (2023). Inferring carbon dioxide emissions from power plants using satellite imagery and machine learning. *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*.

Homburg, C., Lapp, L. O., and Schick, R. (2024). The impact of greenhouse gas emissions on firm value: A loglinear model. *Working Paper*.

Huang, Z., Piotroski, J. D., and Xie, C. L. (2025). Strategic disclosure and investor loss aversion. *Working Paper*.

Ilhan, E., Krueger, P., Sautner, Z., and Starks, L. T. (2023). Climate risk disclosure and institutional investors. *The Review of Financial Studies*, 36(7):2617–2650.

Ilhan, E., Sautner, Z., and Vilkov, G. (2021). Carbon tail risk. *The Review of Financial Studies*, 34(3):1540–1571.

Jaggi, B., Allini, A., Macchioni, R., and Zagaria, C. (2018). The factors motivating voluntary disclosure of carbon information: Evidence based on Italian listed companies. *Organization & Environment*, 31(2):178–202.

Jehiel, P. (2018). Investment strategy and selection bias: An equilibrium perspective on overoptimism. *American Economic Review*, 108(6):1582–1597.

Jin, G. Z., Luca, M., and Martin, D. (2021). Is no news (perceived as) bad news? An experimental investigation of information disclosure. *American Economic Journal: Microeconomics*, 13(2):141–173.

Jouvenot, V. and Krueger, P. (2021). Mandatory corporate carbon disclosure: Evidence from a natural experiment. *Working Paper*.

Juslin, P., Winman, A., and Hansson, P. (2007). The naïve intuitive statistician: A naïve sampling model of intuitive confidence intervals. *Psychological review*, 114(3):678.

Kadach, I., Koo, M., Martin, X., and Zhao, M. (2024). Decipher market responses to Climate TRACE emission data release. *Working Paper*.

Kang, J. K., Stice-Lawrence, L., and Wong, Y. T. F. (2021). The firm next door: Using satellite images to study local information advantage. *Journal of Accounting Research*, 59(2):713–750.

Kim, E.-H. and Lyon, T. P. (2011). Strategic environmental disclosure: Evidence from the DOE's voluntary greenhouse gas registry. *Journal of Environmental Economics and Management*, 61(3):311–326.

Kim, E.-H. and Lyon, T. P. (2015). Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, 26(3):705–723.

Kim, S. (2024). Investor preferences and responses to disclosure: Evidence from carbon net-zero pledges. *Working Paper*.

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

Lepere, M., Maso, S., Dong, Y., and Aikman, D. (2025). Emissions dissonance: Examining how firm-level under-reporting undermines policy. *Working Paper*.

Leuz, C. (2022). How best to report carbon emissions. *Chicago Booth Review (link)*.

Li, Q., Watts, E. M., and Zhu, C. (2024). Retail investors and esg news. *Journal of Accounting and Economics*, 78(2-3):101719.

Matsumura, E. M., Prakash, R., and Vera-Muñoz, S. C. (2014). Firm-value effects of carbon emissions and carbon disclosures. *The Accounting Review*, 89(2):695–724.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2):380–391.

MSCI (2024). Greenhouse gas emissions methodologies. *MSCI (link)*.

Mukherjee, A., Panayotov, G., and Shon, J. (2021). Eye in the sky: Private satellites and government macro data. *Journal of Financial Economics*, 141(1):234–254.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, 85(5):1313–1326.

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

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

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

Rajgopal, S. (2022). Why I support the SEC's proposed climate risk disclosure rules. *Forbes (link)*.

Sautner, Z., Van Lent, L., Vilkov, G., and Zhang, R. (2023). Firm-level climate change exposure. *The Journal of Finance*, 78(3):1449–1498.

Shi, Y., Tang, C. S., and Wu, J. (2023). Are firms voluntarily disclosing emissions greener? *Working Paper*.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, pages 99–118.

Tang, D., Yan, J., and Yao, Y. (2022). The determinants of ESG ratings: Rater ownership matters. *Available at SSRN*.

Tomar, S. (2023). Greenhouse gas disclosure and emissions benchmarking. *Journal of Accounting Research*, 61(2):451–492.

Trucost (2020). Trucost environmental data methodology guide. *S&P Global*.

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, 5:179–194.

Voosen, P. (2023). Al Gore's climate watchdog spots emission gaps. *Science*, 382(6675):1096–1097.

Yang, L., Muller, N. Z., and Liang, P. J. (2021). The real effects of mandatory CSR disclosure on emissions: Evidence from the greenhouse gas reporting program. *NBER Working Paper*, (w28984).

Zhang, L. (2024). Carbon emission underreporting: Evidence from satellite emission data. *Working Paper*.

Zhou, F. S. and Zhou, Y. (2020). The dog that did not bark: Limited price efficiency and strategic nondisclosure. *Journal of Accounting Research*, 58(1):155–197.

Zhu, C. (2019). Big data as a governance mechanism. *The Review of Financial Studies*, 32(5):2021–2061.

A Risk-aversion Investors

In this appendix, I prove that the main analytical results discussed in Section 2 continue to hold qualitatively when firms' valuation is determined by market clearing among risk-averse investors.

Let π denote the short-term market valuation of the firm. It is determined by the fundamental value v_f, net of a valuation penalty proportional to emissions e, $\pi = v_f - \varphi e$, where $\varphi > 0$ captures the sensitivity of firm value to emissions. To isolate the effect of firms' emissions disclosures—and consistent with the notion that earnings-related information is resolved prior to emissions disclosure decisions—I hold v_f fixed. In other words, all uncertainty in π arises from uncertainty in e.

Investors have mean-variance preferences over their terminal payoffs $\omega(\pi - p)$, where ω is their holdings of the firm's asset and p is the firm price. Assuming zero net supply of the firm's asset, market clearing yields $p = \frac{\lambda\mathbb{E}^\nu[\pi|\mathcal{I}]\mathbb{V}[\pi|\mathcal{I}]+(1-\lambda)\mathbb{E}[\pi|\mathcal{I}]\mathbb{V}^\nu[\pi|\mathcal{I}]}{\lambda\mathbb{V}[\pi|\mathcal{I}]+(1-\lambda)\mathbb{V}^\nu[\pi|\mathcal{I}]}$. While not identical, the equilibrium pricing retains the structure of a weighted average of naive and rational investors' beliefs, similar to the reduced-form setup. The difference is that weights deviate from $(\lambda, 1-\lambda)$ by incorporating naive and rational investors' posterior variances. Let p_e denote the penalization on emissions, $p_e = -\varphi\frac{\lambda\mathbb{E}^\nu[e|\mathcal{I}]\mathbb{V}[e|\mathcal{I}]+(1-\lambda)\mathbb{E}[e|\mathcal{I}]\mathbb{V}^\nu[e|\mathcal{I}]}{\lambda\mathbb{V}[e|\mathcal{I}]+(1-\lambda)\mathbb{V}^\nu[e|\mathcal{I}]}$. Firms' problems are therefore characterized by $\max_{d\in\{\hat{s},\varnothing\}} p_e$.

A.1 No Intermediary

In the absence of an intermediary, investors' priors are given by $e \sim \mathcal{N}(\mu, \sigma_e^2)$. Investors form expectations as described in Section 2. That is, if firms disclose, $\mathbb{E}[e|\hat{s}] = \mathbb{E}^\nu[e|\hat{s}] = \frac{\tau_e}{\tau_e+\tau_s}\mu + \frac{\tau_s}{\tau_e+\tau_s}\hat{s}$; if firms do not disclose, $\mathbb{E}[e|\varnothing] = \frac{\tau_e}{\tau_e+\tau_s}\mu + \frac{\tau_s}{\tau_e+\tau_s}\mathbb{E}[\hat{s}|\varnothing]$, $\mathbb{E}^\nu[e|\varnothing] = \mu$. Regarding posterior variances, I likewise assume all investors calculate posteriors variances using the Bayes' rule if firms disclose, i.e., $\mathbb{V}^\nu[e|\hat{s}] = \mathbb{V}[e|\hat{s}] = \frac{1}{\tau_e+\tau_s}$. Furthermore, rational investors form non-disclosure beliefs using rational expectations, $\mathbb{V}[e|\varnothing] = \left(\frac{\tau_s}{\tau_e+\tau_s}\right)^2\mathbb{V}[\hat{s}|\varnothing] + \frac{1}{\tau_e+\tau_s}$, while naive investors ignore the implications of non-disclosure and thus retain the prior variance, $\mathbb{V}^\nu[e|\varnothing] = \frac{1}{\tau_e}$.[39] In equilibrium, firms' disclosure strategy follows a threshold structure, which is formally characterized in the following lemma.

Lemma A.1. *Without an intermediary, firms disclose if and only if $\hat{s} - \mu \leq i^N$. The equilibrium disclosure threshold $i^N \in \mathbb{R}$ solves*

$$i^N = (1 - \Lambda^N)\mathbb{E}[\hat{s} - \mu | \hat{s} - \mu \geq i^N] \tag{A.1}$$

[39] The derivation of $\mathbb{V}[e|\varnothing]$ follows the law of total variance, $\mathbb{V}[e|\varnothing] = \mathbb{V}[\mathbb{E}[e|\hat{s}]|\varnothing] + \mathbb{E}[\mathbb{V}[e|\hat{s}]|\varnothing]$.

where $\Lambda^N \equiv \frac{\lambda \mathbb{V}[e|\varnothing]}{\lambda \mathbb{V}[e|\varnothing]+(1-\lambda)\mathbb{V}^\nu[e|\varnothing]}$, $\mathbb{V}^\nu[e|\varnothing] = \frac{1}{\tau_e}$, $\mathbb{V}[e|\varnothing] = \left(\frac{\tau_s}{\tau_e+\tau_s}\right)^2 \mathbb{V}[\hat{s}-\mu|\hat{s}-\mu \geq i^N] + \frac{1}{\tau_e+\tau_s}$, $\lambda \in (0,1)$. *The uniqueness of the equilibrium is established in Online Appendix III.*

The partial disclosure result remains when investors are risk averse. That said, there is more disclosure in the case of risk aversion, driven by the greater valuation weight placed on rational investors' beliefs under non-disclosure—that is, $\Lambda^N < \lambda$, and thus $1-\Lambda^N > 1-\lambda$.[40] Moreover, under risk aversion, the prior precision τ_e also affects equilibrium disclosure. The lower the prior precision, the more disclosure in equilibrium. The intuition is that, when the prior precision is low, rational investors, relative to naive investors, extract more information from non-disclosure, allowing them to narrow the range of possible emissions realizations to a greater extent. As a result, their beliefs carry greater weight in equilibrium pricing.

A.2 With Intermediary

In the presence of an intermediary, naive investors' valuation is subject to the self-selection bias embedded in the intermediary's estimates $\hat{x}$. Investors' second-moment beliefs follow the structure in the "no intermediary" case, except that the pre-disclosure precision is given by τ_x rather than τ_e. The equilibrium disclosure strategy still follows a threshold structure, and is formally characterized by the following proposition.

Proposition A.1. *With a representative intermediary who does not correct for the selection bias in its estimates* $\hat{x}$*, the equilibrium disclosure strategy follows a threshold structure w.r.t.* $\varepsilon \leq i$ *where* $i \in \mathbb{R}$ *such that*

$$i = (1-\Lambda)\mathbb{E}[\varepsilon|\varepsilon \geq i] + \Lambda \mathbb{E}[\varepsilon|\varepsilon \leq i] + \kappa(\Lambda-\lambda)\mathbb{E}[\varepsilon|\varepsilon \leq i] \tag{A.2}$$

where $\Lambda \equiv \frac{\lambda \mathbb{V}[e|\hat{x},\varnothing]}{\lambda \mathbb{V}[e|\hat{x},\varnothing]+(1-\lambda)\mathbb{V}^\nu[e|\hat{x},\varnothing]}$, $\mathbb{V}^\nu[e|\hat{x},\varnothing] = \frac{1}{\tau_x}$, $\mathbb{V}[e|\hat{x},\varnothing] = \left(\frac{\tau_s}{\tau_x+\tau_s}\right)^2 \mathbb{V}[\varepsilon|\varepsilon \geq i] + \frac{1}{\tau_x+\tau_s}$, $\lambda \in (0,1)$. $\kappa \equiv \frac{\tau_x}{\tau_s}$ *is the relative precision. The uniqueness of the equilibrium is established in Online Appendix III.*

Compared to Proposition 1, the unraveling incentive is stronger and the pooling force weaker for similar reasons discussed above; that is, $\Lambda < \lambda$. Additionally, notice the presence of an additional term, $\kappa(\Lambda-\lambda)\mathbb{E}[\varepsilon|\varepsilon \leq i]$, which further reduces the pooling incentive.[41]

[40] $\Lambda^N < \lambda$ because $\mathbb{V}[e|\varnothing] < \mathbb{V}^\nu[e|\varnothing]$. That is, upon non-disclosure, rational investors form more precise (i.e., lower-variance) posterior beliefs than naive investors, as they can infer and narrow down the range of potential $\hat{s}$ realizations. Mathematically, the difference in posterior variances is $\mathbb{V}^\nu[e|\varnothing] - \mathbb{V}[e|\varnothing] = \frac{1}{\tau_e} - \left(\frac{\tau_s}{\tau_e+\tau_s}\right)^2 \mathbb{V}[\hat{s}|\varnothing] - \frac{1}{\tau_e+\tau_s} > \frac{1}{\tau_e} - \frac{\tau_s}{\tau_e}\frac{1}{\tau_e+\tau_s} - \frac{1}{\tau_e+\tau_s} = 0$. The same argument applies in the presence of an intermediary.

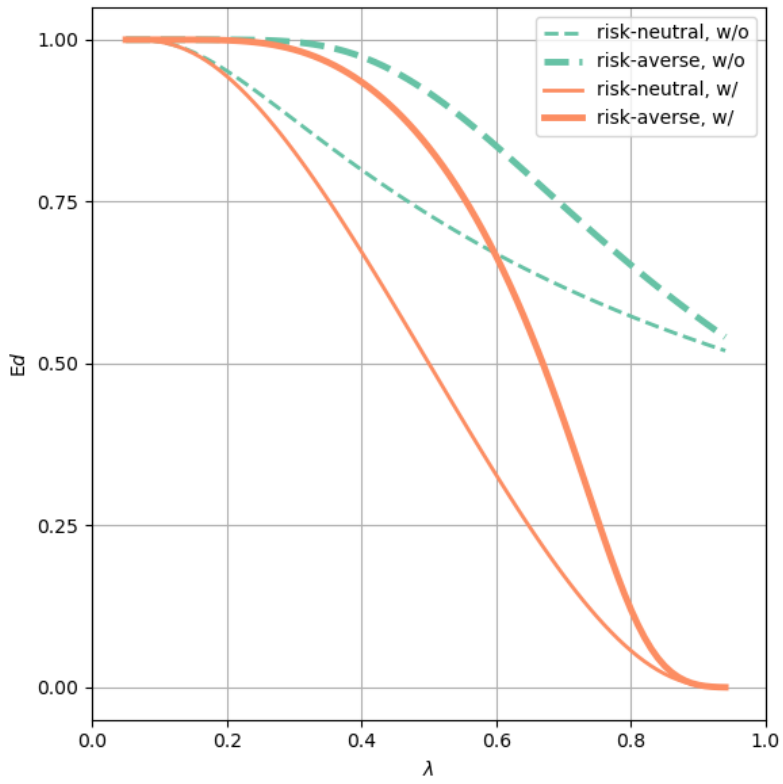
[41] There is still overall positive pooling: $\Lambda + \kappa(\Lambda-\lambda) > 0$, as $\Lambda > \frac{\mathbb{V}[e|\hat{x},\varnothing]}{\mathbb{V}^\nu[e|\hat{x},\varnothing]}\lambda > \frac{\kappa}{\kappa+1}\lambda$.

This arises because the favorable bias, $\mathbb{E}[\varepsilon|\varepsilon \leq i]$, receives a higher valuation weight under disclosure than under non-disclosure, given that $\lambda > \Lambda$. The term effectively adjusts for this difference in weighting. As in the "no intermediary" case, the quality of the intermediary's modeling affects equilibrium disclosure. Crucially, this further implies that—holding the prior precision fixed—the more variation the intermediary's model explains (i.e., higher τ_x), the stronger the deterrence effect. While this modeling feature is appealing, it introduces additional challenges for structural estimation, as the precisions, τ_e, τ_x, and τ_s are hardly identifiable, even in relative terms.

Overall, the core modeling prediction the the intermediary's emissions estimates deter firms' voluntary emissions disclosures remains intact under risk aversion, as graphically illustrated in Figure A.1 below.

Figure A.1: Equilibrium disclosure propensity $\mathbb{E}d$, risk aversion

This figure overlays the numerically solved equilibrium disclosure for $\lambda \in (0, 1)$ from Lemma A.1 and Proposition A.1 onto Figure 2, using the parameterization $\sigma_e = 4$, $\sigma_x = 3$, $\sigma_s = 2$.



B Model Identification

This appendix describes the identification of modeling parameters (λ, σ) in detail. I first establish the identification of modeling parameters λ and σ. The moment condition that uniquely identifies λ is as follows:

$$\mathbb{E}d = \Phi(i(\lambda)) \tag{A.3}$$

where $\mathbb{E}d$ is the shorthand for the disclosure propensity $\mathbb{E}[\mathbb{1}_{d=\hat{s}}]$, Φ is standard normal c.d.f., and $i : (0,1) \to \mathbb{R}$ is an invertible function mapping λ to the standardized equilibrium disclosure threshold $\tilde{\iota}$ characterized in Equation 6a. The function i is invertible because i maps every $\lambda \in (0,1)$ to a unique $\tilde{\iota} \in \mathbb{R}$ (i.e., i is onto) and i is strictly decreasing (i.e., i is one-to-one). The invertibility of i ensures that for any $\tilde{\iota}$, there is a unique λ that maps back to it. The empirical counterpart of $\mathbb{E}d$ is given by the empirical disclosure mass $\bar{d} \equiv \frac{1}{N}\sum_i \mathbb{1}_{d_i=\hat{s}_i}$ where N is total number of firms in a given market and $\mathbb{1}_{d_i=\hat{s}_i}$ equals one if firm i discloses. Knowing $\bar{d}$, the estimated standardized disclosure threshold $\hat{\tilde{\iota}}$ can be computed by inverting Φ, i.e., $\hat{\tilde{\iota}} = \Phi^{-1}(\bar{d})$. Given $\hat{\tilde{\iota}}$, the estimated share of naive investors $\hat{\lambda}$ can be solved by rearranging terms in Equation 6a, i.e., $\hat{\lambda} = \frac{\hat{\tilde{\iota}}-u}{l-u}$ where $u \equiv \frac{\phi(\hat{\tilde{\iota}})}{1-\Phi(\hat{\tilde{\iota}})}$ and $l \equiv \frac{-\phi(\hat{\tilde{\iota}})}{\Phi(\hat{\tilde{\iota}})}$.[42]

The second moment condition examines the second moment of the in-sample estimation error, which equals the variance of a truncated normal distribution. The truncation arises from strategic disclosure, i.e., the sample selection criterion is $\varepsilon \leq \iota$. The in-sample estimation error $\hat{x} - \hat{s}$ equals $-\varepsilon + \mathcal{B}$ where $\mathcal{B} \equiv \mathbb{E}[\varepsilon|\varepsilon \leq \iota]$. The in-sample fit is unbiased. Therefore, the second moment of the in-sample error should be equivalent to the variance of the in-sample error:

$$\mathbb{E}\varepsilon_d^2 = \sigma^2\left(1 - \frac{\tilde{\iota}\phi(\tilde{\iota})}{\Phi(\tilde{\iota})} - \left(\frac{\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}\right)^2\right) \tag{A.4}$$

where $\mathbb{E}\varepsilon_d^2$ is the shorthand for $\mathbb{E}\left[(\hat{x}-\hat{s})^2|d=\hat{s}\right]$ and ϕ and Φ are standard normal p.d.f. and c.d.f., respectively. The empirical counterpart of $\mathbb{E}\varepsilon_d^2$ is given by $\bar{\varepsilon_d^2} \equiv \frac{1}{N\bar{d}}\sum_{d_i\neq\varnothing}(\hat{x}_i - \hat{s}_i)^2$. Together with the estimated threshold $\hat{\tilde{\iota}}$, one can compute the estimated variance of the specification error $\hat{\sigma^2}$ by dividing the empirical variance of the in-sample estimation error $\bar{\varepsilon_d^2}$ by the variance of the truncated standard normal parameterized using $\hat{\tilde{\iota}}$, i.e., $1-\frac{\hat{\tilde{\iota}}\phi(\hat{\tilde{\iota}})}{\Phi(\hat{\tilde{\iota}})}-\left(\frac{\phi(\hat{\tilde{\iota}})}{\Phi(\hat{\tilde{\iota}})}\right)^2$.

Collectively, for any $\bar{d} \in (0,1)$ and $\bar{\varepsilon_d^2} > 0$, the method of moment estimators $(\hat{\lambda}, \hat{\sigma})$ of

[42] The computations of ϕ, Φ, and its inverse Φ^{-1} are available in most statistical computing software distributions.

parameters (λ, σ) are defined as the unique solution to the following system of equations.

$$\begin{aligned} \bar{d} &= \Phi(i(\hat{\lambda})) \\ \bar{\varepsilon_d^2} &= \hat{\sigma^2}\left(1 - \frac{\hat{\tilde{\iota}}\phi(\hat{\tilde{\iota}})}{\Phi(\hat{\tilde{\iota}})} - \left(\frac{\phi(\hat{\tilde{\iota}})}{\Phi(\hat{\tilde{\iota}})}\right)^2\right) \end{aligned} \tag{A.5}$$

Unfortunately, the last parameter of the model, κ, i.e., the relative precision, is not identified because firms' true underlying emissions $\{e_i\}$ are not available. Thus, it is infeasible for the econometrician to disentangle σ_x^2 and σ_s^2. From a modeling perspective, κ can be interpreted as investors' subjective belief of the relative precision.

C Information Loss

The information loss, denoted by $\mathcal{L}$, is defined as the second moment of errors in investors' posterior beliefs—$\mathbb{E}[e|\mathcal{I}]$ for rational investors and $\mathbb{E}^\nu[e|\mathcal{I}]$ for naive investors—about firms' emissions.[43] The information loss, conditional on disclosure or non-disclosure, is given by

$$\mathcal{L}^d \equiv \begin{cases} \mathbb{E}\left[\left(\mathbb{E}^\nu[e|\hat{x},\hat{s}] - e\right)^2\right] & d = \hat{s},\ \text{naive} \\ \mathbb{E}\left[\left(\mathbb{E}[e|\hat{x},\hat{s}] - e\right)^2\right] & d = \hat{s},\ \text{rational} \end{cases} \tag{A.6a}$$

$$\mathcal{L}^{nd} \equiv \begin{cases} \mathbb{E}\left[\left(\mathbb{E}^\nu[e|\hat{x},\varnothing] - e\right)^2\right] & d = \varnothing,\ \text{naive} \\ \mathbb{E}\left[\left(\mathbb{E}[e|\hat{x},\varnothing] - e\right)^2\right] & d = \varnothing,\ \text{rational} \end{cases} \tag{A.6b}$$

Table A.1 presents the closed-form solution of the information loss defined in Equation A.6. The proof of the results is provided in Online Appendix IV. The ex ante information loss is further defined as the weighted average of information loss given disclosure propensity: $\mathcal{L} \equiv \mathbb{E}d\mathcal{L}^d + (1-\mathbb{E}d)\mathcal{L}^{nd}$ where $\mathbb{E}d$ is the expected disclosure propensity. An important remark regarding investors' information loss is that the selection bias does affect naive investors' loss even if firms disclose, even though intermediaries' estimates are unbiased when benchmarked against firms' disclosures.[44]

[43] Despite the risk neutrality assumption in my analytical framework, I use the second moments to capture their information losses for the following reason. When making capital allocation decisions, investors care about each individual firm's expected emissions. However, for a portfolio of firms, both over-allocation and under-allocation cause inefficiencies.

[44] The intuition is that the benchmark that matters with respect to evaluating information loss is the true underlying emissions. As shown in Online Appendix IV, the error in rational investors' posterior beliefs equals $-\frac{\kappa}{\kappa+1}\varepsilon^x + \frac{1}{\kappa+1}\varepsilon^s$. This error is orthogonal to the specification error ε which determines firms' disclosure decisions. Therefore, in expectation, $\mathbb{E}[\mathbb{E}[e|\hat{x},\hat{s}] - e] = 0$ and $\mathbb{E}[\mathbb{E}^\nu[e|\hat{x},\hat{s}] - e] = \frac{\kappa}{\kappa+1}\mathcal{B}$.

Table A.1: Information loss

This table presents the analytical solution of the information loss defined in Equation A.6. Panel A analyzes the case where intermediaries do not correct for the bias and report $\hat{x}$. Panel B analyzes the case where intermediaries correct for the bias and report x. Importantly, all the losses are scaled by $1/\sigma^2$ to mute the variance effect. Panel A also generally applies to cases where intermediaries refine their estimates (Corollary A.1) and disclosure mandates are implemented (Corollary A.2), except that the standardized disclosure threshold and/or the standardized bias needs to be updated accordingly. The loss is the same in the case of disclosure mandates regardless of whether disclosure is voluntary or required, because the error in rational investors' beliefs conditional on disclosure, i.e., $-\frac{\kappa}{\kappa+1}\varepsilon^x+\frac{1}{\kappa+1}\varepsilon^s$, is orthogonal to the specification error ε that determines disclosure.

Panel A: Intermediary reports $\hat{x}$

	Disclosure ($d=\hat{s}$)	Non-disclosure ($d=\varnothing$)
Naive ($\mathbb{E}^N$)	$\left(\frac{\kappa}{\kappa+1}\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}\right)^2+\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$	$\left(\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}-\frac{1}{\kappa+1}\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}\right)^2+\left(1+\frac{\tilde{\iota}\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}-\left(\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}\right)^2\right)\frac{1}{(\kappa+1)^2}+\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$
Rational ($\mathbb{E}$)	$\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$	$\left(1+\frac{\tilde{\iota}\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}-\left(\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}\right)^2\right)\frac{1}{(\kappa+1)^2}+\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$

Panel B: Intermediary reports x

	Disclosure ($d=\hat{s}$)	Non-disclosure ($d=\varnothing$)
Naive ($\mathbb{E}^N$)	$\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$	$\left(-\frac{1}{\kappa+1}\frac{\phi(\iota^{\tilde{C}})}{1-\Phi(\iota^{\tilde{C}})}\right)^2+\left(1+\frac{\tilde{\iota}\phi(\iota^{\tilde{C}})}{1-\Phi(\iota^{\tilde{C}})}-\left(\frac{\phi(\iota^{\tilde{C}})}{1-\Phi(\iota^{\tilde{C}})}\right)^2\right)\frac{1}{(\kappa+1)^2}+\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$
Rational ($\mathbb{E}$)	$\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$	$\left(1+\frac{\tilde{\iota}\phi(\iota^{\tilde{C}})}{1-\Phi(\iota^{\tilde{C}})}-\left(\frac{\phi(\iota^{\tilde{C}})}{1-\Phi(\iota^{\tilde{C}})}\right)^2\right)\frac{1}{(\kappa+1)^2}+\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$

D Supplementary Corollaries

In this appendix, I outline additional corollaries that provide the analytical foundation for my structural estimation and counterfactual analysis.

The following corollary characterizes the equilibrium disclosure strategy if intermediaries de-bias their estimates.

Corollary A.1. *If intermediaries adjust for the selection bias of their estimates, the equilibrium disclosure strategy $\iota^C \in \mathbb{R}$ solves*

$$\iota^C = (1-\lambda)\mathbb{E}\left[\varepsilon|\varepsilon \geq \iota^C\right] \tag{A.7}$$

The standardized disclosure threshold $\tilde{\iota}^C \equiv \frac{\iota^C}{\sigma}$ solves $\tilde{\iota}^C = (1-\lambda)\frac{\phi(\tilde{\iota}^C)}{1-\Phi(\tilde{\iota}^C)}$, with $\tilde{\iota}^C = \tilde{\iota}^N > \tilde{\iota}$, where $\tilde{\iota}$ is the standardized disclosure threshold if intermediaries do not correct for the selection bias and $\tilde{\iota}^N$ is the standardized disclosure threshold if there is no intermediary, both characterized in Corollary 1.

The following corollary characterizes the equilibrium disclosure strategy and intermediaries' bias if $\theta \in [0,1]$ firms are required to disclose. For tractability, let $d_\theta \in \{\hat{s}, \varnothing\}$ denote firms' voluntary disclosure decisions, regardless of whether they are subject to the mandate, and let d'_θ denote the eventual disclosures observed by the market. The propensity of non-disclosure, $1-\mathbb{E}d'_\theta$, equals the product of the propensity of not being drafted for mandatory disclosure, $1-\theta$, and the propensity of not voluntarily disclosing, $1-\mathbb{E}d_\theta$.[45]

Corollary A.2. *If $\theta \in [0,1]$ firms are randomly assigned to disclose their emissions, the equilibrium disclosure threshold $\iota_\theta \in \mathbb{R}$ solves*

$$\iota_\theta = (1-\lambda)\mathbb{E}[\varepsilon|\varepsilon \geq \iota_\theta] + \lambda\mathcal{B}_\theta \tag{A.8}$$

where the selection bias $\mathcal{B}_\theta \equiv \mathbb{E}[\varepsilon|d'_\theta = \hat{s}]$ is given by $\frac{(1-\theta)\Phi(\tilde{\iota_\theta})}{\theta+(1-\theta)\Phi(\tilde{\iota_\theta})}\mathbb{E}[\varepsilon|\varepsilon \leq \iota_\theta]$. Thus, the standardized disclosure threshold $\tilde{\iota_\theta} \equiv \frac{\iota_\theta}{\sigma}$ solves $\tilde{\iota_\theta} = (1-\lambda)\frac{\phi(\tilde{\iota_\theta})}{1-\Phi(\tilde{\iota_\theta})} + \lambda\frac{(1-\theta)\Phi(\tilde{\iota_\theta})}{\theta+(1-\theta)\Phi(\tilde{\iota_\theta})}\frac{-\phi(\tilde{\iota_\theta})}{\Phi(\tilde{\iota_\theta})}$. Moreover, when $\theta = 0$, $\tilde{\iota_\theta} = \tilde{\iota}$; when $\theta = 1$, $\tilde{\iota_\theta} = \tilde{\iota}^C$; for interim $\theta \in (0,1)$, $\tilde{\iota} < \tilde{\iota_\theta} < \tilde{\iota}^C$. $\tilde{\iota}$ is the standardized disclosure threshold if intermediaries do not correct for the selection bias characterized in Corollary 1, $\tilde{\iota}^C$ is the standardized disclosure threshold if intermediaries de-bias their estimates as characterized in Corollary A.1.

[45] Given the random assignment of mandatory disclosure, firms will not disclose if and only if they are not subject to the mandate and they do not gain from voluntary disclosure, with these two conditions occurring independently of each other.

E Proofs of Main Results

E.1 Lemma 1

Proof. Firms disclose if and only if $\mathbb{E}_\lambda[e|d=\hat{s}] \leq \mathbb{E}_\lambda[e|d=\varnothing]$ with

$$\mathbb{E}_\lambda[e|d=\hat{s}] = \frac{\tau_e}{\tau_e+\tau_s}\mu + \frac{\tau_s}{\tau_e+\tau_s}\hat{s}$$
$$\mathbb{E}_\lambda[e|d=\varnothing] = \lambda\mu + (1-\lambda)\mathbb{E}\left\{\frac{\tau_e}{\tau_e+\tau_s}\mu + \frac{\tau_s}{\tau_e+\tau_s}\hat{s}\,\middle|\, d=\varnothing\right\}$$

where $\tau_e \equiv \sigma_e^{-2}$ is the precision of the of the common prior. Rearranging terms yields the following criterion for disclosure

$$\hat{s}-\mu \leq (1-\lambda)\mathbb{E}\{\hat{s}-\mu|d=\varnothing\}$$

where the right-hand side does not depend on $\hat{s}$. Therefore, there exists a threshold equilibrium in which firms disclose if and only if $\hat{s}-\mu \leq \iota^N$ where $\iota^N \in \mathbb{R}$ solves

$$\iota^N = (1-\lambda)\mathbb{E}\{\hat{s}-\mu|\hat{s}-\mu \geq \iota^N\}$$

The uniqueness of ι^N is later established in Corollary 1.

Furthermore, the proof for Lemma A.1 follows the same logic. The uniqueness of i^N that solves Equation A.1 is established in Online Appendix III.

□

E.2 Proposition 1

Proof. Recall that when the number of firms in the economy is large, i.e., $N \to \infty$, we have $\hat{x} = x + \mathbb{E}[\varepsilon|d \neq \varnothing]$ where the second term is the estimation bias. Rearranging terms in Equation 4 yields:

$$\mathbb{E}_\lambda[e|\hat{x}, d=\hat{s}] = x + \lambda\frac{\kappa}{\kappa+1}\mathbb{E}[\varepsilon|d\neq\varnothing] + \frac{1}{\kappa+1}\varepsilon$$
$$\mathbb{E}_\lambda[e|\hat{x}, d=\varnothing] = x + \lambda\mathbb{E}[\varepsilon|d\neq\varnothing] + (1-\lambda)\frac{1}{\kappa+1}\mathbb{E}[\varepsilon|d=\varnothing]$$

Firms disclose if and only if $\mathbb{E}_\lambda[e|\hat{x}, d=\hat{s}] \leq \mathbb{E}_\lambda[e|\hat{x}, d=\varnothing]$ and hence

$$\varepsilon \leq (1-\lambda)\mathbb{E}[\varepsilon|d=\varnothing] + \lambda\mathbb{E}[\varepsilon|d\neq\varnothing]$$

where the right-hand side does not depend on ε. Therefore, there exists an equilibrium characterized by a threshold structure w.r.t. ε such that firms disclose if and only if $\varepsilon \leq \iota$ where $\iota \in \mathbb{R}$ solves the implicit function

$$\iota = (1-\lambda)\mathbb{E}\left[\varepsilon|\varepsilon \geq \iota\right] + \lambda\mathbb{E}[\varepsilon|\varepsilon \leq \iota]$$

The uniqueness of ι is later established in Corollary 1.

Furthermore, the proof for Proposition A.1 follows the same logic. The uniqueness of i that solves Equation A.2 is established in Online Appendix III.

□

E.3 Corollary 1

Proof. Given normality of $\varepsilon \equiv \varepsilon^x + \varepsilon^s \sim \mathcal{N}(0, \sigma^2)$, Equation 5 in Proposition 1 can be rewritten as Equation 6a:

$$\tilde{\iota} = (1-\lambda)\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})} + \lambda\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}$$

where $\tilde{\iota} \equiv \frac{\iota}{\sigma}$. Similarly, given normality of $\varepsilon^N \equiv \varepsilon^e + \varepsilon^s \sim \mathcal{N}(0, \sigma_N^2)$, Equation 2 in Lemma 1 can be rewritten as Equation 6b:

$$\tilde{\iota}^N = (1-\lambda)\frac{\phi(\tilde{\iota}^N)}{1-\Phi(\tilde{\iota}^N)}$$

where $\tilde{\iota}^N \equiv \frac{\iota^N}{\sigma_N}$.[46]

First, I prove the uniqueness of $\tilde{\iota}$ that solves Equation 6a and prove that $\frac{\partial\tilde{\iota}}{\partial\lambda} < 0$ for $\lambda \in (0,1)$. The same properties of $\tilde{\iota}^N$ trivially follow. Let $I(z,\lambda) \equiv (1-\lambda)\frac{\phi(z)}{1-\Phi(z)} + \lambda\frac{-\phi(z)}{\Phi(z)} - z$. For any $\lambda \in (0,1)$, we have $\lim_{z\to-\infty} I(z,\lambda) > 0$ and $\lim_{z\to+\infty} I(z,\lambda) < 0$ since $\lim_{z\to-\infty}\frac{\phi(z)}{1-\Phi(z)} = 0$, $\lim_{z\to-\infty}\frac{-\phi(z)}{\Phi(z)} = -\infty$ and $\lim_{z\to+\infty}\frac{\phi(z)}{1-\Phi(z)} = +\infty$, $\lim_{z\to+\infty}\frac{-\phi(z)}{\Phi(z)} = 0$. Therefore, there always exists some $\tilde{\iota} \in \mathbb{R}$ such that $I(\tilde{\iota},\lambda) = 0$. In addition, given the properties of truncated normal distributions, both $l(z) = \frac{-\phi(z)}{\Phi(z)} - z$ and $u(z) = \frac{\phi(z)}{1-\Phi(z)} - z$ are strictly decreasing in z. Hence, $\frac{\partial I}{\partial\tilde{\iota}} < 0$ for $\lambda \in (0,1)$, which means that there is a unique $\tilde{\iota} \in \mathbb{R}$ that solves the implicit function.

To prove that $\tilde{\iota} < \tilde{\iota}^N$, suppose $\tilde{\iota} \geq \tilde{\iota}^N$. Hence, $(1-\lambda)\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})} + \lambda\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})} - (1-\lambda)\frac{\phi(\tilde{\iota}^N)}{1-\Phi(\tilde{\iota}^N)} \geq 0$. However, $\tilde{\iota} \geq \tilde{\iota}^N$ also implies that $\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})} - \frac{\phi(\tilde{\iota}^N)}{1-\Phi(\tilde{\iota}^N)} \leq 0$. Meanwhile, $\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})} < 0$ for any $\tilde{\iota} \in \mathbb{R}$. Collectively, $\tilde{\iota} \geq \tilde{\iota}^N$ must lead to $(1-\lambda)\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})} + \lambda\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})} - (1-\lambda)\frac{\phi(\tilde{\iota}^N)}{1-\Phi(\tilde{\iota}^N)} < 0$, that is, a contradiction. Therefore, $\tilde{\iota} < \tilde{\iota}^N$. Lastly, as Φ is monotonically increasing, we have

[46] Specifically, $\sigma^2 = \sigma_x^2 + \sigma_s^2$ and $\sigma_N^2 = \sigma_e^2 + \sigma_s^2$. Since $\sigma_e^2 > \sigma_x^2$, $\sigma^2 > \sigma_N^2$.

$\Phi(\tilde{\iota}) < \Phi(\tilde{\iota}^N)$.

Moreover, for any $\lambda \in (0,1)$ and $\tilde{\iota} \in \mathbb{R}$

$$\frac{\partial I}{\partial \lambda} = -\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})} - \frac{\phi(\tilde{\iota})}{\Phi(\tilde{\iota})} < 0$$

Therefore, for any $\lambda_1 < \lambda_2$, $\lambda_1, \lambda_2 \in (0,1)$, suppose that $\tilde{\iota}_1 \in \mathbb{R}$ uniquely solves $I(\tilde{\iota}_1, \lambda_1) = 0$ and $\tilde{\iota}_2 \in \mathbb{R}$ uniquely solves $I(\tilde{\iota}_2, \lambda_2) = 0$. It must be the case that $\tilde{\iota}_1 > \tilde{\iota}_2$, because if otherwise, i.e., $\tilde{\iota}_1 \leq \tilde{\iota}_2$

$$0 = I(\tilde{\iota}_1, \lambda_1) \geq I(\tilde{\iota}_2, \lambda_1) > I(\tilde{\iota}_2, \lambda_2) = 0$$

that is, a contradiction. Hence, $\tilde{\iota}$ is monotonically decreasing in $\lambda \in (0,1)$.

□

Figure 1: Timeline

This figure explains the timeline of the voluntary disclosure game analyzed in Section 2. $\hat{s}$ is firms' emissions measurement. $\hat{x} = \hat{f}(z)$ is the intermediary's estimates calculated using the fitted $\hat{f}$ and observed firm characteristics z. The game repeats every period but is in essence a stage game due to the i.i.d. assumptions.

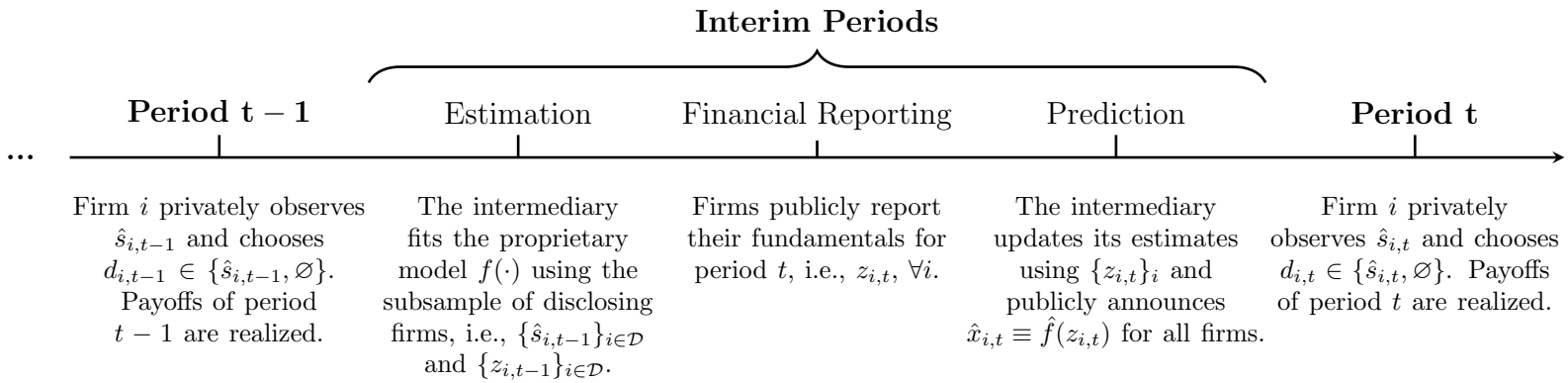


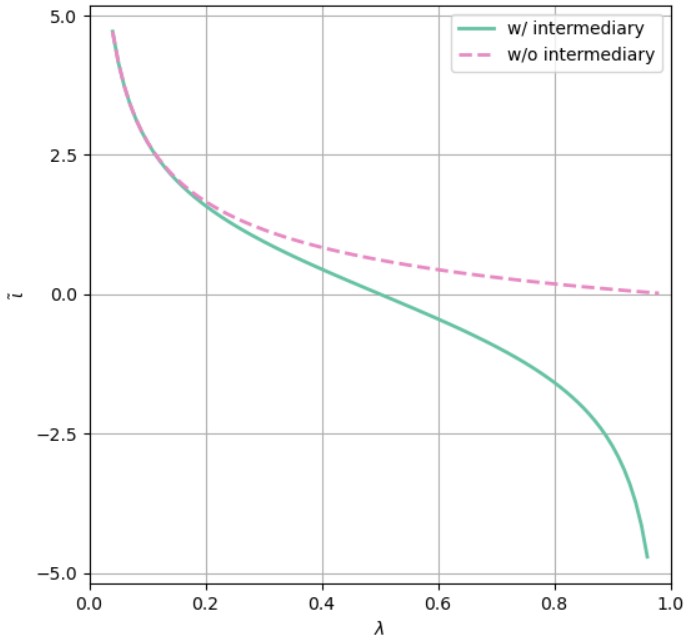
Figure 2: Equilibrium standardized disclosure threshold, $\tilde{\iota}$, and average disclosure propensity, $\mathbb{E}d$

This figure plots the numerically solved equilibrium disclosure thresholds (standardized), $\tilde{\iota}$, and disclosure mass, $\mathbb{E}d$, following the implicit functions outlined in Equation 6:

$$\begin{cases} \tilde{\iota} = (1-\lambda)\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})} + \lambda\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})} & \text{with the intermediary} \\ \tilde{\iota}^N = (1-\lambda)\frac{\phi(\tilde{\iota}^N)}{1-\Phi(\tilde{\iota}^N)} & \text{without the intermediary} \end{cases}$$

where $\lambda \in (0,1)$ is the share of naive investors. $\phi(\cdot)$ and $\Phi(\cdot)$ are standard normal p.d.f. and c.d.f., respectively.

Panel A: The standardized threshold, $\tilde{\iota}$



Panel B: The average disclosure propensity, $\mathbb{E}d$

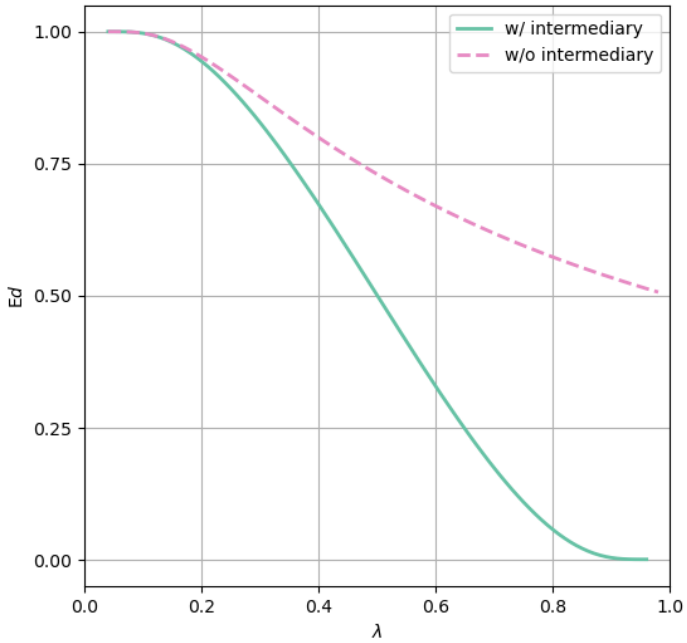
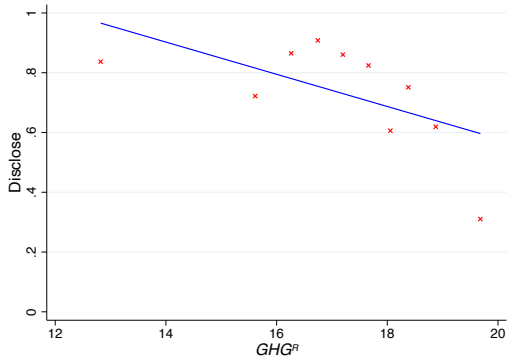


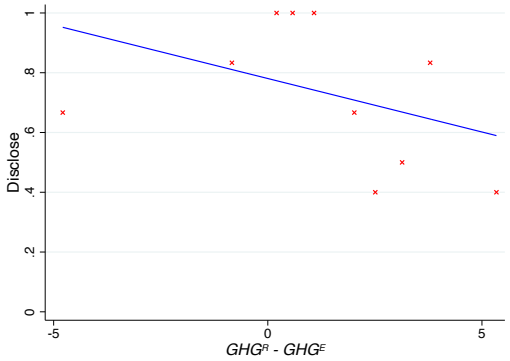
Figure 3: Selective disclosure of scope 1 emissions

This figure shows the relationship between the likelihood of voluntary disclosure (*Disclose*) and "ground truth" emissions proxied by aggregated facility-level emissions (GHG^R). Three aggregation methods are considered: full summation (*full*), proportional summation (*prop.*), and pseudo consolidation (*cons.*). The subplots are `binscatter` plots that display ten average values of disclosure-emissions combinations and insert a linear fit. Panels on the left control for scale of operations using firm fundamentals. Panels on the right subtract estimated emissions (GHG^E) from aggregated emissions (GHG^R). All variables are arcsin-transformed and are winsorized at the 5% and 95% levels to accommodate the smaller sample size.

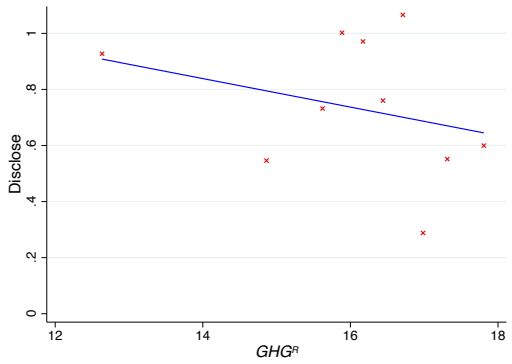
Panel A: *full*, w/ controls



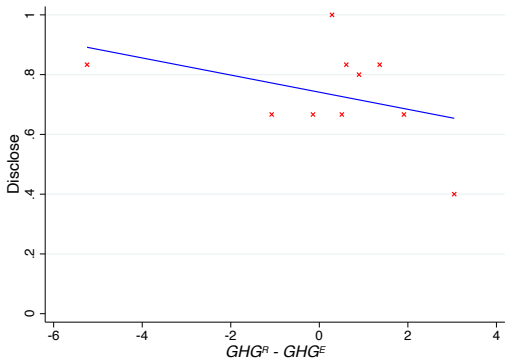
Panel B: *full*, error



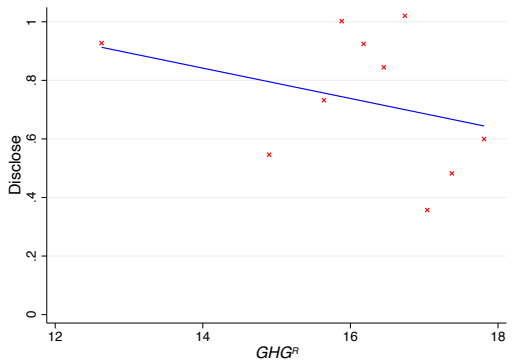
Panel C: *prop.*, w/ controls



Panel D: *prop.*, error



Panel E: *cons.*, w/ controls



Panel F: *cons.*, error

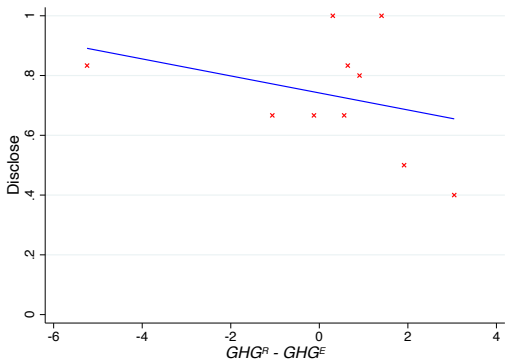


Figure 4: The distribution of estimated vs. disclosed emissions, Bloomberg

This figure is excerpted from Figure 2 in Bloomberg (2021). The figure compares the distributions of the estimated emissions, pre- and post-recalibration, versus self-reported (empirical) emissions. The orange line represents the calibrated distribution of scope 1 emissions estimates. The blue line represents the recalibrated distribution of scope 1 emissions estimates that Bloomberg uses for its ultimate estimation outputs.

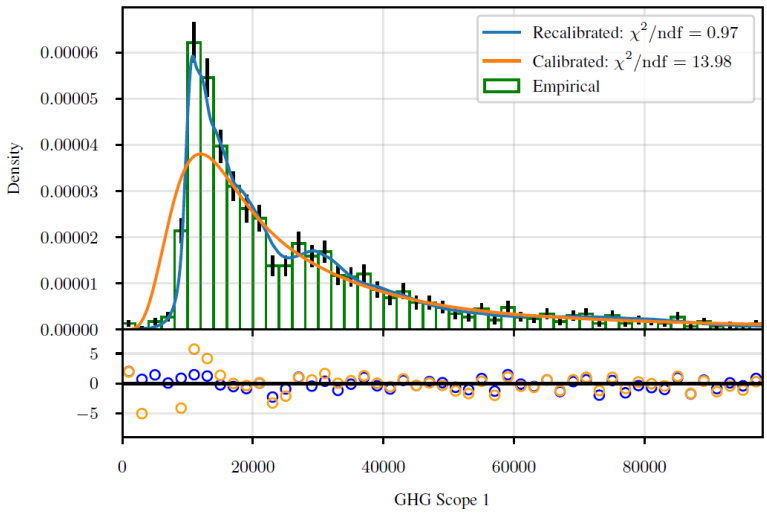


Figure 2: Comparison of the predicted mixture distribution for Reporting Companies GHG Scope 1 Model before and after recalibration versus observations. The lower panel shows the distance between the predicted density and the observed density, normalized by the standard error of the counts in the bin.

Figure 5: The estimation bias for always- vs. newly disclosing U.K. firms

This figure compares the estimation errors computed as the arcsin-transformed estimated emissions (GHG^E) minus the arcsin-transformed disclosed emissions (GHG^A) for always-disclosing firms versus newly disclosing firms after the U.K.'s 2019 "comply-or-explain" emissions reporting regulation. Firms' transformed emissions are winsorized at the 5% and 95% levels to accommodate the smaller sample size.

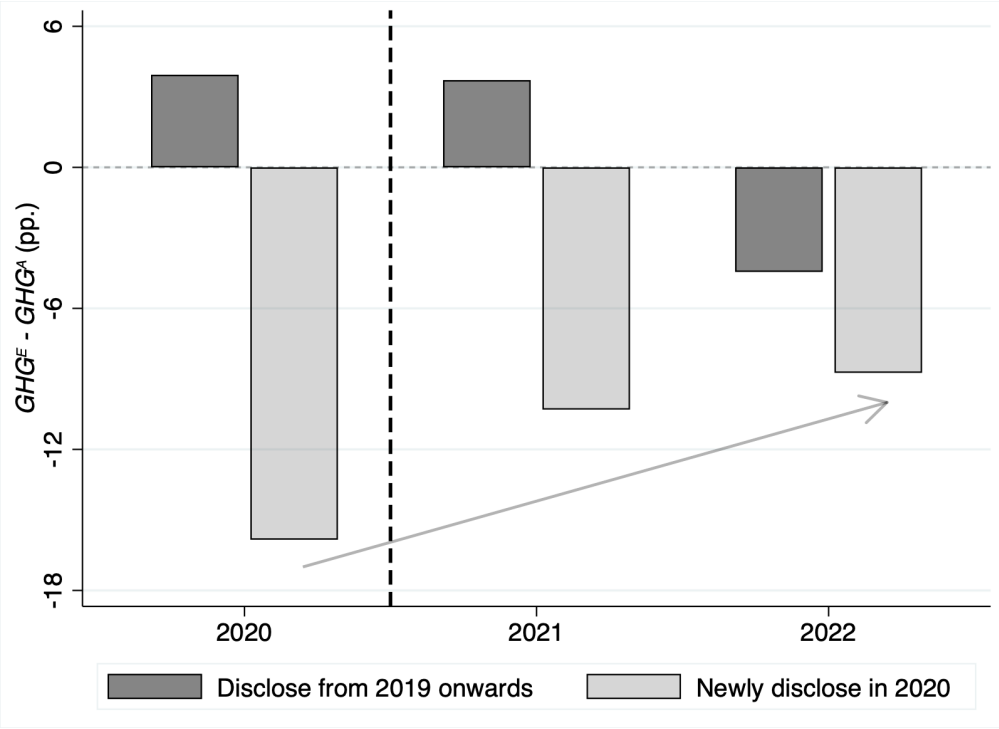


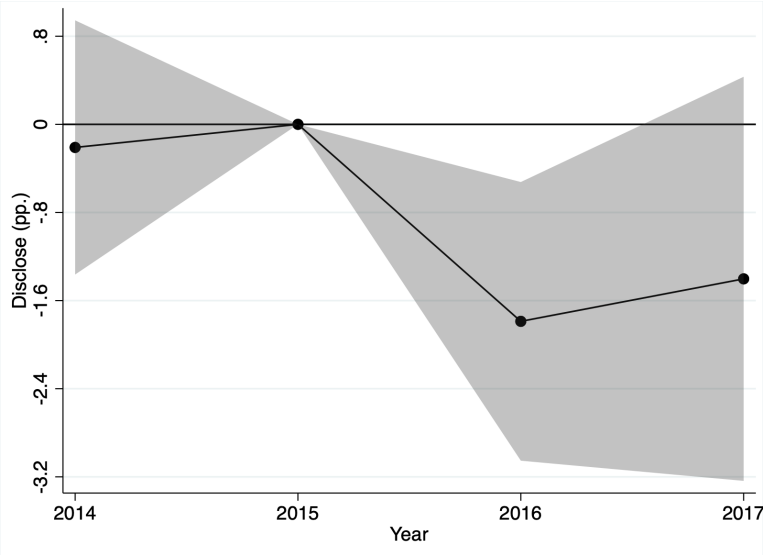
Figure 6: The disclosure deterrence effect of the 2016 Trucost expansion

This figure plots the dynamic deterrence effect of the 2016 Trucost coverage expansion over the sample period 2014 to 2017 by estimating the following regression:

$$Disclose_{i,t} = \alpha_i + \alpha_{j,t} + \sum_{\tau \neq 2016} \beta_\tau Treated_{c(i)} \times \mathbb{1}_{t=\tau} + \epsilon_{i,t}$$

where $\mathbb{1}_{t=\tau}$ equals one if year t for the observation is precisely year τ. The year 2016 is considered the baseline year and omitted from the regression. The indicator variable *Disclose* is scaled up by 100, i.e., in the percentage-point (pp.) term. Panel A uses the country-level treatment *Expansion*, while Panel B uses the firm-level treatment *NewCoverage*. The error bars indicate 90% confidence intervals based on standard errors clustered at the country level (Panel A) or at the firm level (Panel B).

Panel A: Dynamic treatment effect, *Expansion*



Panel B: Dynamic treatment effect, *NewCoverage*

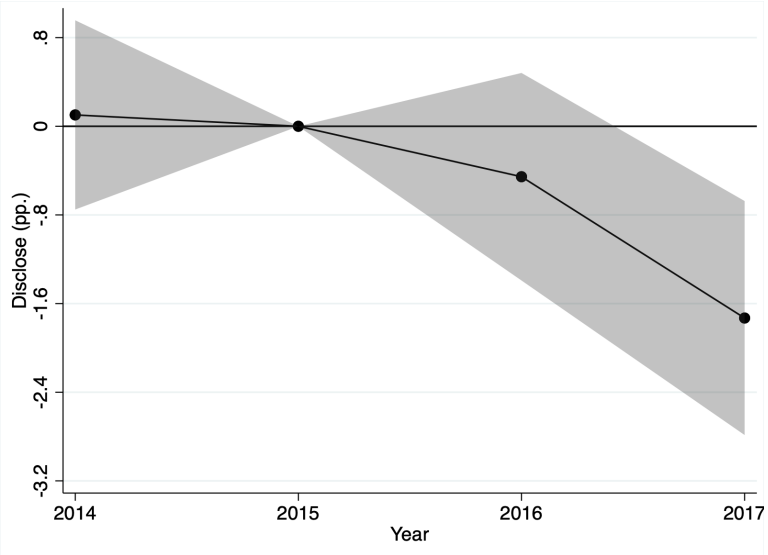
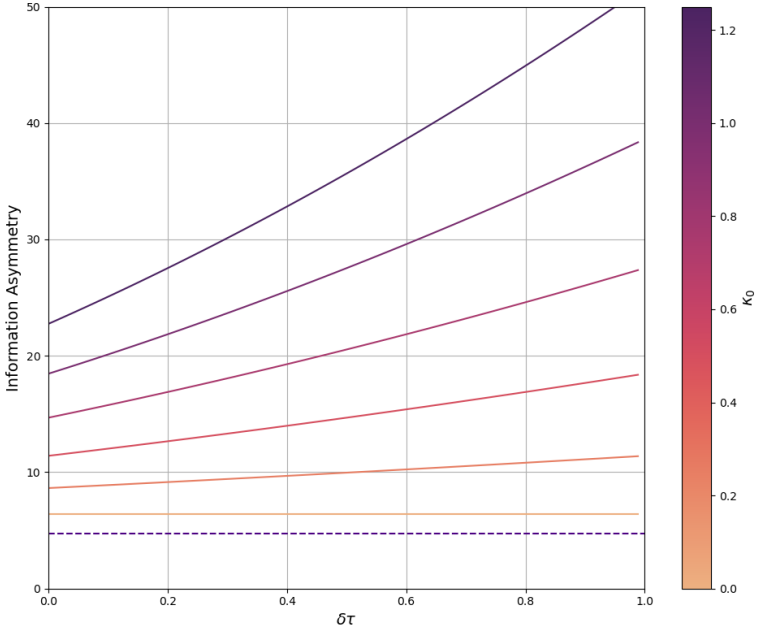


Figure 7: Information asymmetry between naive and rational investors

This figure plots the consequential information asymmetry between naive and rational investors when intermediaries refine their estimates (Panel A) or disclosure mandates are introduced (Panel B). The information asymmetry is computed as the ratio of naive investors' and rational investors' information loss. $\delta\tau$ is the percentage increase in the precision of intermediaries' emissions modeling. θ is the fraction of firms subject to mandatory disclosure rules. Deeper line colors correspond to higher values of κ_0, the initial relative precision of intermediaries' emissions modeling compared to firms' emissions measurements. The dashed line represents the information asymmetry when intermediaries correct for the selection bias. In computing investors' information loss, I subtract the optimal information loss in a frictionless scenario, i.e., $\mathcal{L}_0 \equiv \frac{1}{\left(\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}\right)^2}$.

Panel A: Refined emissions models, $\delta\tau$



Panel B: Disclosure mandates, θ

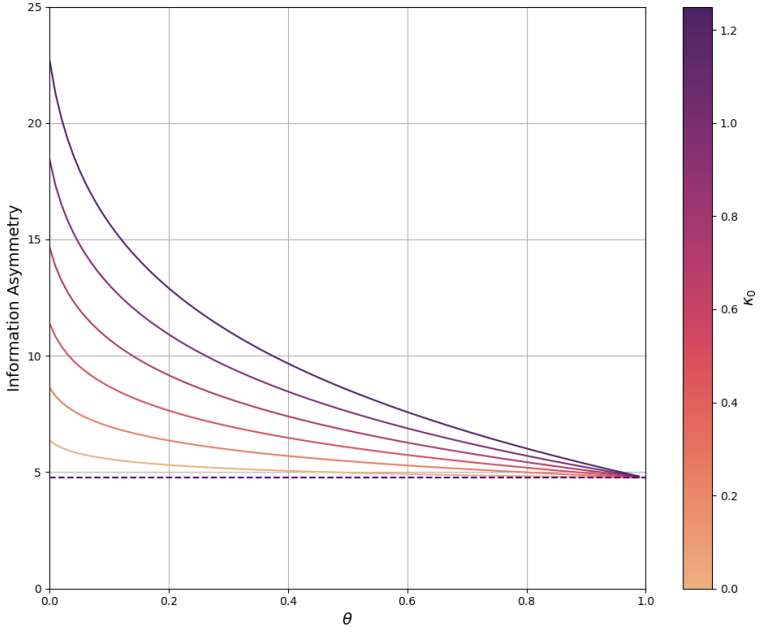


Table 1: Descriptive Statistics

This table reports descriptive statistics for various samples used in the empirical analysis. The Bloomberg sample spans from 2020 to 2022. The Trucost sample spans from 2014 to 2017. All emissions variables and firm-level characteristics are transformed using the inverse hyperbolic sine function. Estimation errors are calculated by subtracting transformed disclosed emissions from transformed estimated emissions, i.e., $GHG^E - GHG^A$. Firm fundamentals data are obtained from Compustat and CRSP. All variables in the Bloomberg sample are winsorized at the 1% and 99% levels.

Variables	N	Mean	Std.	10th	25th	50th	75th	90th
Bloomberg U.S. sample, 2020-22 monthly								
GHG^E	97,186	9.00	3.71	4.16	7.43	9.37	11.07	13.35
GHG^A	27,037	11.35	3.44	7.09	9.11	11.41	13.71	15.90
$GHG^E - GHG^A$	27,037	0.05	0.62	-0.17	-0.05	-0.00	0.04	0.20
ME	97,186	7.57	2.33	4.44	5.89	7.63	9.17	10.68
$Assets$	97,186	7.53	2.21	4.55	5.97	7.53	9.02	10.43
$Sales$	97,186	6.41	3.03	1.16	4.67	6.97	8.55	9.92
OP	97,186	3.19	4.74	-3.94	-2.19	4.88	6.83	8.30
PPE	97,186	5.82	2.82	2.00	3.80	5.97	7.89	9.46
$Capex$	97,186	3.45	2.46	0.13	1.24	3.42	5.24	6.83
$Debt$	97,186	5.64	3.02	1.02	3.39	6.02	7.97	9.42
Trucost global sample, 2014-17 annual								
$Disclose$	49,690	9.71	29.61	-	-	-	-	-
$NewCoverage$	49,690	0.51	0.50	-	-	-	-	-
$Expansion$	49,690	1.68	0.69	0.54	1.49	1.78	2.27	2.34

Table 2: Emissions Pricing

This table examines whether emissions are priced. Panel A reports the results for estimating Equation 7 and its variations. Columns (1) - (2) focus on all firms; columns (3) - (5) focus on disclosing firms only. Control variables include total assets, sales, operating profits, PP&E, capital expenditure, and debt. All variables are transformed using the inverse hyperbolic sine function. Panel B further reports the results examining whether emissions pricing is more sensitive for firms with higher climate exposure à la Sautner et al. (2023). All continuous variables are winsorized at the 1% and 99% levels. Standard errors are clustered at the firm level and are displayed in parentheses. *, **, and *** indicate statistical significance of the mean differences at the 10%, 5%, and 1% levels, respectively.

Panel A: Emissions pricing, disclosed vs. estimated

	(1)	(2)	(3)	(4)	(5)
	Dep. Var. = ME				
	All Firms		*Disclosing Firms*		
GHG^E	-0.097***	-0.088***	-0.068***		-0.068***
	(0.009)	(0.012)	(0.015)		(0.015)
Disclose		0.134			
		(0.120)			
$GHG^E \times Disclose$		0.004			
		(0.010)			
GHG^A				-0.055***	
				(0.013)	
$GHG^E - GHG^A$					-0.006
					(0.032)
N	97,186	97,186	27,037	27,037	27,037
Industry FE	x	x	x	x	x
Year-Month FE	x	x	x	x	x
Control	x	x	x	x	x
Cluster	Firm	Firm	Firm	Firm	Firm
Adj. R^2	0.833	0.834	0.810	0.809	0.810

Table 2: Emissions Pricing

Panel B: Emissions pricing, cross-sectional results

	(1)	(2)	(3)	(4)
		Dep. Var. $= ME$		
	$Risk$	$Risk^{OP}$	$Risk^{RG}$	$Risk^{PH}$
GHG	-0.092***	-0.093***	-0.093***	-0.094***
	(0.010)	(0.010)	(0.010)	(0.010)
$Relevance$	4.135***	4.334**	9.876***	-1.580
	(1.332)	(2.193)	(3.595)	(8.956)
$GHG \times Relevance$	-0.303***	-0.339**	-0.636**	-0.074
	(0.098)	(0.165)	(0.257)	(0.653)
N	52,729	52,729	52,729	52,729
Industry FE	x	x	x	x
Year-Month FE	x	x	x	x
Control	x	x	x	x
Cluster	Firm	Firm	Firm	Firm
Adj. R^2	0.829	0.829	0.829	0.829

Table 3: Selective disclosure of scope 1 emissions

This table reports the results of estimating the Probit model variants of the linear probability models outlined in Equation 8. Columns (1) - (3) estimate the Probit model with aggregated Climate TRACE emissions (GHG^R) and an array of controls to control for firms' scale of operations. Columns (4) - (6) control for expected emissions by subtracting Bloomberg's emissions estimates GHG^E from the aggregated Climate TRACE emissions GHG^R. Three aggregation methods are considered: full summation (*full*), proportional summation (*prop.*), and pseudo consolidation (*cons.*). Robust standard errors are displayed in parentheses. *, **, and *** indicate statistical significance of the mean differences at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Var. = *Disclose*					
	full	*prop.*	*cons.*	*full*	*prop.*	*cons.*
GHG^R	-0.3910***	-0.3036*	-0.3119**			
	(0.150)	(0.157)	(0.159)			
$GHG^R - GHG^E$				-0.1514*	-0.1239	-0.1229
				(0.088)	(0.090)	(0.090)
Const.	-1.3671	-1.8759	-1.8836	0.8804***	0.6821***	0.6841***
	(3.503)	(3.173)	(3.159)	(0.251)	(0.188)	(0.189)
N	57	57	57	57	57	57
Controls	x	x	x			
Pseudo-R^2	0.395	0.352	0.354	0.0616	0.0270	0.0267

Table 4: Estimation errors for U.K. firms after the 2019 regulation

This table reports the results of examining whether the signed estimation error is lower for newly disclosing firms or for always-disclosing firms after the U.K. 2019 "comply-or-explain" emissions reporting regulation. Three reporting cycles after the regulation are considered: FY2020, FY2021, and FY2022. Newly disclosing firms did not disclose in FY2019 but started to disclose in FY2020; always-disclosing firms disclosed in both FY2019 and FY2020. Columns (1) - (3) focus on signed estimation errors of transformed emissions; Columns (4) - (6) focus on percentage signed error. Firms' transformed emissions are winsorized at the 5% and 95% levels to accommodate the smaller sample size. Standard errors are clustered at the firm level and are displayed in parentheses. *, **, and *** indicate statistical significance of the mean differences at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Var. $= GHG^E - GHG^A$			Dep. Var. $= GHG^E/GHG^A - 1$		
	FY2020	FY2021	FY2022	FY2020	FY2021	FY2022
NewlyDisclose	-0.1978*	-0.1104	-0.0524	-0.0216*	-0.0129	-0.0018
	(0.107)	(0.098)	(0.107)	(0.012)	(0.014)	(0.017)
Const.	0.0394*	0.0371	-0.0445	0.0086***	0.0137**	-0.0021
	(0.021)	(0.024)	(0.033)	(0.003)	(0.006)	(0.004)
N	465	457	429	465	457	429
Cluster	GHG^A tier	GHG^A tier	GHG^A tier	GHG^A tier	GHG^A tier	GHG^A tier
R^2	0.0210	0.0062	0.0007	0.0165	0.0019	0.0000

Table 5: The disclosure deterrence effect of the 2016 Trucost expansion

This table tabulates the disclosure deterrence effect of the 2016 Trucost coverage expansion over the sample period 2014 to 2017. *Treat* is defined as i) the increase in number of Compustat firms covered scaled by the pre-2016 coverage at the country level (*Expansion*); or ii) whether the firm is newly covered by Trucost starting in 2016 at the firm level (*NewCoverage*). The dependent variable *Disclose* equals one if firms disclose their scope 1 emissions directly to CDP or in their annual filings. The indicator variable *Disclose* is scaled up by 100. *, **, and *** indicate statistical significance of the mean differences at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
			Dep. Var. = *Disclose*			
		Treat = Expansion			*Treat = NewCoverage*	
Treat × *Post*	-1.147*	-1.445**	-1.234*	-1.495**	-1.068***	-1.131***
	(0.588)	(0.632)	(0.709)	(0.717)	(0.414)	(0.431)
N	49,657	49,657	48,303	48,303	48,303	48,303
Country-Industry FE	x	x				
Firm FE			x	x	x	x
Year FE	x		x		x	
Industry-Year FE		x		x		x
Cluster	Country	Country	Country	Country	Firm	Firm
Adj. R^2	0.217	0.215	0.685	0.686	0.685	0.685

Table 6: Baseline structural estimation results

This table presents structural estimation results using Bloomberg data on U.S. firms' emissions from 2020 to 2022. It first reports method of moments estimators characterized in Appendix B, then reports estimates of four economic constructs, i.e., the overall selection bias, the bias for non-disclosing firms, naive investors' information loss, and rational investors' information loss. The estimated overall bias $\hat{\mathcal{B}}$ and the estimated bias among non-disclosing firms $\hat{\mathcal{B}^{nd}}$ are given by

$$\hat{\mathcal{B}} = \hat{\sigma}\frac{-\phi(\hat{\iota})}{\Phi(\hat{\iota})}$$
$$\hat{\mathcal{B}^{nd}} = \hat{\sigma}\left(\frac{-\phi(\hat{\iota})}{\Phi(\hat{\iota})} - \frac{\phi(\hat{\iota})}{1-\Phi(\hat{\iota})}\right)$$

The information loss is calculated by applying $\kappa = 0$ and is normalized as in Table A.1. Bootstrapped standard errors are reported in parentheses.

λ	σ	$\mathcal{B}$	$\mathcal{B}^{nd}$	$\mathcal{L}$ naive	$\mathcal{L}$ rational
0.6540	0.4652	-0.6018	-0.7925	2.6163	0.4124
(0.0025)	(0.0174)	(0.0236)	(0.0300)	(0.0250)	(0.0046)

Table 7: Counterfactual analyses

This table presents the consequences of refinements in intermediaries' estimates and the roles of disclosure mandates. The quantification is conducted using estimators from Table 6. Panel A reports estimates of the disclosure mass, naive investors' information loss, and rational investors' information loss if intermediaries correct for the selection bias. The information loss is calculated by applying $\kappa = 0$ and is normalized as in Table A.1. Panel B reports estimates of the intermediaries' bias and the disclosure propensity in the presence of the θ mandate. Bootstrapped standard errors are reported in parentheses.

Panel A: Corrected estimates

$\mathbb{E}d$	$\mathcal{L}$ naive	$\mathcal{L}$ rational
0.6405	0.4940	0.1051
(0.0012)	(0.0033)	(0.0006)

Panel B: Disclosure mandates

	θ				
	10%	20%	30%	40%	50%
$\mathcal{B}_\theta$	-0.4117	-0.2952	-0.2164	-0.1596	-0.1167
	(0.0157)	(0.0112)	(0.0082)	(0.0060)	(0.0044)
$\mathbb{E}d'_\theta$	0.4486	0.5625	0.6480	0.7181	0.7783
	(0.0021)	(0.0016)	(0.0013)	(0.0010)	(0.0008)

Online Appendix

for Intermediaries and Emissions Disclosure

**The Online Appendix is not intended for publication.*

I Estimation under Self-selection

In this online appendix, I illustrate how self-selection in firms' disclosure decisions affects intermediaries' estimation. Consider the Heckman (1979) linear two-equation setup:

$$\begin{aligned} \hat{s} &= f(\boldsymbol{z}) + \varepsilon = \alpha + \boldsymbol{\beta}\boldsymbol{z} + \varepsilon \\ u &= g(\boldsymbol{v}) + \eta = \gamma + \boldsymbol{\delta}\boldsymbol{v} + \eta \end{aligned} \tag{OA.1}$$

where error terms are jointly normal: $(\varepsilon, \eta)^\top \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$; $\boldsymbol{z} \perp \varepsilon, \eta$ is a vector of firm-level features drive firms' emissions; $\boldsymbol{v} \perp \varepsilon, \eta$ is a vector of firm-level features that may overlap with $\boldsymbol{z}$ that drives firms' utilities of disclosing, u. Firms choose to disclose if and only if $u \geq 0$. Hence, as in Heckman (1979), estimating corporate emissions using the censored sample consisting of only disclosing firms results in $\mathbb{E}[\hat{s}|\boldsymbol{z}, u \geq 0] = \alpha + \boldsymbol{\beta}\boldsymbol{z} + \mathbb{E}[\varepsilon|\eta \geq -\gamma - \boldsymbol{\delta}\boldsymbol{v}]$ where the last term is the omitted regression term due to the selection. In the following discussion, I use $\psi \equiv \mathbb{E}[\varepsilon|\eta \geq -\gamma - \boldsymbol{\delta}\boldsymbol{v}]$ to represent the omitted term.

I.1 *Exogenous* disclosure utilities

First, I analyze the scenario where the stochastic disclosing utility (i.e., η) is uncorrelated with the specification error (i.e., ε). Firms' utilities of disclosure may still be correlated with observable firm-level features $\boldsymbol{v}$. For example, larger firms with greater publicity are more pressured to disclose. Regardlessly, the selection bias is zero, i.e., $\psi = 0$. Consequently, regressing observed emissions $\{\hat{s}\}_{u\geq 0}$ on these disclosing firms observable features $\{\boldsymbol{z}\}_{u\geq 0}$ still yield consistent estimates of f as long as disclosure propensity is not trivial.[1]

[1] Relatedly, if firms are subject to disclosure mandates that are exogenous to the specification error ε, estimation using data from disclosing firms also yields consistent estimates of f.

I.2 *Endogenous* disclosure utilities

Second, I analyze the scenario where the stochastic disclosing utility (i.e., η) is correlated with the specification error (i.e., ε), with the correlation coefficient given by ϱ. For example, firms with a more environmentally-conscious culture—an arguably difficult-to-measure feature—may simultaneously generate lower emissions and see more values in disclosure. The latent characteristic an organization's consciousness of environmental issues is part of ε due to its unobservability and is part of η due to the environmentally-conscious ones' preferences towards disclosure. Given ϱ, the omitted term is given by $\psi = \varrho\sigma\frac{\phi(y)}{\Phi(y)}$ where $y = \frac{\gamma+\boldsymbol{\delta v}}{\sigma_\eta}$.

If $\boldsymbol{v}$ and $\boldsymbol{z}$ do not correlate at all, regressing observed emissions $\{\hat{s}\}_{u\geq 0}$ on these disclosing firms observable features $\{\boldsymbol{z}\}_{u\geq 0}$ yield biased estimate of α—the intercept is biased up to a constant—but consistent estimate of $\boldsymbol{\beta}$. Conversely, if $\boldsymbol{v}$ and $\boldsymbol{z}$ are correlated (e.g., $z_1 = v_1$), both α and $\boldsymbol{\beta}$ will be biased. The exact biases on estimated parameters are given by regressing the omitted term ψ on $\boldsymbol{z}$ for observations whose $u \geq 0$.

I.3 Multiple selection criteria

Suppose there are two sample selection criteria. Firms still *voluntarily* disclose if $u \geq 0$. In addition, firms are *required* to disclose if they meet certain criteria. For example, firms are randomly selected to provide mandatory disclosures, or they are subject to the disclosure mandate if their total assets, denoted by $a = z_1 = v_1$, exceed a threshold a_0. There is no omitted self-selection bias if firms disclose due to the mandate, as firms' applicability to the mandate is orthogonal to the specification error ε. Let ξ denote firms' applicability—firms disclose if $u \geq 0$ or $\xi = 1$. In this case, regressing observed emissions $\{\hat{s}\}_{u\geq 0 \text{ or } \xi=1}$ on these disclosing firms observable features $\{\boldsymbol{z}\}_{u\geq 0 \text{ or } \xi=1}$ yield biased estimate of f. The exact biases on estimated parameters are given by regressing the effective omitted term $(1-\xi)\psi$ on $\boldsymbol{z}$ for observations whose $u \geq 0$ or $\xi = 1$.

I.4 Simulations

I conduct Monte Carlo simulation analysis to illustrate key insights from the preceding discussions. I simulate observed emissions (i.e., $\hat{s}$), voluntary disclosure utilities (i.e., u), and applicability to mandates (i.e., ξ) in an economy with 200,000 firms. Table OA.1 summarizes the most generalized set of parameters of the simulated economy. Firm characteristic $z_1 = v_1$ can be seen as firm size that simultaneously drives corporate emissions and disclosures. The negative correlation between η and ε intends to mimic the baseline model in which the specification error ε underlies firms' strategic disclosures. For illustrative purposes, I also consider alternative scenarios in which disclosure utility is not related to any factors driving firms' emissions (i.e., the baseline case studied in Section 2) or disclosure mandate is randomly assigned to firms (i.e., the baseline case studied in Section 4.3).

Table OA.1: Simulation analysis parametrization

This table presents the parametrization of the following data generating process:

$$\hat{s} = f(\boldsymbol{z}) + \varepsilon = \alpha + \boldsymbol{\beta}\boldsymbol{z} + \varepsilon$$
$$u = g(\boldsymbol{v}) + \eta = \gamma + \boldsymbol{\delta}\boldsymbol{v} + \eta$$

Firms voluntarily disclose if $u \geq 0$. Firms are subject to the mandate if $\xi = 1$.

$\boldsymbol{z}$	$\boldsymbol{\mu}_{\boldsymbol{z}} = (20, 8, 4)^\top$, $\boldsymbol{\Sigma}_{\boldsymbol{z}} = \begin{pmatrix} 25 & 4 & 0 \\ 4 & 9 & 0 \\ 0 & 0 & 16 \end{pmatrix}$
f	$\alpha = 10$, $\boldsymbol{\beta} = \left(1, 2, -\frac{1}{2}\right)^\top$
ε	$\varepsilon = \varepsilon^x + \varepsilon^s$, $\sigma_x = 4$, $\sigma_s = 3$
$\boldsymbol{v}$	$v_1 = z_1$, $v_2 = z_2$, $\mu_{v_3} = 2$, $\sigma_{v_3} = 2$
g	$\gamma = -2$, $\boldsymbol{\delta} = \left(\frac{1}{4}, -\frac{1}{2}, 1\right)^\top$
η	$\sigma_\eta = 2$, $\varrho = -.8$
ξ	$\xi = 1$ if $a = z_1 = v_1 \geq a_0 = 25$

Table OA.2 tabulates the results of estimating corporate emissions and biases in coefficients, if any, using data from only disclosing firms. The findings confirm earlier discussions on how sample selection criteria influence the estimation of corporate emissions.

Table OA.2: Regression analysis based on the simulated sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
u	$\eta \perp \varepsilon$	$\gamma + \eta$	$\gamma + \boldsymbol{\delta v} + \eta$		$\gamma + \eta$		$\gamma + \eta$	
ξ					$\xi = 1$ randomly		$\xi = 1$ if $a \geq a_0$	
Dep. Var.	$\hat{s}$	$\hat{s}$	$\hat{s}$	ψ	$\hat{s}$	$(1-\xi)\psi$	$\hat{s}$	$(1-\xi)\psi$
const.	10.027***	4.124***	7.871***	-2.203***	7.520***	-2.754***	-3.348***	-13.446***
	(0.065)	(0.090)	(0.060)	(0.017)	(0.100)	(0.057)	(0.094)	(0.034)
z_1	1.000***	0.991***	1.118***	0.121***	0.986***	-0.003	1.452***	0.456***
	(0.003)	(0.004)	(0.003)	(0.001)	(0.005)	(0.003)	(0.004)	(0.001)
z_2	2.000***	2.001***	1.754***	-0.246***	2.002***	0.006	1.997***	-0.004
	(0.005)	(0.007)	(0.005)	(0.001)	(0.008)	(0.004)	(0.007)	(0.002)
z_3	-0.503***	-0.511***	-0.504***	-0.001	-0.506***	-0.002	-0.504***	0.001
	(0.004)	(0.005)	(0.003)	(0.001)	(0.006)	(0.003)	(0.005)	(0.002)
N	124,107	31,448	123,667	123,667	58,223	58,223	58,275	58,275
R^2	0.763	0.869	0.789	0.251	0.740	0.000	0.856	0.675

Columns 1 through 4 examine the self-selection problem in the context of voluntary disclosure only. In column 1, estimation using the censored sample almost perfectly recovers the true coefficients, as disclosure utility is exogenous to the specification error ε. The setting underlying column 2 is the closest to the baseline model discussed in Section 2.[2] Results in column 2 suggest that, when disclosure utility is not related to any factors driving firms' emissions (e.g., size), the estimation bias is a constant, which equals the constant omitted term $\psi = \varrho\sigma\frac{\phi(y)}{\Phi(y)}$, where $y = \frac{\gamma}{\sigma_\eta}$.[3] When determinants of firms' disclosure utilities are correlated with their emissions drivers (e.g., firm size), both α and $\boldsymbol{\beta}$ are biased, as seen in column 3. Column 4 further confirms that the exact biases on model parameters are given by regressing the omitted term $\psi = \varrho\sigma\frac{\phi(y)}{\Phi(y)}$, where $y = \frac{\gamma+\boldsymbol{\delta v}}{\sigma_\eta}$, on $\boldsymbol{z}$ using the censored sample.[4] Importantly, the estimation bias, which is linear in $\boldsymbol{z}$, does not equal the omitted term ψ, which is non-linear in $\boldsymbol{z}$. The estimation bias is the "best" linear approximation of the actual omitted term.

Columns 5 through 8 examine the self-selection problem in the context of both voluntary disclosure and mandatory disclosure. The setting underlying columns 5 to 6 is the closest to the baseline model plus randomized assignment of disclosure mandates discussed in Section 4.3. Consistent with Corollary A.2, the bias shrinks as not all firms self-select into the disclosing sample. The estimation bias is still a constant, which equals the mean of the effective omitted term $(1-\xi)\psi$ among disclosing firms, i.e., $\mathbb{E}[(1-\xi)\psi|u \geq 0 \text{ or } \xi = 1] = \frac{(1-\theta)\Phi(y)}{\theta+(1-\theta)\Phi(y)}\varrho\sigma\frac{\phi(y)}{\Phi(y)}$ where $y = \frac{\gamma}{\sigma_\eta}$. If disclosure mandates are non-randomized based on some running variable $a = z_1$ which is also an emissions driver, the effective omitted term $(1-\xi)\psi$ will be correlated with the running variable z_1. Therefore, the bias is a function of z_1, as seen

[2]The baseline model can be seen as the special case in which $\eta = -\varepsilon$ and $\gamma = \iota$ is endogenously determined through Equation 5.

[3]Strictly speaking, the bias converges to the constant omitted term as the number of disclosing firms approaches infinity. For simplicity, I make the assumption from this point forward that there is an infinitely large number of firms in the economy and a non-trivial disclosure mass such that the estimates always approach their limiting values.

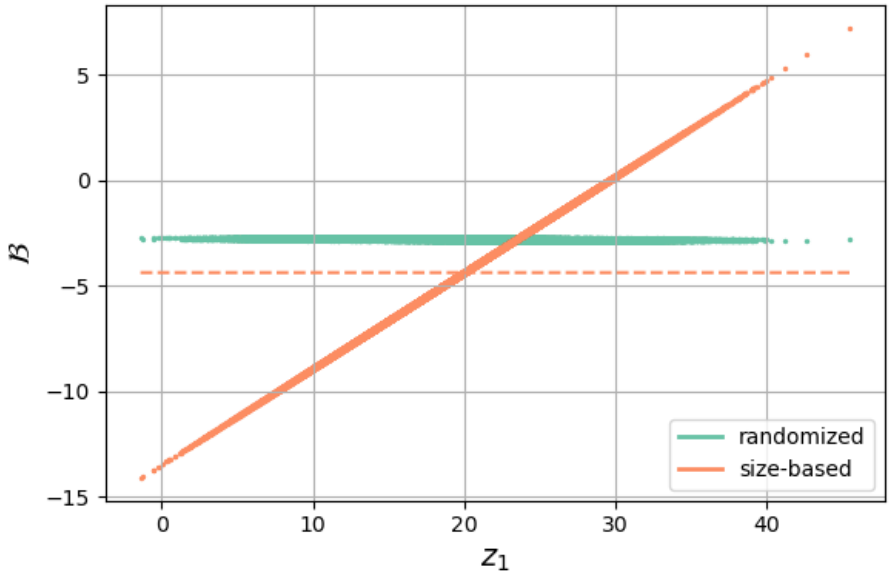
[4]In the simulation example, both z_1 and z_2 are simultaneously driving emissions and disclosing utilities. Therefore, estimation bias will load on z_1 and z_2 but not z_3, as seen in column 4.

in columns 7 to 8. Importantly, the estimation bias is, on average, larger under size-based mandates compared to randomized mandates. The bias is particularly pronounced when extrapolating the fitted model to out-of-sample firms that are smaller in size, as shown in the bottom-left of Figure OA.1.

Reverse-engineering the bias is challenging if sample selection criteria—voluntary or mandatory—are correlated with observable drivers of corporate emissions (i.e., $\boldsymbol{z}$). The estimation biases in these scenarios are functions of $\boldsymbol{z}$. More precisely, the biases are functions of the subset of emissions drivers $\boldsymbol{z}$ that also determine sample selection. De-biasing requires knowledge about disclosure utilities (to compute the omitted term ψ) and factors that drive both corporate emissions and emissions disclosure decisions (to estimate the biases on coefficients of f). In the following discussion of model extensions, I assume that both firms and rational investors are aware of the overlap between emissions drivers and determinants of disclosing utilities.

Figure OA.1: Estimation bias under randomized vs. size-based mandates

This figure plots the estimation bias under randomized versus size-based mandates that require the same amount of firms to disclose. The estimation biases are computed based on estimated coefficient from column 6 and column 8 of Table OA.2. The dashed oranage line reflects the unconditional mean of the estimation bias under size-based disclosure mandates.



II Model Extensions

In this online appendix, I analyze the extensions of my baseline model outlined in Section 2.

II.1 The Dye (1985) friction

First, suppose there is probability p that the firm is not informed about its emissions, i.e., the firm does not observe its emissions measurement $\hat{s}$. In this case, rational investors' valuation of firms' emissions conditional on non-disclosure changes from $x + \frac{1}{\kappa+1}\mathbb{E}[\varepsilon|d = \varnothing]$ to the weighted average of the scenario when the firm is uninformed and the scenario when the firm is informed that choose to withhold, i.e., $\frac{px+(1-p)\Pr(d=\varnothing)\left(x+\frac{1}{\kappa+1}\mathbb{E}[\varepsilon|d=\varnothing]\right)}{p+(1-p)\Pr(d=\varnothing)}$ where $\Pr(d = \varnothing)$ denotes the probability of non-disclosure. Hence, the equilibrium disclosure strategy is characterize as follows: for $\lambda \in (0, 1)$ and $p \in (0, 1)$, there exists a threshold $\iota_p \in \mathbb{R}$ such that firms disclose if and only if $\varepsilon \leq \iota_p$:

$$\iota_p = (1 - \lambda)\mathcal{P} + \lambda\mathbb{E}[\varepsilon|\varepsilon \leq \iota_p]$$

where rational punishment for non-disclosure $\mathcal{P}$ is given by

$$\mathcal{P} = \frac{(1 - p)(1 - \Phi(\tilde{\iota}_p))\mathbb{E}[\varepsilon|\varepsilon \geq \iota_p]}{p + (1 - p)(1 - \Phi(\tilde{\iota}_p))}$$

where $\tilde{\iota}_p \equiv \frac{\iota(p)}{\sigma}$ is the standardized disclosure threshold.

Figure OA.2 plots the numerically solved equilibrium disclosure strategy in the presence of the Dye friction. Holding the share of naive investors constant, an increase in the information friction p reduces firms' disclosing incentives as firms who are informed about their high emissions can further pool with uninformed firms valued at the expected emissions x. Such effect only channels through rational investors' valuation, though. Consequently, the impact of the Dye friction is only salient when the share of naive investors is small.

II.2 Firm-specific disclosure utilities

In this online appendix, I analyze scenarios in which there are factors other than investors' emissions pricing that enter firms' utility functions. I use $u = \eta + \gamma + \boldsymbol{\delta v}$ to represent each individual firm's disclosing utility on top of the emissions pricing. There are three components in the firm-specific disclosing utility: $\eta \sim \mathcal{N}(0, \sigma_\eta^2), \eta \perp \varepsilon$ captures the stochastic disclosing utility; $\gamma \in \mathbb{R}$ captures a fixed disclosing utility ($\gamma > 0$) or disutility ($\gamma < 0$) à la Verrecchia (1983); $\boldsymbol{\delta v}$ captures firm-specific disclosing utilities that vary with firms' observable features $\boldsymbol{v}$. Given the extra disclosing utility u, the firm's problem changes from $\min_{d\in\{\hat{s},\varnothing\}} \mathbb{E}_\lambda[e|\hat{x}, d]$ to $\min_{d\in\{\hat{s},\varnothing\}} \mathbb{E}_\lambda[e|\hat{x}, d] + \mathbb{1}_{d=\hat{s}}u$. Put differently, firms choose to disclose if and only if

$$\varepsilon \leq (1-\lambda)\mathbb{E}[\varepsilon|d = \varnothing] + \lambda\mathbb{E}[\varepsilon|d = \hat{s}] + (\kappa+1)u$$

The intuition with the $\kappa + 1$ scalar on u is as follows. Holding σ constant, the more precise intermediaries' emissions modeling is relative to firms' emissions measurement (i.e., higher κ), the smaller the updates investors make after observing disclosed emissions, while the extra disclosure utility is unaffected. Therefore, there is—on a relative basis—more salient impact of the extra disclosure utility when κ is higher. In the following, I analyze the impacts of these three components to the equilibrium disclosure strategy separately.

If $u = \eta$, there exists a threshold $\iota_\eta \in \mathbb{R}$ such that firms disclose if and only if $\varepsilon - (\kappa+1)\eta \leq \iota_\eta$:

$$\iota_\eta = (1-\lambda)\mathbb{E}[\varepsilon|\varepsilon - (\kappa+1)\eta \geq \iota_\eta] + \lambda\mathbb{E}[\varepsilon|\varepsilon - (\kappa+1)\eta \leq \iota_\eta]$$

$\tilde{\iota}_\eta \equiv \frac{\iota_\eta}{(\sigma^2+(\kappa+1)^2\sigma_\eta^2)^{\frac{1}{2}}}$ denoting the standardized disclosure threshold.

The consequence of the stochastic disclosing utility η is that it introduces "fuzziness" into the sample selection criterion $\varepsilon \leq \iota_\eta + (\kappa+1)\eta$. The "fuzziness" simultaneously mitigates the estimation bias due to self-selection and the sharpness in rational investors' valuation of non-disclosure. Therefore, an increase in the variation of η flattens the sensitivity of the

equilibrium disclosure threshold to the share of naive investors, as seen in Figure OA.3.[5] If the variation in *eta* is large enough, the strategic disclosure incentive will be overpowered and firms will disclose quasi-randomly.

If $u = \gamma$, there exists a threshold $\iota_\gamma \in \mathbb{R}$ such that firms disclose if and only if $\varepsilon \leq \iota_\gamma$:

$$\iota_\gamma = (1-\lambda)\mathbb{E}[\varepsilon|\varepsilon \geq \iota_\gamma] + \lambda\mathbb{E}[\varepsilon|\varepsilon \leq \iota_\gamma] + (\kappa+1)\gamma$$

with $\tilde{\iota}_\gamma \equiv \frac{\iota_\gamma}{\sigma}$ denoting the standardized disclosure threshold.

Intuitively, a common disclosing utility (disutility) will shift the disclosure threshold up (down), as more (less) firms will find it in their interests to disclose. Such patterns are seen in Figure OA.4.

If $u = \boldsymbol{\delta v}$, the estimation bias will be a linear function of $\boldsymbol{v}$, as seen in Table OA.2. The linear function is an approximation of the true non-linear relationship between the omitted term, which involves normal p.d.f. and normal c.d.f., and $\boldsymbol{v}$. First, suppose the estimator is able to capture any non-linear relationship between corporate emissions and emissions drivers in the disclosing sample. In this case, the equilibrium disclosure strategy is characterized by a firm-specific threshold $\iota_{\boldsymbol{\delta v}} \in \mathbb{R}$ such that firms disclose if and only if $\varepsilon \leq \iota_{\boldsymbol{\delta v}}$:

$$\iota_{\boldsymbol{\delta v}} = (1-\lambda)\mathbb{E}[\varepsilon|\varepsilon \geq \iota_{\boldsymbol{\delta v}}] + \lambda\mathbb{E}[\varepsilon|\varepsilon \leq \iota_{\boldsymbol{\delta v}}] + (\kappa+1)\boldsymbol{\delta v}$$

with $\tilde{\iota}_{\boldsymbol{\delta v}} \equiv \frac{\iota_{\boldsymbol{\delta v}}}{\sigma}$ denoting the standardized disclosure threshold.

For each firm, the firm-specific disclosure utility $\boldsymbol{\delta v}$ affects its disclosing incentives in a manner similar to the common disclosure utility γ. However, unlike the prior cases, the disclosure threshold and the corresponding disclosure propensity vary across firms. In ag-

[5] From Figure OA.3 onwards, I narrow the range of λ for the following reason. In `scipy`, the normal p.d.f. and normal c.d.f. functions (`scipy.stats.norm.pdf` and `scipy.stats.norm.cdf`, respectively) do not work well with large values, which affects evaluations of means of truncated normal distributions. The calculated expectations do not approach infinity as they should be for extreme lower-tail and upper-tail truncated normal distributions.

gregate, the average disclosure propensity is given by $\int \Phi(\tilde{\iota}_{\boldsymbol{\delta v}})dF(\boldsymbol{v})$ where F is the joint c.d.f. of $\boldsymbol{v} \sim \mathcal{N}(\boldsymbol{\mu_v}, \boldsymbol{\Sigma_v})$. The solid lines in Figure OA.5 illustrate the aggregate disclosure propensity for different parametrizations of a univariate $v \sim \mathcal{N}(\mu_v, \sigma_v^2)$. Similar to how γ affects the disclosure propensity, the average disclosure propensity shifts up (down) when μ_v is higher (lower). Similar to how σ_η affects the disclosure propensity, the sensitivity of average disclosure propensity to λ flattens when σ_v is larger.

The more realistic scenario, however, is that the estimator is not able to fully capture the non-linear relationship between corporate emissions and emissions drivers in the disclosing sample (e.g., using linear approximation). If the estimator sticks to its linear emissions model, the estimates will be noisier, effectively decreasing κ. Hence, the dashed lines are always slightly above the solid lines in Figure OA.5. The approximation error is smaller if v is deterring disclosure or the variation in v is smaller. The approximation error is also smaller if rational investors are the dominant majority, as the approximation error only affects the equilibrium through naive investors' emissions pricing. Overall, the approximation error does not cause substantial deviations in equilibrium outcomes.

II.3 Disclosure mandates

In this online appendix, I extend the analysis in Section 4.3 by further discussing impacts of non-randomized disclosure mandates. The majority of forthcoming mandatory emissions disclosure regulations are applied to firms based on their size. For example, EU's Corporate Sustainability Reporting Directive (Directive 2022/2464) and Califonia's Climate Corporate Data Accountability Act (S.B. 253) apply to *large* firms only. Specifically, the CSRD will be phased in starting FY2024 for listed EU firms with over 500 employees, as well as banks and insurance companies. By FY2025, the directive will extend to EU firms that meet at least two of the following criteria: more than 250 employees, €40 million in turnover, or €20 million in total assets. California's S.B. 253 requires businesses with $1 billion or more in annual revenue that operate in California to calculate and publicly and report their GHG

emissions starting FY2025.

The simulation result in Table OA.2 suggests that the estimation biases differ depending on whether the mandate is randomized or non-randomized (e.g., size-based). While a randomized disclosure mandate lead to a constant estimation bias across all firms, the bias under a size-based disclosure mandate a function of firm size. Suppose the disclosure mandate requires firms to disclose if their size, denoted by $z \sim \mathcal{N}(\mu_z, \sigma_z^2)$ and is part of their emissions drivers $\boldsymbol{z}$, exceeds a certain threshold, denoted by z_0. The equilibrium disclosure strategy is characterized by a firm-specific threshold $\iota_{z,z_0} \in \mathbb{R}$ such that firms *voluntarily* disclose if and only if $\varepsilon \leq \iota_{z,z_0}$ and they are not subject to the mandate (i.e., $z \leq z_0$):

$$\iota_{z,z_0} = (1-\lambda)\mathbb{E}[\varepsilon|\varepsilon \geq \iota_{z,z_0}] + \lambda \mathcal{B}_{z,z_0}$$

where $\mathcal{B}_{z,z_0} = \hat{a} + \hat{b}z$ comes from estimating the linear model $(1 - \mathbb{1}_{z \geq z_0})\sigma \frac{-\phi(\tilde{\iota}_{z,z_0})}{\Phi(\tilde{\iota}_{z,z_0})} \sim a + bz$ with $\tilde{\iota}_{z,z_0} \equiv \frac{\iota_{z,z_0}}{\sigma}$ denoting the standardized disclosure threshold.

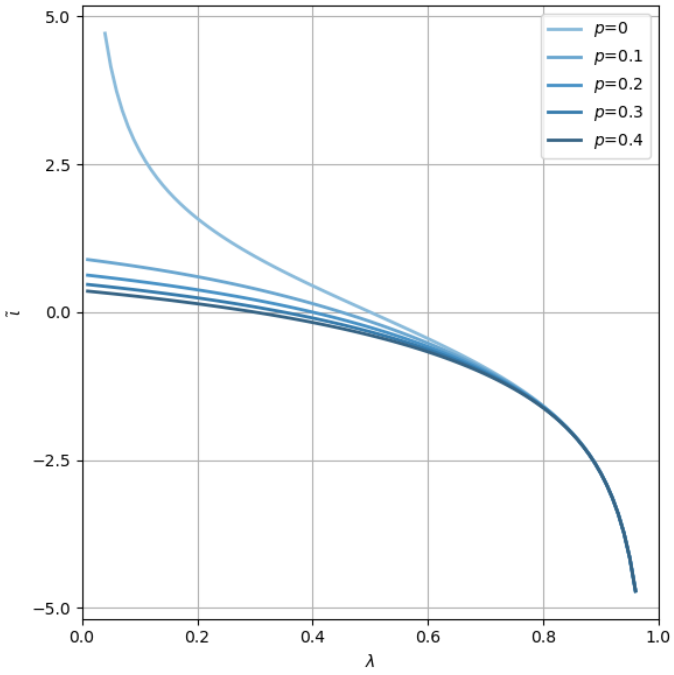
Figure OA.6 compares the average disclosure propensity and average incremental disclosure under two scenarios: when the mandate is randomized versus when it is size-based.[6] The results indicate that while size-based mandates still help mitigate the selection bias in intermediaries' estimates and unleash more voluntary disclosures, their effects are less pronounced compared to randomly requiring the same proportion of firms to disclose. The difference arises because the selection bias is, on average, larger when the mandate is size-based compared to when it is randomized. The bias is larger when extrapolating the fitted model to out-of-sample firms that are smaller in size, as illustrated in Figure OA.1. Overall, size-based disclosure mandates shall be expected to be less effective in complementing intermediaries' estimates and incentivizing voluntary disclosures compared to randomized disclosure mandates.

[6] In calibrating the average disclosure propensity under size-based disclosure mandates, I assume $F = \Phi$, i.e., z follows standard normal. In untabulated analyses, I confirm that μ_z and σ_z does not affect the average disclosure propensity in equilibrium.

Figure OA.2: Equilibrium disclosure strategy with the Dye friction p

This figure plots the numerically solved equilibrium disclosure thresholds (standardized), $\tilde{\iota}_p$, and the corresponding disclosure propensity, $\mathbb{E}d_p$, in the presence of the Dye friction denoted by p.

Panel A: The standardized threshold, $\tilde{\iota}_p$



Panel B: The disclosure propensity, $\mathbb{E}d_p$

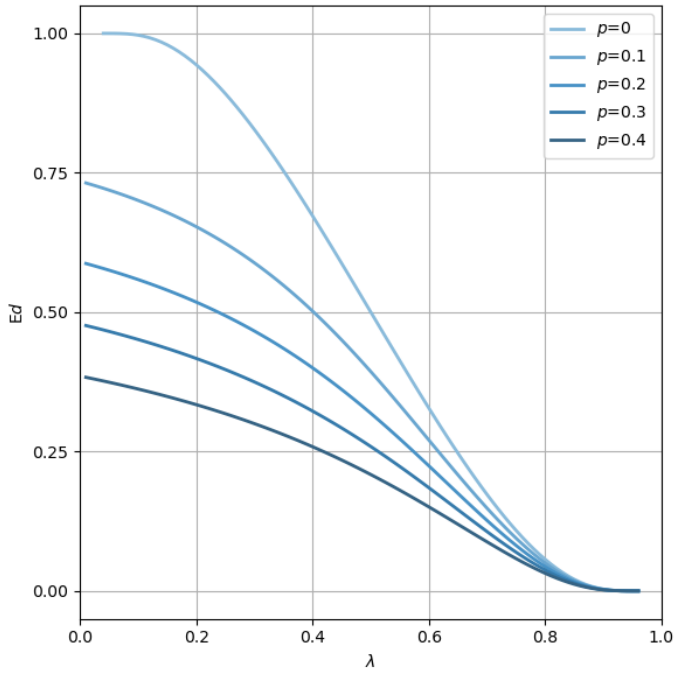
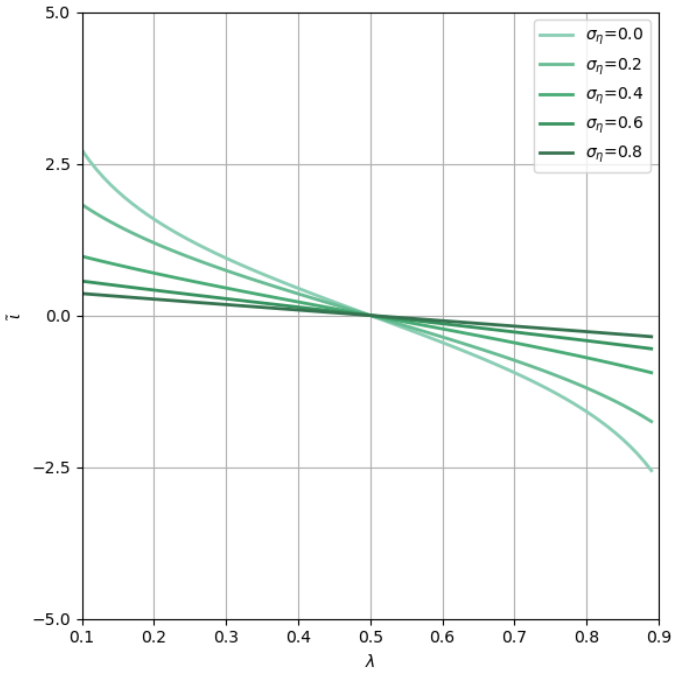


Figure OA.3: Equilibrium disclosure strategy with the stochastic utility η

This figure plots the numerically solved equilibrium disclosure thresholds (standardized), $\tilde{\iota}_\eta$, and the corresponding disclosure propensity, $\mathbb{E}d_\eta$, in the presence of stochastic disclosing utility η. The standard deviation of η, σ_η, is standardized by σ and κ is set to $\frac{1}{2}$ in the following plots.

Panel A: The standardized threshold, $\tilde{\iota}_\eta$



Panel B: The disclosure propensity, $\mathbb{E}d_\eta$

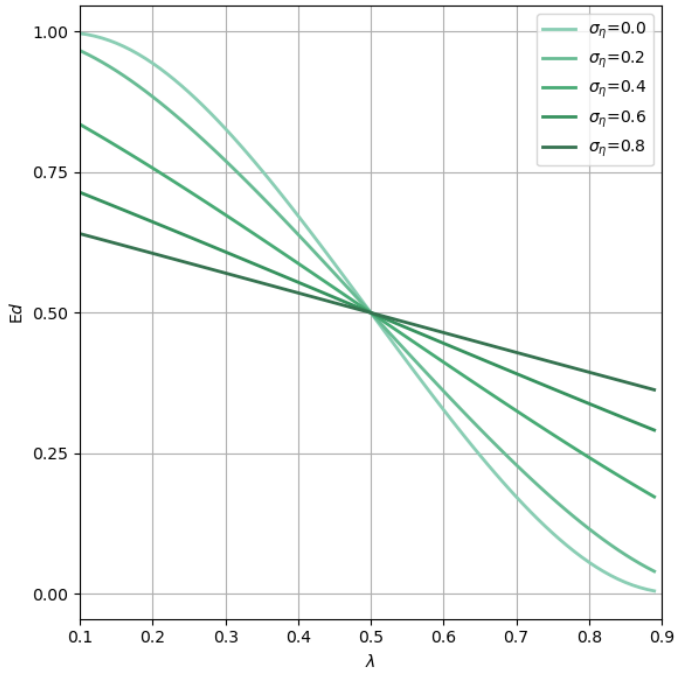
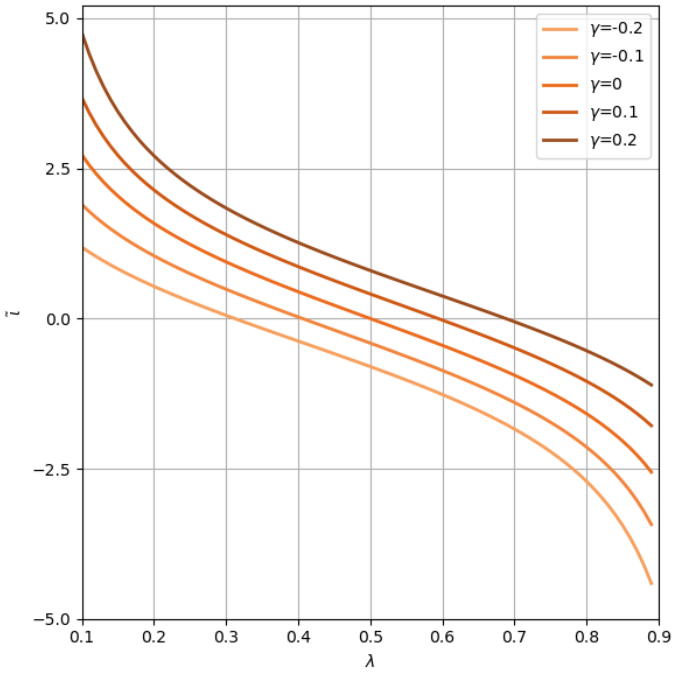


Figure OA.4: Equilibrium disclosure strategy with common disclosure utility γ

This figure plots the numerically solved equilibrium disclosure thresholds (standardized), $\tilde{\iota}_\gamma$, and the corresponding disclosure propensity, $\mathbb{E}d_\gamma$, in the presence of the common disclosing utility $\gamma \in \mathbb{R}$. γ is standardized by σ and κ is set to $\frac{1}{2}$ in the following plots.

Panel A: The standardized threshold, $\tilde{\iota}_\gamma$



Panel B: The disclosure propensity, $\mathbb{E}d_\gamma$

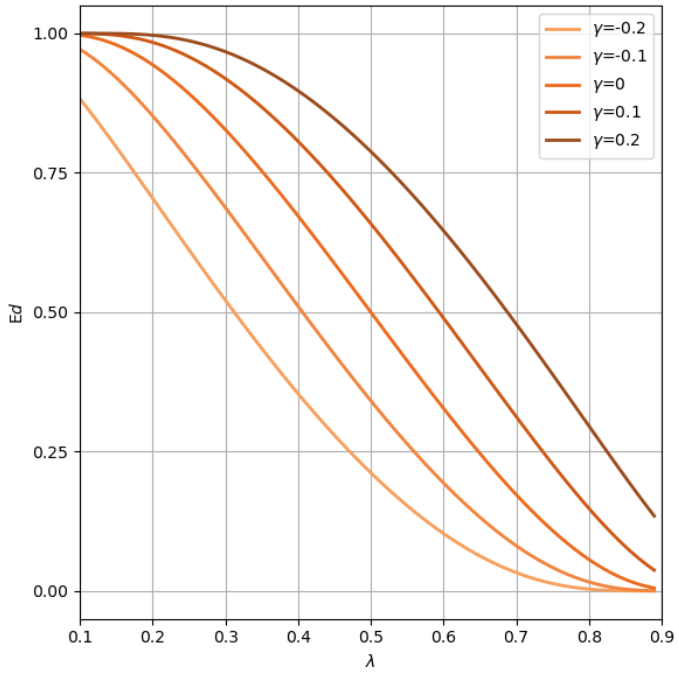
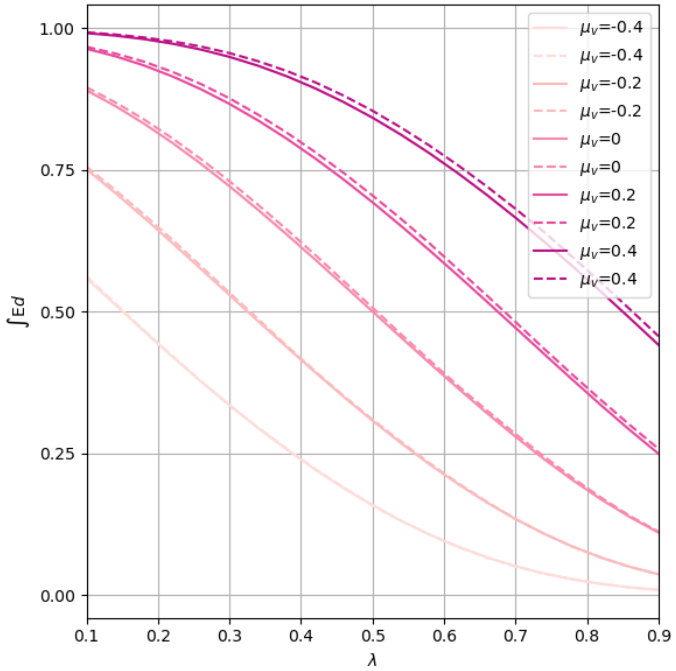


Figure OA.5: Equilibrium disclosure strategy with firm-specific utility δv

This figure plots the numerically solved aggregate disclosure propensity, $\int_v \mathbb{E}d_{\delta v}dF(v)$, in the presence of the firm-specific disclosing utility δv where $v \sim \mathcal{N}(\mu_v, \sigma_v^2)$ is also a determinant of firms' emissions. v is standardized by σ, δ is set to one, and κ is set to $\frac{1}{2}$ in the following plots. In Panel A, $\sigma_v = 0.3$, in Panel B, $\mu_v = 0.2$. The solid lines are calibrated assuming that the estimator can fully capture any non-linear relationship between corporate emissions and v in the disclosing sample. The dashed lines are calibrated assuming that the estimator sticks to the linear emissions model.

Panel A: Average disclosure propensity and μ_v



Panel B: Average disclosure propensity and σ_v

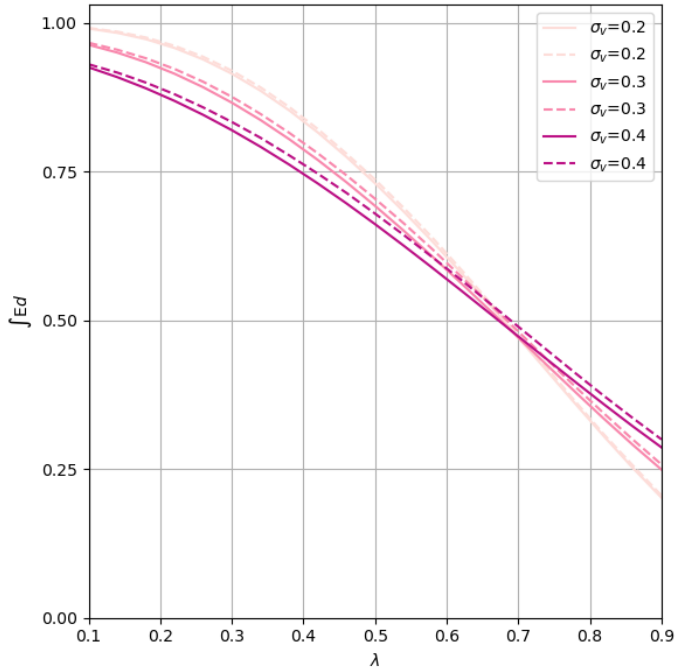
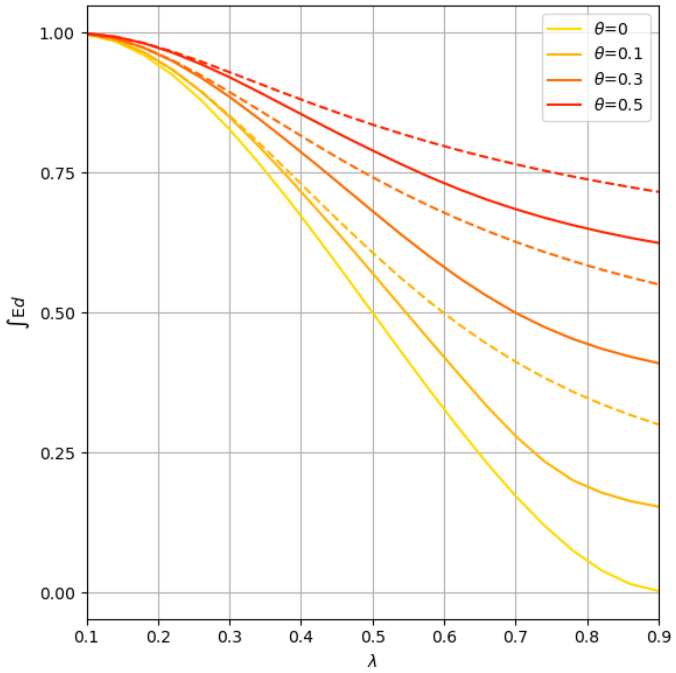


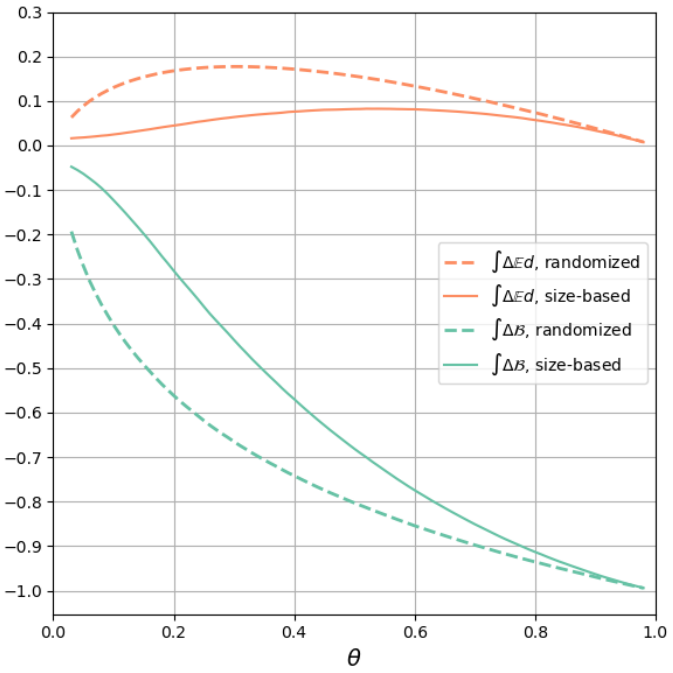
Figure OA.6: Equilibrium disclosure for randomized vs. size-based mandates

This figure plots the numerically solved aggregate disclosure propensity, $\int_v \mathbb{E}d_{z,z_0}d\Phi(z)$ under randomized vs. size-based disclosure mandates. For each pair of comparison, I use θ to denote the fraction of firms that are required to disclose, either randomly or because their size is above a threshold $z_0 = -\Phi^{-1}(\theta)$. In Panel A, I compared the average disclosure propensity under the randomized mandate (dashed lines) versus the size-based mandate (solid lines). Using $\hat{\lambda} = 0.65$ from Table 6, Panel B plots the incremental disclosure $\int \Delta\mathbb{E}d$, defined as the eventual disclosure (mandatory plus voluntary) minus the hypothetical disclosure computed by simply stacking the mandatory disclosure over the voluntary disclosure without considering the interaction, and the shrinkage in bias $\int \Delta\mathcal{B}$, defined as the reduction in the magnitude of the average bias.

Panel A: Average disclosure propensity: randomized vs. size-based



Panel B: Incremental disclosure and shrinkage in bias: randomized vs. size-based



III Uniqueness of Equilibrium under Risk Aversion

In this section, I prove the existence and the uniqueness of the equilibria characterized in Lemma A.1 and Proposition A.1. In the following proof, I work with standardized thresholds in the spirit of Corollary 1. Let ϕ and Φ denote standard normal p.d.f. and c.d.f., respectively.

Let $F(x) = (1-\Lambda(x))u(x) - x$ where $u(x) = \frac{\phi(x)}{1-\Phi(x)}$, $\Lambda(x) \equiv \lambda^N$. $\frac{d\Lambda}{dx} < 0$ as the upper-tail truncated variance is decreasing as the threshold increases. Showing that $i^N \in \mathbb{R}$ that solves Equation A.1 is unique is equivalent to showing the uniqueness of $x \in \mathbb{R}$ such that $F(x) = 0$.

To see this, let $G(a, x) = (1-a)u(x) - x$, $a \in (0, 1)$. We know that $\frac{\partial G}{\partial a} < 0$ and $\frac{\partial G}{\partial x} < 0$. For $x \in \mathbb{R}$, $\Lambda(x) \in (\frac{\kappa}{\kappa+1}\lambda, \lambda)$. Let $\underline{a} \equiv \frac{\kappa}{\kappa+1}\lambda$, $\overline{a} \equiv \lambda$. Let $\underline{x}_0, \overline{x}_0$ be the unique solutions to $G(\overline{a}, x) = 0$ and $G(\underline{a}, x) = 0$; that is $G(\overline{a}, \underline{x}_0) = 0$, $G(\underline{a}, \overline{x}_0) = 0$. Hence, $\underline{x}_0 < \overline{x}_0$.

For any $x \leq \underline{x}_0$,

$$F(x) = G(\Lambda(x), x) > G(\overline{a}, x) \geq G(\overline{a}, \underline{x}_0) = 0$$

where the first (strict) inequality is due to $\Lambda(x) < \overline{a}$, $\frac{\partial G}{\partial a} < 0$, and the second inequality is due to $\frac{\partial G}{\partial x} < 0$. Similarly, for any $x \geq \overline{x}_0$,

$$F(x) = G(\Lambda(x), x) < G(\underline{a}, x) \leq G(\underline{a}, \overline{x}_0) = 0$$

The above two conditions suggest that $\exists x \in (\underline{x}_0, \overline{x}_0)$ such that $F(x) = 0$. Since $F(\underline{x}_0) = G(\Lambda(\underline{x}_0), \underline{x}_0) > 0$, it must be the case that $\exists \underline{x}_1 > \underline{x}_0$ such that $G(\Lambda(\underline{x}_0), \underline{x}_1) = 0$. Likewise, since $F(\overline{x}_0) = G(\Lambda(\overline{x}_0), \overline{x}_0) < 0$, $\exists \overline{x}_1 < \overline{x}_0$ such that $G(\Lambda(\overline{x}_0), \overline{x}_1) = 0$. Moreover, $\underline{x}_1 < \overline{x}_1$.

For any $x \in (\underline{x}_0, \underline{x}_1]$, $\Lambda(x) < \Lambda(\underline{x}_0)$ as $\frac{d\Lambda}{dx} < 0$. Therefore,

$$F(x) = G(\Lambda(x), x) > G(\Lambda(\underline{x}_0), x) \geq G(\Lambda(\underline{x}_0), \underline{x}_1) = 0$$

and for any $x \in [\overline{x}_1, \overline{x}_0)$, $\Lambda(x) > \Lambda(\overline{x}_0)$ as $\frac{d\Lambda}{dx} < 0$. Therefore,

$$F(x) = G(\Lambda(x), x) < G(\Lambda(\overline{x}_0), x) \leq G(\Lambda(\overline{x}_0), \overline{x}_1) = 0$$

The above two conditions further suggest that $\exists x \in (\underline{x}_1, \overline{x}_1)$ such that $F(x) = 0$. Note that the interval narrows. Repeat the above processes until the interval converge, there must be a unique x^* that solves $F(x) = 0$. Moreover, $\frac{dF}{dx}|_{x=x^*} < 0$. $\square$

Let $F(x) = (1 - \Lambda(x))u(x) + \Lambda l(x) + \kappa(\Lambda(x) - \lambda)l(x)$ where $u(x) = \frac{\phi(x)}{1-\Phi(x)}$, $l(x) = \frac{-\phi(x)}{\Phi(x)}$, $\Lambda(x) \equiv \Lambda$ with $\frac{d\Lambda}{dx} < 0$. Showing that $i \in \mathbb{R}$ that solves Equation A.2 is unique is equivalent to showing the uniqueness of $x \in \mathbb{R}$ such that $F(x) = 0$.

To see this, let $G(a, x) = (1 - a)u(x) + al(x) + \kappa(a - \overline{a})l(x) - x$, $a \in (\underline{a}, \overline{a})$. From Corollary 1 we know that $(1 - a)u(x) + al(x) - x$ is strictly decreasing in both a and x. Importantly $\kappa(a - \overline{a}) < 0$ and $\frac{dl}{dx} > 0$ suggest that $\frac{\partial G}{\partial x} < 0$; $l(x) < 0$ for any $x \in \mathbb{R}$ suggests that $\frac{\partial G}{\partial a} < 0$. Let $\underline{x}_0, \overline{x}_0$ be the unique solutions to $G(\overline{a}, x) = 0$ and $G(\underline{a}, x) = 0$. Hence, $\underline{x}_0 < \overline{x}_0$. For any $x \leq \underline{x}_0$,

$$F(x) = G(\Lambda(x), x) > G(\overline{a}, x) \geq G(\overline{a}, \underline{x}_0) = 0$$

and for any $x \geq \overline{x}_0$,

$$F(x) = G(\Lambda(x), x) < G(\underline{a}, x) \leq G(\underline{a}, \overline{x}_0) = 0$$

The above two conditions suggest that $\exists x \in (\underline{x}_0, \overline{x}_0)$ such that $F(x) = 0$. Since $G(\Lambda(\underline{x}_0), \underline{x}_0) > 0$, it must be the case that $\exists \underline{x}_1 > \underline{x}_0$ such that $G(\Lambda(\underline{x}_0), \underline{x}_1) = 0$. Likewise, since $G(\Lambda(\overline{x}_0), \overline{x}_0) < 0$, $\exists \overline{x}_1 < \overline{x}_0$ such that $G(\Lambda(\overline{x}_0), \overline{x}_1) = 0$. Moreover, $\underline{x}_1 < \overline{x}_1$.

For any $x \in (\underline{x}_0, \underline{x}_1]$, $\Lambda(x) < \Lambda(\underline{x}_0)$ as $\frac{d\Lambda}{dx} < 0$. Therefore,

$$F(x) = G(\Lambda(x), x) > G(\Lambda(\underline{x}_0), x) \geq G(\Lambda(\underline{x}_0), \underline{x}_1) = 0$$

and for any $x \in [\overline{x}_1, \overline{x}_0)$, $\Lambda(x) > \Lambda(\overline{x}_0)$ as $\frac{d\Lambda}{dx} < 0$. Therefore,

$$F(x) = G(\Lambda(x), x) < G(\Lambda(\overline{x}_0), x) \leq G(\Lambda(\overline{x}_0), \overline{x}_1) = 0$$

The above two conditions further suggest that $\exists x \in (\underline{x}_1, \overline{x}_1)$ such that $F(x) = 0$. Note that the interval narrows. Repeat the above processes until the interval converge, there must be a unique x^* that solves $F(x) = 0$. Moreover, $\frac{dF}{dx}|_{x=x^*} < 0$. $\square$

IV Proof of Table A.1

Proof. This section provides the proof of results tabulated in Table A.1. First, when the intermediary does not correct the bias, naive investors' pricing error reads

$$\mathbb{E}^N[e|\mathcal{I}] - e = \begin{cases} \frac{\kappa}{\kappa+1}\left(-\varepsilon^x + \mathbb{E}[\varepsilon|\varepsilon \leq \iota]\right) + \frac{1}{\kappa+1}\varepsilon^s & d = \hat{s} \\ -\varepsilon^x + \mathbb{E}[\varepsilon|\varepsilon \leq \iota] & d = \varnothing \end{cases}$$

while rational investors' pricing error reads

$$\mathbb{E}[e|\mathcal{I}] - e = \begin{cases} \frac{\kappa}{\kappa+1}(-\varepsilon^x) + \frac{1}{\kappa+1}\varepsilon^s & d = \hat{s} \\ -\varepsilon^x + \frac{1}{\kappa+1}\mathbb{E}[\varepsilon|\varepsilon \geq \iota] & d = \varnothing \end{cases}$$

Notice that $-\frac{\kappa}{\kappa+1}\varepsilon^x + \frac{1}{\kappa+1}\varepsilon^s$ and $\varepsilon \equiv \varepsilon^x + \varepsilon^s$ are orthogonal as

$$\begin{aligned} \operatorname{cov}\left(-\frac{\kappa}{\kappa+1}\varepsilon^x + \frac{1}{\kappa+1}\varepsilon^s, \varepsilon^x + \varepsilon^s\right) &= -\frac{\kappa}{\kappa+1}\sigma_x^2 + \frac{1}{\kappa+1}\sigma_s^2 \\ &= \frac{-\sigma_s^2\sigma_x^2 + \sigma_x^2\sigma_s^2}{\sigma_x^2 + \sigma_s^2} \\ &= 0 \end{aligned}$$

where the independence between ε^x leads to $\varepsilon^s \operatorname{cov}(\varepsilon^x, \varepsilon^s) = 0$ and κ is defined as $\frac{\tau_x}{\tau_s}$ and hence $\frac{\sigma_s^2}{\sigma_z^2}$. Therefore, the conditional expectation and variance of $-\frac{\kappa}{\kappa+1}\varepsilon^x + \frac{1}{\kappa+1}\varepsilon^s$ are equivalent to their corresponding unconditional values.

$$\begin{aligned} &\mathbb{E}\left[-\frac{\kappa}{\kappa+1}\varepsilon^x + \frac{1}{\kappa+1}\varepsilon^s\right] = 0 \\ &\mathbb{V}\left(-\frac{\kappa}{\kappa+1}\varepsilon^x + \frac{1}{\kappa+1}\varepsilon^s\right) = \left(\frac{\kappa}{\kappa+1}\right)^2\sigma_x^2 + \left(\frac{1}{\kappa+1}\right)^2\sigma_s^2 = \left(\sigma\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2 \end{aligned}$$

where the last term is simplified from $\frac{\sigma_s^4\sigma_x^2+\sigma_x^4\sigma_s^2}{(\sigma_s^2+\sigma_x^2)^2}$. Utilizing these results to evaluate second

moments of investors' pricing error conditional on disclosure yields

$$\mathbb{E}\left[\left(\mathbb{E}^N[e|\hat{x},\hat{s}]-e\right)^2\right]=\sigma^2\left(\frac{\kappa}{\kappa+1}\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}\right)^2+\sigma^2\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$$
$$\mathbb{E}\left[\left(\mathbb{E}[e|\hat{x},\hat{s}]-e\right)^2\right]=\sigma^2\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2$$

Evaluating investors' pricing error conditional on non-disclosure requires analyzing properties of ε^x conditional on ε. Recall that $-\frac{\kappa}{\kappa+1}\varepsilon^x+\frac{1}{\kappa+1}\varepsilon^s$ and $\varepsilon\equiv\varepsilon^x+\varepsilon^s$ are orthogonal, which means

$$(1+\kappa)\mathbb{E}[\varepsilon^x|\varepsilon\leq\iota]=\mathbb{E}[\varepsilon|\varepsilon\leq\iota]-(\kappa+1)\mathbb{E}\left[-\frac{\kappa}{\kappa+1}\varepsilon^x+\frac{1}{\kappa+1}\varepsilon^s\right]=\sigma\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}$$
$$(1+\kappa)\mathbb{E}[\varepsilon^x|\varepsilon\geq\iota]=\mathbb{E}[\varepsilon|\varepsilon\geq\iota]-(\kappa+1)\mathbb{E}\left[-\frac{\kappa}{\kappa+1}\varepsilon^x+\frac{1}{\kappa+1}\varepsilon^s\right]=\sigma\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}$$

and

$$\begin{aligned}(1+\kappa)^2\mathbb{V}(\varepsilon^x|\varepsilon\leq\iota)&=\mathbb{V}(\varepsilon|\varepsilon\leq\iota)+(\kappa+1)^2\mathbb{V}\left(-\frac{\kappa}{\kappa+1}\varepsilon^x+\frac{1}{\kappa+1}\varepsilon^s\right)\\&=\sigma^2\left(1-\frac{\tilde{\iota}\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}-\left(\frac{\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}\right)^2\right)+(\kappa+1)^2\sigma^2\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2\\(1+\kappa)^2\mathbb{V}(\varepsilon^x|\varepsilon\geq\iota)&=\mathbb{V}(\varepsilon|\varepsilon\geq\iota)+(\kappa+1)^2\mathbb{V}\left(-\frac{\kappa}{\kappa+1}\varepsilon^x+\frac{1}{\kappa+1}\varepsilon^s\right)\\&=\sigma^2\left(1+\frac{\tilde{\iota}\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}-\left(\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}\right)^2\right)+(\kappa+1)^2\sigma^2\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2\end{aligned}$$

Utilizing these results to evaluate second moments of investors' pricing error conditional on

non-disclosure yields

$$\begin{aligned}\mathbb{E}\left[\left(\mathbb{E}^N[e|\hat{x},\varnothing]-e\right)^2\right]=&\sigma^2\left(\frac{-\phi(\tilde{\iota})}{\Phi(\tilde{\iota})}-\frac{1}{\kappa+1}\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}\right)^2\\&+\sigma^2\left(1+\frac{\tilde{\iota}\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}-\left(\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}\right)^2\right)\frac{1}{(\kappa+1)^2}+\sigma^2\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2\\\mathbb{E}\left[\left(\mathbb{E}[e|\hat{x},\varnothing]-e\right)^2\right]=&\sigma^2\left(1+\frac{\tilde{\iota}\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}-\left(\frac{\phi(\tilde{\iota})}{1-\Phi(\tilde{\iota})}\right)^2\right)\frac{1}{(\kappa+1)^2}+\sigma^2\left(\frac{1}{\sqrt{\kappa}+\frac{1}{\sqrt{\kappa}}}\right)^2\end{aligned}$$

Thus far, I have proved the results in Panel A of Table A.1. When the intermediary does correct the bias, naive investors' pricing error reads

$$\mathbb{E}^N[e|\mathcal{I}]-e=\begin{cases}\frac{\kappa}{\kappa+1}\left(-\varepsilon^x\right)+\frac{1}{\kappa+1}\varepsilon^s & d=\hat{s}\\-\varepsilon^x & d=\varnothing\end{cases}$$

while rational investors' pricing error reads

$$\mathbb{E}[e|\mathcal{I}]-e=\begin{cases}\frac{\kappa}{\kappa+1}(-\varepsilon^x)+\frac{1}{\kappa+1}\varepsilon^s & d=\hat{s}\\-\varepsilon^x+\frac{1}{\kappa+1}\mathbb{E}[\varepsilon|\varepsilon\geq\iota] & d=\varnothing\end{cases}$$

The remaining steps proving Panel B of Table A.1 are skipped as they are similar to those proving Panel A of Table A.1.

□