Firm Operations, Biodiversity Loss, and Corporate Disclosure
*

Atreya Dey[1]

[1] *University of Edinburgh Business School*

For the latest version, click here.

Abstract

This study investigates whether biodiversity loss is internalized through market pricing or government intervention. Empirical results reveal that emissions from toxic facilities increase lead concentrations and decrease animal populations within a 12-kilometer radius. The US equity market ignores this biodiversity loss, as firms emitting pollutants near biodiverse regions attract no additional risk premium and face economically small penalties from regulatory action. A moral hazard issue likely drives these observations: firms located near ecologically sensitive areas underreport their environmental impacts both in financial disclosures and to regulatory oversight bodies.

Keywords: Biodiversity loss, Equity returns, Corporate disclosure

JEL classification codes: G30, Q57, D83, D62.

*I am grateful for valuable discussions and feedback from Diego Guisande (discussant), Caroline Flammer, Luca Taschini, Glen Gostlow, and Tim Newbold. I thank Baillie Gifford & Co for their generous financial support. The author's email address is `atreya.dey@ed.ac.uk` .

Voluntary or regulatory disclosures are mechanisms for revealing private information about corporate activities, particularly for issues regarding Corporate Social Responsibility (CSR). These mechanisms reduce information asymmetry between firms and market participants, enabling risk pricing and capital allocation. In a Grossman and Hart (1980) world, where information is verifiable and quantifiable, such as the case of carbon emissions (Bolton and Kacperczyk, 2021), firms voluntarily disclose their information in equilibrium. Emerging CSR issues, such as biodiversity loss resulting from firm operations, are challenging to measure due to heterogeneity across species, ecosystems, assessment methodologies, and monitoring frameworks. The unverifiable nature of biodiversity can create perverse incentives for firms to strategically reveal information, distorting disclosures (Christensen et al., 2021) and affecting how information is incorporated into markets (Avramov et al., 2022).

Market participants show increasing concern over a firm's impact on biodiversity, as reflected in the growing demand for disclosure on these effects (TNFD, 2023), which influences how this externality is resolved. Several economic channels explain the incorporation of biodiversity loss into markets. Agents' sustainability preferences may shift asset prices (Pástor et al., 2021). Transition risks arise as regulatory policies prompt firms to internalize externalities (Pastor and Veronesi (2012); Hsu et al. (2023)). Biodiversity degradation may also affect economic productivity and resilience through disrupting ecosystem services (Giglio et al., 2024). Conversely, markets may not capitalize biodiversity information if investors perceive limited impact on expected cash flows.[1] Nonetheless, underlying these mechanisms is the reliance on information, which, when not available through regulatory channels, is accessed through currently limited voluntary disclosures. Figure 1, for example, documents that over 40% of Carbon Disclosure Project (CDP) respondents provide no biodiversity data, hindering assessment of corporate environmental impacts.

This paper asks whether biodiversity loss, directly attributable to a firm's pollutive

[1]Empirical evidence predominantly supports this latter view (Garel et al. (2023); Coqueret et al. (2025)), with risk premia emerging only after the Kunming-Montreal Global Biodiversity Framework in October 2021, increasing the salience of biodiversity risks for investors.

operations, is internalized within the firm through asset prices or government regulation. This novel setting allows an exploration of this emerging issue while evaluating the function of mandatory and voluntary environmental disclosures in addressing information asymmetries.

The study addresses the question in four stages. First, I establish a causal link between toxic emissions and environmental degradation, documenting elevated lead concentrations in water sources within a 9-kilometer radius of new facilities. Second, I demonstrate biological impact through observed animal population declines within a 12-kilometer radius. Third, I assess market pricing of biodiversity risk by comparing returns of firms emitting near protected against unprotected areas, and investigate whether regulatory interventions internalize these externalities. These steps are necessary to finally evaluate firm disclosure of these environmental effects.

I focus on industrial pollution as it represents a meaningful channel for biodiversity degradation and is regulated. The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services ranks pollution as the third most significant driver of biodiversity loss, following land use change and resource extraction (Watson et al., 2019; Jaureguiberry et al., 2022). The U.S. Environmental Protection Agency documents that annual industrial waste releases exceed three billion pounds through the Toxics Release Inventory (TRI) program. These emissions generate biological externalities through multiple channels: reduced genetic diversity (Rillig et al., 2019), aquatic ecosystem degradation Dale (2001), and soil biodiversity loss Rillig et al. (2019). Market and regulatory responses to these externalities are evolving, with increasing emphasis on biodiversity valuation and measurement.[2]

The empirical strategy relies on firm pollution data from the TRI program and exploits variation in facility construction time. Facility construction dates are derived from a 30-meter resolution geospatial land cover database and jump-detection algorithm spanning 1985-2024. The introduction of these "pioneering facilities" into the environment generates plausibly

[2]The Kunming-Montreal Global Biodiversity Framework establishes quantifiable targets for biodiversity preservation (target 1) and risk disclosure (target 15) through 2030. From the article, "Investment sector seeks to put a value on biodiversity", Financial Times December 9th 2024.

exogenous variation in local ecosystem exposure to toxic emissions.

I first establish the spatial dispersion of toxic emissions following facility construction by examining chemical transport to nearby water bodies. This focus on aquatic ecosystems is motivated by their role in supporting biodiversity, with freshwater systems containing 25% of global animal diversity (Dudgeon et al., 2006). I implement a spatial difference-in-differences strategy comparing lead concentrations in proximate versus distant locations from TRI facilities. The estimates reveal that facility construction induces a 4 µg/L increase in mean lead concentration within a 9-kilometer radius relative to measurements at 10-30 kilometers over the subsequent four years. This increment exceeds the EPA's chronic exposure criterion for aquatic life of 2.5 µg/L.[3] These results establish a direct causal link between facility construction and environmental degradation through toxic emissions.

Next, I use the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) database to examine the relationship between animal populations and industrial facilities (Hudson et al., 2017). PREDICTS aggregates biodiversity survey data across sites with varying intensities of anthropogenic land use. I match facility locations with species abundance data using geographic coordinates to estimate the impact of TRI facilities on animal populations at different spatial scales. The empirical results indicate that the introduction of a TRI facility reduces animal populations by 25% of the mean within a 12-kilometer radius relative to populations observed at 13 to 30 kilometers.

The analysis thus far demonstrates that firm facilities plausibly increase toxicity and decrease animal populations in nearby areas. I next examine whether firms internalize their toxic externalities through asset prices or regulatory penalties. I obtain spatial data on biodiversity hotspots from the World Database on Protected Areas (WDPA). The WDPA designates protected areas (PAs) that conserve biodiversity and maintain ecological processes (Elsen et al., 2018). Using spatial boundary data, I identify facilities and firms legally producing toxic emissions within versus beyond a 12-kilometer radius of PAs.

[3]The EPA criterion is available here.

To understand whether the equity market prices this risk, I use Fama-Macbeth regressions (Fama and MacBeth, 1973) to estimate risk premiums associated with emissions proximity to protected areas. The biodiversity footprint measure is constructed as the ratio of emissions within 12-kilometers of protected areas to total firm TRI emissions. The specification controls for the pollution premium (Hsu et al., 2023) through emission intensity. The empirical results indicate no significant risk premium for firms with larger biodiversity footprints near protected regions. Additional tests show these firms do not systematically exceed analyst expectations, suggesting firms do not unexpectedly overperform compared to peers. Moreover, their returns remained unaffected by increased biodiversity attention following the Kunming Declaration (Garel et al., 2023). The evidence suggests the equity market does not price firms' biodiversity impacts.

Another way that externalities can be resolved is through government intervention, e.g., regulation and fines. Using the EPA's Enforcement and Compliance History Online (ECHO) database, I estimate differences in regulatory penalties between facilities proximate to protected areas versus those more distant. The empirical results indicate an economically small though statistically significant fine differential of around $2,500, suggesting minimal regulatory action or potential legal constraints on enforcement for firms near PAs (Bellon, 2021).

I next explore the mechanisms underlying the non-pricing of biodiversity externalities, focusing on disclosure patterns. Using ECHO data, I find that facilities near protected areas (PAs) self-report fewer violations compared to those not near such regions. To evaluate potential underreporting, I use biodiversity risk scores from Giglio et al. (2021), which quantify firms' biodiversity risk disclosures in their 10-K filings. The empirical results show no relationship between toxic emissions near biodiverse regions and firms' self-reported risk measures. In fact, emissions intensity correlates with these measures, suggesting challenges in differentiating between carbon emissions and biodiversity impacts in financial reports. Analysis of CDP biodiversity questionnaire responses yields similar non-correlation with firms' biodiversity footprints.

The evidence collectively implies systematic underreporting of operational biodiversity impacts by firms with larger toxic footprints. Because investors use disclosures as a major channel to incorporate information, this practice may explain the lack of evidence for pricing this risk. Indeed, the suggestive evidence that facilities tend to not self-report their violations to regulatory authorities raises questions about the usefulness of voluntary and financial disclosures in assessing difficult to measure CSR activities. The voluntary disclosures could amount to "cheap talk" (Crawford and Sobel (1982); Bingler et al. (2022); Gostlow (2020)) and cause moral hazard problems, as there is little incentive for firms to reduce information asymmetries given the difficulties in measuring biodiversity loss.

As this an emerging topic, there are few other articles that measure the biodiversity footprint of firms to understand whether this loss is priced. I contribute to the literature examining the relationship between asset prices and the operational impacts of firms on the broader environment. A related article by Hsu et al. (2023) uncovers a premium for firms that generate high toxic emission intensity in comparison to those with lower intensity. This paper uses similar techniques; however, the primary difference is in identifying the effects of emissions in the context of biodiversity. Another paper by Garel et al. (2023) uses a more general measure provided by an external data provider to assess firm impact on biodiversity loss. While their results demonstrate that corporate biodiversity footprint is not priced in the cross-section of stock returns, they illustrate a relationship between attention and firms with greater footprints. This paper instead directly links biodiversity loss to firm operations and uncovers little evidence that the market or governments resolve these losses.

This study also adds to the body of work on CSR disclosure. Ilhan et al. (2023) conduct a survey of institutional investors, revealing a consensus that inadequate reporting of climate-related risks results in asset mispricing. Matsumura et al. (2014) examine S&P 500 companies, investigating the impact of carbon emissions on firm valuation and the influence of voluntary public emissions disclosure via the CDP on firm value. Here, I explain inadequate pricing through a channel of underdisclosure and demonstrate the presence of moral hazard, even in the presence of trusted third-party mediators like the government or the CDP (Ewerhart and Lareida, 2024).

More generally, this analysis extends the emerging biodiversity-finance literature surveyed by Karolyi and Tobin-de la Puente (2023). Flammer et al. (2023) analyze private capital deployment in biodiversity conservation through blended financing structures. Giglio et al. (2023) construct biodiversity risk measures from firm disclosures and find evidence of risk pricing.

1 Data

The data used in this paper are drawn from a variety of sources. I first describe the environmental data used and then the firm level accounting and return data.

1.1 Environmental and Spatial Data

1.1.1 Water Quality Portal

I obtain a historical time series water quality information of concentrations of lead from the Water Quality Portal (WQP). The WQP is a public platform where diverse groups such as federal, state, tribal and local governments, academia, non-governmental organizations, and private individuals submit project information and sampling results.

Due to the variety of information the WQP collects, the reporting for each chemical is not fully standardized. I therefore keep all observations that are either measured in nanograms or micrograms per liter and transform measurements in nanograms to micrograms. I then aggregate the observations per station to a yearly frequency by taking the mean of the measurements. Furthermore, due to variations in monitoring periods, I select observations part of "regular screening" at monitoring sites with more than 10 observations. I also remove observations that are related to industrial releases such as "industrial effluent", "industrial waste", "wastewater treatment plant effluent", and "waste water". Including these observations would likely bias estimates as they might only occur due to the facility itself.

1.1.2 animal Population

I collect data on animal populations from the Projecting Responses of Ecological Diversity In Changing Terrestrial Systems (PREDICTS) database (Hudson et al., 2017). This largely cross-sectional database, curated by the Natural History Museum of London, compiles data from 480 published scientific studies and government-monitored populations to estimate animal populations across various regions. For a survey to be included, it must describe species abundance or ecological assemblage diversity. The data must come from published work or use formally documented sampling methods. Consistent sampling methods are applied across sites per study, though varied effort levels are accepted if documented. Geographic coordinates and land-use classifications accompany each sampling site.

Sampling effort influences measurements in the study. Thus, for the 1% of records dependent on effort-sensitive metrics, where sampling effort varies, adjustments are necessary. I standardize sampling effort in each study by assigning a relative effort value of 1 to the most-sampled site and adjusting the efforts of other sites proportionately.[4]

1.1.3 Protected Areas

As a proxy for biodiverse regions, I use the World Database on Protected Areas (WDPA) because protected areas are known to contain more species and individuals than from unprotected areas (Gray et al., 2016). The database has been put forth by the UN Environment Programme and the International Union for Conservation of Nature (IUCN) to denote information on global marine and land protected regions across the world. It contains spatial information on the geographical area which has been recognized by either the IUCN or the Convention on Biological Diversity (CBD). It includes areas such as: nature reserves,

[4]The code used to standardize the measurements are available from Tim Newbold.

wilderness areas, national park, natural monuments, managed habitats and species, protected landscapes or seascapes, and managed resources. Overall, this database is meant to proxy for biodiversity hotspots or havens. An image of the protected areas in the United States is displayed in Figure 3.

I focus on IUCN-designated protected areas with classifications Ia, Ib, II, III, IV, and V based on their land management goals. Specifically, Category Ia represents a strict nature reserve, Category Ib a wilderness area, Category II a national park, Category III a natural monument, Category IV a habitat or species management area, and Category V a protected landscape or seascape.

1.1.4 Toxic Release Inventory and ECHO

I use the publicly available Toxic Release Inventory (TRI) database, managed by the US Environmental Protection Agency (EPA), to identify industrial facilities emitting toxic pollutants. The TRI database was established in 1986 through the Emergency Planning, Community Right to Know Act (EPCRA), which was prompted by the Bhopal disaster and subsequent incidents of chemical spills at American Union Carbide plants. Under the mandate, manufacturing plants (classified under Standard Industrial Classifications 2000 to 3999) with over 9 full-time employees and exceeding specified thresholds for the use or production of toxic substances are obligated to report their releases to the EPA. The TRI dataset specifically incorporates various parameters such as the report year, quantity of chemical pollutants in pounds, chemical category identifiers, geographical location, and respective corporate identifications. A spatial map of the facilities located in the contiguous United States is presented in Figure 2.

The total log emissions by all TRI facilities is presented as the blue line in Figure 4. The sharp rise of emissions in 1998 through 2000 is marked by inclusion of persistent bioaccumalitve toxins and seven industry sectors: metal and coal mining facilities, electric power generators, commercial hazardous waste treatment operations, solvent recovery facilities, petroleum bulk terminals and wholesale chemical distributors. The dashed red line represents the waste emitted within 12 km of a protected area.

I also draw on the EPA's Enforcement and Compliance History Online (ECHO) database, which integrates facility-level records of compliance with U.S. environmental laws into a comprehensive record for each facility. The ECHO Exporter consolidates various types of information: it combines descriptive details such as the facility's name, address, and industry classification; aggregates five years of compliance history; and records information on the penalties issued to the facility for noncompliance.

1.1.5 Landcover and Pioneers

The first set of empirical tests rely on newly constructed TRI facilities to estimate their effects on lead concentration and biodiversity loss. To estimate the construction date of a facility, I apply a widely used spatial land cover data and a simple jump-detection algorithm. Specifically, I use the Annual National Land Cover Database (NLCD) impervious surface product which provides the fractional or percent surface area of a 30-meter map pixel that is covered with processed materials or structures (pavement, concrete, rooftops, and other constructed materials) available from 1985 to 2023 for the contiguous US.

I calculate the total impervious fraction of a 500 meter circle surrounding the coordinates given by the TRI dataset. Figure 5 shows the log growth in the fractional imperviousness of the 500 meter region in the years surrounding the first report to the TRI at year 0. Overall, the growth in man-made surfaces drops sharply after the first reporting date which adds validity to using this spatial data for this analysis. There appears to be a fairly consistent growth in surface area before the first reporting further showing that naively using the first reporting date of the facility may lead to biased estimates.

To estimate the first construction year of the facility, I use a simple jump-detection algorithm.[5] Specifically, I classify the first two standard deviation increase above the mean,

[5] Jump detection algorithms have been used in the economics and financial literature to detect changes in economic activity (van der Kroft et al. (2024); Lee and Mykland (2008)).

0.473, before the first reporting date as the beginning construction year. Figure 6 presents the difference between the first reporting year and the estimated construction year using this jump detection algorithm. The figure shows that the difference is contained around 2 years before the first reporting date and linearly decreases. Figure 7 presents graphically the growth in built-up area after using the jump-detection method to move the reporting date to the estimated construction year. Overall, after comparing Figures 5 and 7, the method appears to effectively capture a sharp rise in construction efforts in the 500 m region.

The TRI facility filtering process is as follows. The TRI dataset in total contains information on 63,062 facilities from 1987 to 2023. After using the jump-detection algorithm, 6,302 facilities remain in the sample which I call "pioneers". The majority of the removed facilities have no satellite based observable changes in built up area within the 500 m region, suggesting a long-standing facility. For comparison to other literature, Currie et al. (2015) use 1,600 TRI plant openings and closings.

1.2 Financial Data

The sample of firms used in this are those in the intersection of three databases: Compustat, the Center for Research in Security Prices (CRSP), and the TRI database, following the approach of Xiong and Png (2019) and Hsu et al. (2015). I sourced accounting data from Compustat and gathered stock price information from CRSP. The selected companies for our study are those that have complete TRI data, intact standard industrial classification (SIC) codes, and domestic common shares (symbolized as SHRCD = 10 and 11) on public exchanges such as NYSE, AMEX, or NASDAQ. I exclude financial companies with four-digit SIC codes between 6000 and 6999, which encompass sectors such as finance, insurance, trusts, and real estate. The final dataset includes 1,642 firms that have unbalanced return data from October 1987 to December 2023.

2 Environmental Effects

2.1 Lead releases

I first identify whether the introduction of TRI facilities significantly increases pollutants into their environments. Identification of the treatment effect comes with issues as it is difficult cleanly estimate the effects of the TRI facility while satisfying the only through condition. In this case, other industrial pollutants and anthropogenic reasons can add to the pollutants in the area. Furthermore, TRI facilities only have to report to the EPA if the facility hires more than 9 full time employees which makes it difficult to establish the true treatment effect of the facility.[6] Moreover, there is evidence of misreporting of total chemical releases (Currie et al., 2015). To alleviate concerns of endogeneity, I use shock based inference to show that the introduction of these facilities does plausibly increase lead concentration in their surrounding area.

To identify a shock to local pollutants, I estimate the year that a TRI facility is constructed using spatial land cover data and a jump-detection algorithm outlined in Section 1.1.5. I then estimate the relationship between the introduction of these facilities and detectable lead in the water bodies in its proximity. The primary empirical method uses a difference-in-differences strategy to compare chemical measurements in water sources near facilities to those marginally farther away.

I focus on water pollution for two reasons. First, while Currie et al. (2015) provides an empirically defined bound for *air* emissions for the eight pollutants, much less is known regarding the transport and geographic extent of influence of chemicals released into water or land. Substantial evidence suggests that chemical deposits in air, water, or land can travel significantly farther. For example, Svendsen et al. (2007) found high concentrations of heavy metal deposits up to 10 km away from an industrial zinc smelter. Other evidence from Burney (2020) estimates negative effects on crop yields due to air pollution up to 25

[6]Currie et al. (2015) bypasses this limitation by using the opening date from Census Bureau data.

km away from coal plants.[7] This large variation in distances makes it difficult to establish a prior when considering the distance traveled by pollutants.

Second, aquatic ecosystems sustain a substantial fraction of global biodiversity. Freshwater habitats, though constituting less than 1% of the Earth's surface, house almost 30% of vertebrate species, including 51% of known fish species (WWF, 2020). These ecosystems also hold intrinsic value for the environment and are critical for human life, livelihoods, cultures, and economies (Declaration, 2007). Freshwater environments, alongside inland and coastal wetlands, deliver more ecosystem services compared to open oceans, woodlands, grasslands, and temperate forests (Farmer, 2012).

I generate a panel dataset by linking lead monitoring points to the construction date of nearby industrial facilities. This involves matching the longitude and latitude of each monitoring location with those of pioneering industrial facilities within specific distance thresholds. By spatially correlating these points, I assess the relationship between industrial emissions and water quality across different distances.

I use a differences-in-differences strategy that compares the chemicals detected by a monitoring site that has nearby toxic facilities with those slightly farther away. For example, a dummy variable, $Dist$, is created, which equals 1 if a monitoring site (m) is within 500m to 6 km of a facility (f) and 0 if the nearest facility is 7 to 30 km away. This slight variation in area is important, as these differentiated regions should be relatively similar, thus minimizing unwanted bias in the estimation. To estimate the relationship between distance of monitoring stations and chemical releases, I perform various regressions in the style of the following linear model:

$$Y_{m,f,t} = \beta_0 + \beta_1 \omega_{m,f,t} + \beta_2 Const_{f,t} + \beta_3 \mathrm{Dist}_m + \beta_4 (Const_{f,t} \times \mathrm{Dist}_m) + \eta_{m,f} + \tau_t + \varepsilon_{m,f,t}, \quad (1)$$

where $Y_{m,f,t}$ is the outcome variable representing lead levels in micrograms per liter (µg/L)

[7] More evidence from the National Oceanic Atmospheric Administration shows that air pollutants can potentially travel hundreds of miles and be deposited onto various surfaces.https://oceanservice.noaa.gov/education/tutorial_pollution/07input.html

detected at monitoring site, m, near facility, f, for year t. $Dist_m$ serves as a dummy variable distinguishing monitoring stations based on the proximity of their nearest facility within different distance thresholds. To more cleanly estimate the effects of the TRI facility on lead emissions, I focus on the years surrounding the first reporting date of the facility. $Const_{f,t}$ is therefore a dummy variable equal to one if it is the first, second, third, fourth, or fifth year after the construction date identified using the jump-detection algorithm, [0,4], and zero if between two years to five years prior [-2,-5]. I also include monitoring-facility pair and year fixed effects denoted by $\eta_{m,f}$ and τ_t to guarantee that the variation comes from within these pairs around the construction of the facility. The standard errors are two-way clustered on the facility and monitoring station.

In this linear model, β_2 represents whether there is a significant difference between between the measured chemicals before and after the facility is introduced. The coefficient of interest, β_4, estimates the difference in the effect of the TRI facility on on chemical levels between monitoring sites that are near a facility and those that are farther away. If the coefficient is positively significant, this suggests that facilities close to monitoring sites show a larger increase in chemical detections following a TRI facility entering the location compared to those farther away.

The results of the regressions are presented in Table 2 and Figure 8 for various distance pairings. The plot shows a positive and significant coefficient for β_4 from specification 2 between 5 km and 10 km of a TRI facility. Three years following a TRI facility's construction date, the mean concentration of lead within a 9-kilometer radius increases by 4 µg/L compared to levels measured within a 11-30 kilometer radius. For comparison, the chronic exposure limit for aquatic life recommended by the EPA is 2.5 µg/L.

However, there is evidence that these estimates from standard two-way fixed effects models may be biased as there are are more than two time periods and monitoring stations are treated at different points in time (Roth et al., 2023). To addresses this issue with staggered differences-in-differences, I use a stacked difference-in-difference strategy for the distance pairing of [8, 9-30] km (Cengiz et al., 2019). Here the base period is considered the year before the construction date identified by the jump-detection algorithm. Specifically, all observations are aligned with the event-time of construction and not-yet-treated observations act as a control set in each cohort.

Similar to the prior DiD estimation, the introduction of the facility is related to a significant increase in lead concentrations of slightly less than 4 µg/L within the 8 km range (Figure and Table). The increase is persistent up to five years after the construction date. Capturing this connection between the operations of toxic plants and the levels of chemical pollutants in nearby areas acts to confirm that pollutants are indeed disseminating into the local environment. This extends the results of Currie et al. (2015) who find that air emissions are particularly worsened in a one mile radius to toxic facilities.

2.2 Animal population

To complement the analysis on chemical releases, I conduct a similar empirical exercise associating surveyed animal populations to the construction of nearby TRI facilities. Soils and water containing these contaminants have the potential to be absorbed by plants and animals. These chemicals include endocrine disruptors which are known to have adverse effects on the developmental and reproductive ability of animals (Schantz and Widholm, 2001).

One method to measure animal populations is the PREDICTS database developed by Hudson et al. (2017). This holds records of thousands of monitored animal populations worldwide from 1996 to 2022 and is useful to understand species abundance and occurrence in different regions.

To estimate the relationship between animal populations and pioneering facilities, I employ a similar DiD strategy used in Section 2.1 that leverages marginal differences in the proximity of monitored populations to these facilities. Using various distance pairings, I create a dummy variable, $Dist$, which equals 0 if survey site i is closer to a toxic facility and 1 if the nearest facility is slightly further away. I then perform regressions following this specification:

$$\begin{aligned} Population_{m,f,t} = \beta_0 + \beta_1 \omega_{m,f,t} + \beta_2 Construct_{f,t} \\ + \beta_3 \text{Dist}_m + \beta_4 (Construct_{f,t} \times \text{Dist}_m) + \eta_f + \theta_s + \tau_t + \varepsilon_{m,f,t}, \end{aligned} \tag{2}$$

where the dependent variable is the number animals surveyed in a site, m, at year, t. $Dist_m$ serves as a dummy variable distinguishing survey sites based on the proximity of their nearest facility within different distance thresholds. I focus on the years surrounding the construction date of the facility. $Const_{f,t}$ is therefore a dummy variable equal to one if it is the first, second, third, fourth or fifth year of the construction date identified using the jump-detection algorithm, [0,4], and zero if between two years to five years prior [-2,-5]. τ_t is a trend variable which captures the general decrease animal populations due to land loss and other factors. I also include fixed effects for the facility, η_f and study, θ_s. Including facility-site pair fixed effects would subsume the majority of the data as the sites in each study in the PREDICTS data are largely cross-sectional rather than time series. I also include other fixed-effects such as land-use intensity and ecoregion of the site. The standard errors are clustered are the facility-year level.

I explore the interaction term for various distance pairings in Figure 10 and Table 4. The standard errors are large in the first distance pairings due to the imbalance of treated and untreated observations of biodiversity sites within 6 km of facilities.[8] Generally, I show that animal populations decline by 20 in the five years after a TRI facility is constructed. To put this into perspective, the average species surveyed are 55.48, therefore reducing species by 36% of the mean or 25% of the variance. Overall, the consequences of these facilities is evident—not only do these chemicals persist in the environment but they substantially reduce biodiversity in nearby locations.

I also conduct a stacked difference-in-difference strategy for the distance pairing of [8, 9-

[8]There are 540 treated observations within a 500 m to 6 km range of facilities, while 9,875 untreated observations within 7 km - 30 km of facilities.

30] km as a robustness check. The base period is considered the year before the construction date identified by the jump-detection algorithm. Results are presented in graphical form in Figure 11 and Table 5 which show lower counts of animals in the years after construction.

3 Is the biodiversity footprint internalized?

The prior sections provided evidence of lead emissions being released within a 10 km radius and species counts declining significantly in around a 12 km radius. Using this range, I construct a measure of firm level biodiversity footprint by first indicating the facilities that are withing a 12 km radius of protected areas, which are known to have greater biodiversity than other regions. Then, I aggregate all toxic emissions, either nearby protected areas or not, to the facility level. I link these facilities to the ultimate owner (Xiong and Png, 2019) to obtain a time series of legal emissions emitted either nearby protected or unprotected areas by each firm. Then I calculate the square root of total emissions nearby PAs divided by the total emissions of the firm to obtain a firm's biodiversity footprint.

Emissions measures, i.e., the biodiversity footprint and emissions intensity, are available in a lagged fashion at the firm level in the fourth quarter of each calendar year for the previous calendar year. For example, a firm reporting their emissions to the EPA in October of 2016 would be reporting the emissions from each of their facilities between January 2014 through December 2015. Therefore, I use the information as it is available to the broader market from the quarter it is released to the third quarter of the next calendar year, akin to a predictive exercise. Specifically, for a calendar quarter frequency, the biodiversity footprint is available for firm i for all of its facilities in the beginning of quarter four, t_{Q4}, as follows,

$$\text{Biodiversity Footprint}_{i,[t_{\mathrm{Q4}},t_{\mathrm{Q4}}+3]} = \sqrt{\frac{\text{Emissions Nearby Protected Areas}_{i,t_{\mathrm{Q4}}}}{\text{Total Emissions by Firm}_{i,t_{\mathrm{Q4}}}}}, \tag{3}$$

until until the next three quarters $t_{\text{Q4}+3}$. Similarly, emissions intensity,

$$\text{Emissions Intensity}_{i,[t_{\text{Q4}},t_{\text{Q4}}+3]} = \sqrt{\frac{\text{Total Emissions}_{i,t_{\text{Q4}}}}{\text{Total Assets}_{i,t_{\text{Q4}}}}}, \tag{4}$$

is calculated as the total emissions across all facilities owned by firm i in t_{Q4} divided by the total assets of the firm reported at the beginning of t_{Q4}.

Figure 4 presents the log waste released nearby protected areas and the total waste released. In Appendix A, I show that the correlation between this "biodiversity footprint" and emissions intensity used in prior studies (Hsu et al., 2023) is low, differentiating this from prior work. I use this biodiversity footprint measure to understand whether it is priced in the equity market or whether the externality is internalized through regulation or taxation. Generally, I find little empirical evidence of this loss being incorporated.

3.1 Risk premiums

I estimate risk premiums using two-pass regressions (Fama and MacBeth, 1973). This involves first estimating beta exposure to firm characteristics and Fama-French factors, then estimating premiums using cross-sectional regressions of returns on betas each month. Returns are recorded on a monthly frequency and firm characteristics at a quarterly frequency with a lag of one quarter. The Biodiversity Footprint and Emissions Intensity are available at the beginning of the fourth quarter for the following year.

Table 6 produces the estimates from two separate Fama-Macbeth regressions. The first column excludes industry fixed effects while the second column includes Fama-French 10 industry fixed effects. I find that neither the biodiversity footprint nor the emissions intensity measures produce statistically significant risk premiums when using equally-weighted returns. This finding is in line with Garel et al. (2023) who show that a systematic risk premium does not exist for firms with large biodiversity footprints according to the Iceberg Data Labs.

3.2 Attention to biodiversity

To test whether greater attention towards biodiversity risks triggers changes in returns for those firms with greater footprints, I use an event study methodology around the Kunming Declaration, an event previously found to affect abnormal returns (Garel et al., 2023).

Using the day of the adoption of the Kunming Declaration, October 13, 2021, I calculate abnormal returns defined as the returns above each firm's beta exposures to the market, small-minus-big, and high-minus-low, portfolios estimated over a window of [-531, -31] days prior to the event. I perform panel regressions of the following specification,

$$\text{Daily abnormal return }_{i,t} = \beta_0 + \beta_1 \text{Biodiversity Footprint}_i \times \ \text{Post}_t + \eta + \gamma_t + \varepsilon_{i,t}, \quad (5)$$

where I regress daily abnormal returns on an interaction term between the biodiversity footprint and an indicator variable *Post*. The indicator is equal to 1 if denoting the days after the event, inclusive of the event and zero if before the event. I include Fama-French 10 industry fixed effects, η, day fixed effects, γ_t and cluster at the firm-level. The standalone post event fixed effect is therefore subsumed by the day fixed effect. β_1 represents the difference in abnormal returns for firms with a greater biodiversity footprint.

Table 7 presents regression results across various time thresholds from 7 to 3 days surrounding the event. The results show that the returns for firms that produce all of their emissions nearby biodiverse areas increase by 30 basis points surrounding the Kunming Declaration. Nonetheless, there appears to be little statistical significance of these estimates as only using the days, [-4, +3], produce a coefficient significant at the 10 percent level. Interestingly, the coefficient on the interaction term is positive, the opposite of the relationship found in Garel et al. (2023).

3.3 Government intervention

In this section, I identify whether this externality may be internalized through government intervention by way of federal enforcement cases through the EPA. I use the ICIS-FE&C database, which offers information on federal administrative and judicial cases related to various environmental statutes. These include the Clean Air Act (CAA), Clean Water Act (CWA), Resource Conservation and Recovery Act (RCRA), Safe Drinking Water Act (SDWA), Section 313 of the Emergency Planning and Community Right-to-Know Act (EPCRA), Toxic Substances Control Act (TSCA), Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA or Superfund), and the Marine Protection, Research, and Sanctuaries Act (MPRSA). The database contains information on the cases put forth by the EPA at the facility level, the resulting penalties, compliance costs, and recovery amounts awarded to the EPA. I merge this database with the TRI database.

I aim to determine if facilities emitting near protected areas incur different costs when facing enforcement cases. I calculate a cost variable that aggregates: the total federal penalty assessed or agreed upon, the total cost of all Supplemental Environmental Projects (SEPs) for the case, the dollar value sum of compliance action amounts, and the cost recovery amount ordered or agreed to be repaid to the U.S. EPA, due to the Superfund per administrative or judicial settlement. This variable is conditional on a case being made against the facility. The square root of the *Costs* variable is then used in a specification as follows:

$$\text{Costs}_{i,t} = \beta_0 + \beta_1 \text{Protected} + \beta_2 \text{Releases} + \beta_3 \text{Releases}_i \times \text{ Protected}_t + \gamma X_{i,t} + \eta + \rho + \varepsilon_{i,t}, \quad (6)$$

where *Releases* are the square root of the total chemical releases of the facility and *Protected* is an indicator variable if the facility is 12 km away from a protected area. $X_{i,t}$ represents 2010 U.S. Census variables such as number of houses, minority population, and population density within a 5 km radius of the facility. η represents industry by year fixed effects and ρ

are state fixed effects. Standard errors are clustered at the facility level.

Table 8 presents regression results with varying use of fixed effects and controls. Across the three columns, greater releases are associated with higher costs. The coefficients in the second row suggests that facilities near protected areas incur slightly larger costs of slightly below \$35,000. However, there is a negative coefficient on the interaction term, indicating that greater releases from a facility nearby a protected region is negatively related to costs. This suggests a size effect in that greater releases by larger firms may push down the costs they incur through legal channels. For an average facility that releases 313[2] unitless emissions into the environment, the additional fines are slightly below \$2,500. Overall, the costs associated with releasing emissions nearby biodiverse regions appears economically minimal.

4 Disclosures

Next, I attempt to understand why the biodiversity footprint is not resolved either through equity markets or government intervention. I focus on a reporting channel as firms incentives to voluntarily report their violations may be low as evidence indicates that agents will often conceal information detrimental to their performance. In essence, a moral hazard problem can arise if the firm has incentives to misrepresent or underreport their negative environmental impacts.

4.1 Reporting to government

I seek to understand whether there these facilities voluntarily disclose their violations to the EPA. The ICIS-FE&C database contains information on whether the enforcement action was the result of self-disclosure by the facility. After conditioning on a facility having a case with the EPA, I calculate the number of enforcement actions that a facility has per year. Then, I calculate *Disclosures* as the ratio of voluntary disclosures to total enforcement actions taken over the same year. If *Disclosures* is equal to 1, this would mean that all enforcement actions would be a result of self-disclosures from he facility.

Using *Disclosures*, I define a linear probability specification of the following form:

$$\begin{aligned} \text{Disclosure}_{i,t} &= \beta_0 + \beta_1 \text{Protected} + \beta_2 \text{Releases} \\ &+ \beta_3 \text{Releases}_i \times \ \text{Protected}_t + \gamma X_{i,t} + \eta + \rho + \varepsilon_{i,t}, \end{aligned} \tag{7}$$

where *Releases* are the square root of the total chemical releases of the facility and *Protected* is an indicator variable if the facility is 12 km away from a protected area. $X_{i,t}$ represents 2010 U.S. Census variables such as number of houses, minority population, and population density within a 5 km radius of the facility. η represents industry by year fixed effects and ρ are state fixed effects. Standard errors are clustered at the facility level.

Table 9 presents results using varying fixed effects at the facility level. The dependent variable, bounded between 0 and 1, reflects this likelihood. The results indicate a significant negative relationship between the square root of pollutant releases and self-reporting behavior. Furthermore, facilities situated near protected areas are less likely to self-report violations, with an associated coefficient of 3.1% decrease in likelihood compared to facilities not nearby such areas. Interestingly, the interaction term between protected status and releases is positive, highlighting that the adverse effect of releases on reporting probability is mitigated when a facility is near a protected area.

The empirical results reveal a nuanced landscape of moral hazard in environmental self-reporting practices, particularly in the context of facilities proximal to protected, biodiverse areas. The analysis suggests that facilities releasing a smaller number of emissions exhibit greater moral hazard, as they are less likely to self-report violations. However, the interaction term suggests that as a facility's emissions increase, potentially signaling a more significant environmental catastrophe, the likelihood of self-reporting heightens. Thus, larger emissions, which could attract greater regulatory scrutiny and public concern, compel facilities to disclose violations despite the moral hazard influences present in protected regions.

4.2 Reporting in financial statements

I use the self-reported scores developed by Giglio et al. (2023) to measure biodiversity risk of the firms in my sample. They use regular expression searches to extract biodiversity-related sentences from 10-K statements with a biodiversity dictionary. They develop three different scores: an aggregate count that combines biodiversity opportunities and risk, a score to capture when a firm expresses a concern in the 10-K about regulations, and a sentiment based measure that captures negative sentiment or rising risk. The first two scores are binary variables indicating whether biodiveristy or regulation risk is discussed at all and the third is a discrete variable which is a summation of the number of positive and negative sentiments measured. As an added caveat, the percent of mentions of biodiversity in 10-Ks as measured by Giglio et al. (2023), is very low. Specifically, the aggregate and regulation measures account for only 4.25% and 3.04% of occurrences in my sample, respectively.

Using the three scores as dependent variables, I perform ordinary least squares regressions with firm and year fixed effects with the percent of total emissions in protected areas as the dependent variable. The first two columns of Table 10 are linear probability models with the third being simple OLS regression. For the first column, a one-standard deviation increase in the percent of emissions significantly decreases the number of mentions of biodiversity in 10-Ks by 0.02. The results of the second column illustrate that a similar increase in emissions in biodiverse areas leads to 0.01 fewer mentions of biodiversity regulation for firms. Finally, I find no significant relationship between negative mentions of biodiversity risk and emissions.

4.3 Voluntary disclosures

As an additional exercise, I collect data from the CDP's 2023 report on biodiversity. I extract responses to the question, "Does your organization have activities located in or near biodiversity-sensitive areas during the reporting year?". Specifically, I create a binary variable which is equal to 1 if the company reports that it has operations near biodiversity-sensitive areas and 0 if it reports that it does not or does not assess it. I regress this binary variable on the biodiversity footprint, emissions intensity, and lagged quarterly firm level data either in 2020, 2021, 2022, or 2023. Table 11 presents the results of these linear probability models. Overall, I find no relationship between toxic footprints or emissions intensity on reporting if a firm has activities nearby biodiverse sensitive areas.

These results suggest that these polluting firms do not self-disclose the impact of their emissions footprint on biodiversity in their 10-K filings. This finding relates to prior literature that find that 10-K disclosures can be an unreliable measure of material risks. Bingler et al. (2022) find that climate disclosures are mostly “cheap talk” and Gostlow (2020) who shows that Form 8-K filings are better predictors of firm performance during the COVID-19 crisis than disclosures in 10-Ks. Taking this interpretation into account, the market could be underpricing this biodiversity risk due to the lack of clarity on emission footprints or moral hazard.

5 Conclusions

This study provides evidence that chemical emissions detrimentally affect the environment, directly causing biodiversity loss. The analysis reveals that the US equity market fails to price this externality, while facilities located near biodiverse regions incur marginally higher regulatory penalties compared to others. The consistent market response likely stems from inadequate disclosure of environmental impacts through governmental channels, voluntary disclosures, or financial reports. Specifically, suggestive evidence indicates that facilities near biodiverse regions often do not self-report violations, raising concerns of moral hazard.

References

Avramov, Doron, Si Cheng, Abraham Lioui, and Andrea Tarelli, "Sustainable investing with ESG rating uncertainty," *Journal of financial economics*, 2022, *145* (2), 642–664.

Bellon, Aymeric, "Fresh start or fresh water: The impact of environmental lender liability," Technical Report, Working paper 2021.

Bingler, Julia Anna, Mathias Kraus, Markus Leippold, and Nicolas Webersinke, "Cheap talk and cherry-picking: What ClimateBert has to say on corporate climate risk disclosures," *Finance Research Letters*, 2022, *47*, 102776.

Bolton, Patrick and Marcin T Kacperczyk, "Carbon disclosure and the cost of capital," *Available at SSRN 3755613*, 2021.

Burney, Jennifer A, "The downstream air pollution impacts of the transition from coal to natural gas in the United States," *Nature Sustainability*, 2020, *3* (2), 152–160.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Christensen, Hans B, Luzi Hail, and Christian Leuz, "Mandatory CSR and sustainability reporting: Economic analysis and literature review," *Review of accounting studies*, 2021, *26* (3), 1176–1248.

Coqueret, Guillaume, Thomas Giroux, and Olivier David Zerbib, "The biodiversity premium," *Ecological Economics*, 2025, *228*, 108435.

Crawford, Vincent P and Joel Sobel, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1431–1451.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker, "Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings," *American Economic Review*, 2015, *105* (2), 678–709.

Dale, Barrie, "Marine dinoflagellate cysts as indicators of eutrophication and industrial pollution: a discussion," *Science of the total environment*, 2001, *264* (3), 235–240.

Declaration, Brisbane, "Environmental Flows are Essential for Freshwater Ecosystem Health and Human Well-Being. Declaration of the 10th International River Symposium and International Environmental Flows Conference, 3–6 September 2007: Brisbane, Australia," 2007.

Dudgeon, David, Angela H Arthington, Mark O Gessner, Zen-Ichiro Kawabata, Duncan J Knowler, Christian Lévêque, Robert J Naiman, Anne-Hélène Prieur-Richard, Doris Soto, Melanie LJ Stiassny et al., "Freshwater biodiversity: importance, threats, status and conservation challenges," *Biological reviews*, 2006, *81* (2), 163–182.

Elsen, Paul R, William B Monahan, and Adina M Merenlender, "Reply to You et al.: The World Database on Protected Areas is an invaluable resource for global conservation assessments and planning," *Proceedings of the National Academy of Sciences*, 2018, *115* (39), E9029–E9030.

Ewerhart, Christian and Julia Lareida, "Voluntary disclosure in asymmetric contests," *Review of Economic Studies*, 2024, p. rdae001.

Fama, Eugene F and James D MacBeth, "Risk, return, and equilibrium: Empirical tests," *Journal of political economy*, 1973, *81* (3), 607–636.

_ **and Kenneth R French**, "A five-factor asset pricing model," *Journal of financial economics*, 2015, *116* (1), 1–22.

Farmer, Andrew, "The economics of ecosystems and biodiversity for water and wetlands," *Institute for European Environmental Policy (IEEP). https://www. ramsar. org/sites/default/files/documents/pdf/cop11/ppt/cop11-ppt-192-farmer. pdf*, 2012.

Flammer, Caroline, Thomas Giroux, and Geoffrey Heal, "Biodiversity finance," Technical Report, National Bureau of Economic Research 2023.

Garel, Alexandre, Arthur Romec, Zacharias Sautner, and Alexander F Wagner, "Do Investors Care About Biodiversity?," *Swiss Finance Institute Research Paper*, 2023, (23-24).

Giglio, Stefano, Matteo Maggiori, Krishna Rao, Johannes Stroebel, and Andreas Weber, "Climate change and long-run discount rates: Evidence from real estate," *The Review of Financial Studies*, 2021, *34* (8), 3527–3571.

_ **, Theresa Kuchler, Johannes Stroebel, and Olivier Wang**, "The economics of biodiversity loss," Technical Report, National Bureau of Economic Research 2024.

_ **,** _ **,** _ **, and Xuran Zeng**, "Biodiversity Risk," Technical Report, National Bureau of Economic Research 2023.

Gostlow, Glen, "The materiality and measurement of physical climate risk: evidence from form 8-K," 2020.

Gray, Claudia L, Samantha LL Hill, Tim Newbold, Lawrence N Hudson, Luca Börger, Sara Contu, Andrew J Hoskins, Simon Ferrier, Andy Purvis, and Jörn PW Scharlemann, "Local biodiversity is higher inside than outside terrestrial protected areas worldwide," *Nature communications*, 2016, *7* (1), 12306.

Grossman, Sanford J and Oliver D Hart, "Disclosure laws and takeover bids," *The Journal of Finance*, 1980, *35* (2), 323–334.

Hsu, Angel, Andrew S Moffat, Amy J Weinfurter, and Jason D Schwartz, "Towards a new climate diplomacy," *Nature Climate Change*, 2015, *5* (6), 501–503.

hsuan Hsu, Po, Kai Li, and Chi yang Tsou, "The pollution premium," *The Journal of Finance*, 2023, *78* (3), 1343–1392.

Hudson, Lawrence N, Tim Newbold, Sara Contu, Samantha LL Hill, Igor Lysenko, Adriana De Palma, Helen RP Phillips, Tamera I Alhusseini, Felicity E Bedford, Dominic J Bennett et al., "The database of the PREDICTS (projecting responses of ecological diversity in changing terrestrial systems) project," *Ecology and evolution*, 2017, *7* (1), 145–188.

Ilhan, Emirhan, Philipp Krueger, Zacharias Sautner, and Laura T Starks, "Climate risk disclosure and institutional investors," *The Review of Financial Studies*, 2023, *36* (7), 2617–2650.

Jaureguiberry, Pedro, Nicolas Titeux, Martin Wiemers, Diana E Bowler, Luca Coscieme, Abigail S Golden, Carlos A Guerra, Ute Jacob, Yasuo Takahashi, Josef Settele et al., "The direct drivers of recent global anthropogenic biodiversity loss," *Science advances*, 2022, *8* (45), eabm9982.

Karolyi, G Andrew and John Tobin de la Puente, "Biodiversity finance: A call for research into financing nature," *Financial Management*, 2023.

Lee, Suzanne S and Per A Mykland, "Jumps in financial markets: A new nonparametric test and jump dynamics," *The Review of Financial Studies*, 2008, *21* (6), 2535–2563.

Matsumura, Ella Mae, Rachna Prakash, and Sandra C Vera-Muñoz, "Firm-value effects of carbon emissions and carbon disclosures," *The accounting review*, 2014, *89* (2), 695–724.

Pastor, Lubos and Pietro Veronesi, "Uncertainty about government policy and stock prices," *The journal of Finance*, 2012, *67* (4), 1219–1264.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor, "Sustainable investing in equilibrium," *Journal of financial economics*, 2021, *142* (2), 550–571.

Rillig, Matthias C, Masahiro Ryo, Anika Lehmann, Carlos A Aguilar-Trigueros, Sabine Buchert, Anja Wulf, Aiko Iwasaki, Julien Roy, and Gaowen Yang, "The role of multiple global change factors in driving soil functions and microbial biodiversity," *Science*, 2019, *366* (6467), 886–890.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Schantz, Susan L and John J Widholm, "Cognitive effects of endocrine-disrupting chemicals in animals.," *Environmental health perspectives*, 2001, *109* (12), 1197–1206.

Svendsen, Monica Lian, Eiliv Steinnes, and Hans A Blom, "Vertical and horizontal distributions of Zn, Cd, Pb, Cu, and Hg in uncultivated soil in the vicinity of a zinc smelter at Odda, Norway," *Soil & Sediment Contamination*, 2007, *16* (6), 585–603.

TNFD, "Recommendations of the Taskforce on Nature-Related Financial Disclosures," 2023.

van der Kroft, Bram, Juan Palacios, Roberto Rigobon, and Siqi Zheng, "Timing sustainable engagement in real asset investments," Technical Report, National Bureau of Economic Research 2024.

Watson, Robert, Ivar Baste, Anne Larigauderie, Paul Leadley, Unai Pascual, Brigitte Baptiste, Sebsebe Demissew, Luthando Dziba, Günay Erpul, Asghar Fazel et al., "Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services," *IPBES Secretariat: Bonn, Germany*, 2019, pp. 22–47.

WWF, "Living Planet Report 2020: Deep Dive into Freshwater," 2020. Accessed: 2024-08-14.

Xiong, Xi and Ivan PL Png, "Location of US manufacturing, 1987-2014: A new dataset," *Available at SSRN 3401582*, 2019.

6 Figures

Figure 1: Carbon Disclosure Project question: Where have you published information about your organization's response to biodiversity-related issues?

Figure 1 presents information from the Carbon Disclosure Project for the question "Where have you published information about your organization's response to biodiversity-related issues?". Answers at the firm level are either "Regulatory", "Voluntary", "None".

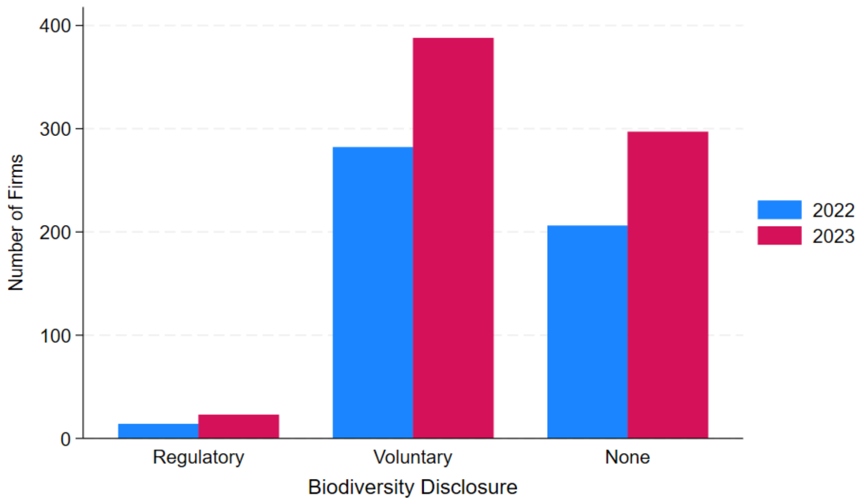


Figure 2: Toxic Release Inventory Facility Locations

Figure 2 presents the locations of Toxics Release Inventory facilities in the contiguous United States as reported by the Environmental Protection Agency.

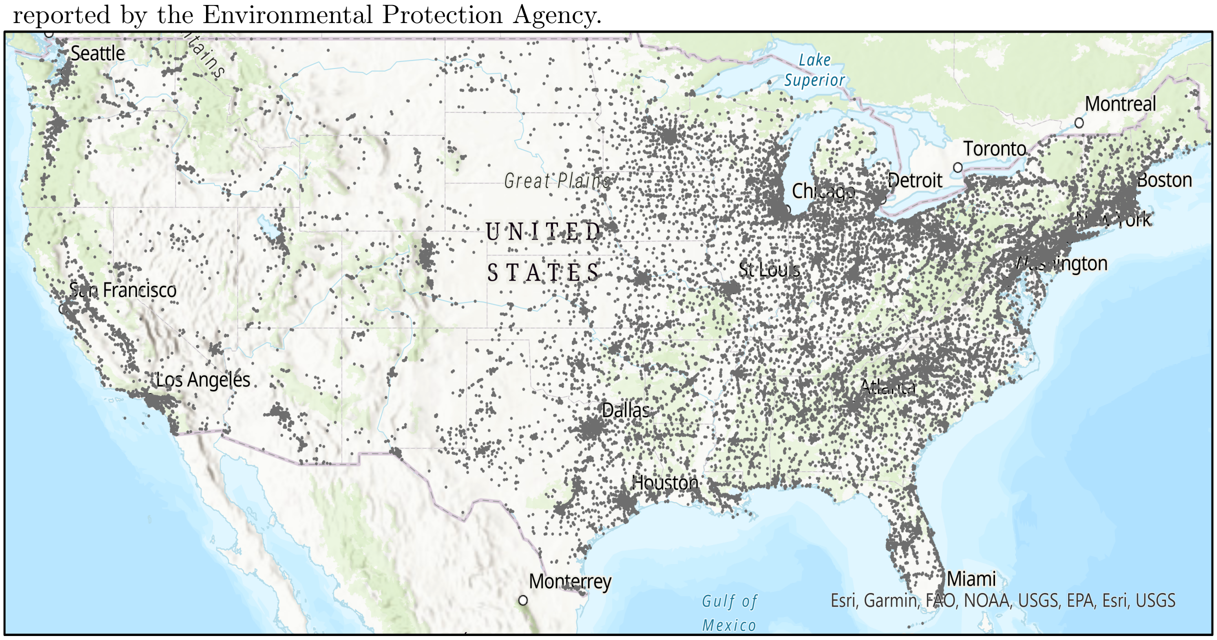


Figure 3: Protected Areas in the Contiguous United States

Figure 3 presents a spatial map of protected areas in the contiguous United States as defined by the World Database on Protected Areas.

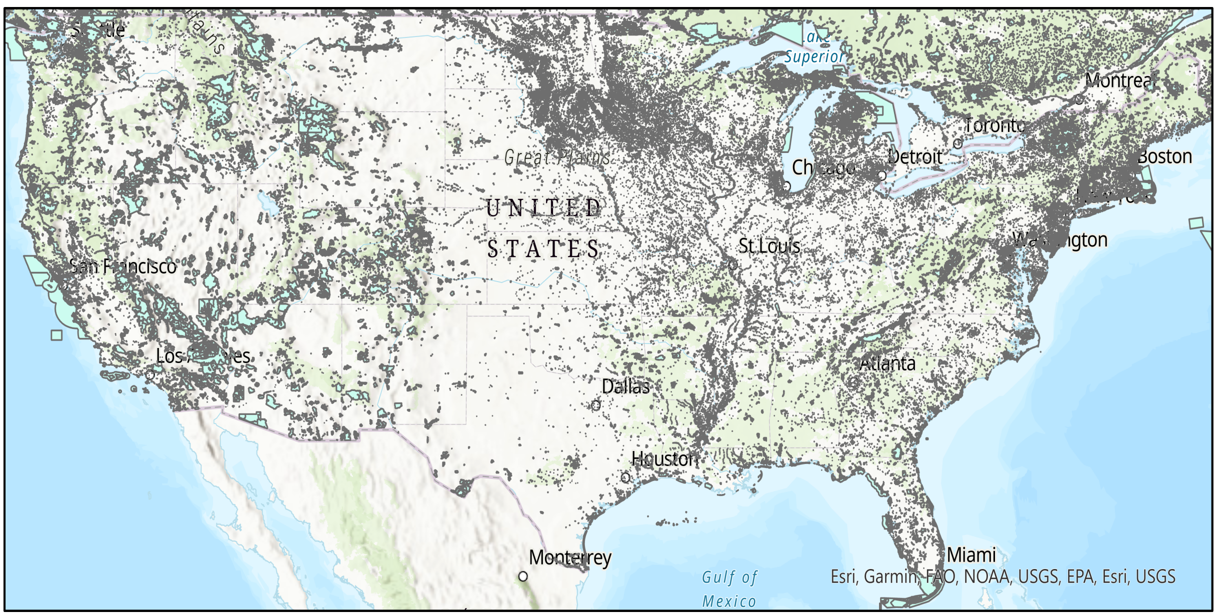


Figure 4: Total Waste Released Near Protected Areas

Figure 4 presents a time-series of the total log emissions released by all TRI facilities in the US and whether the emissions are within three miles of a protected area (solid line). The dashed line is the summation of all emissions that are more than three miles away from a protected region. The left graph weights the chemical release by its toxicity according to the the EPA's Risk-Screening Environmental Indicators (RSEI) model. The graph on the right presents the unweighted totals.

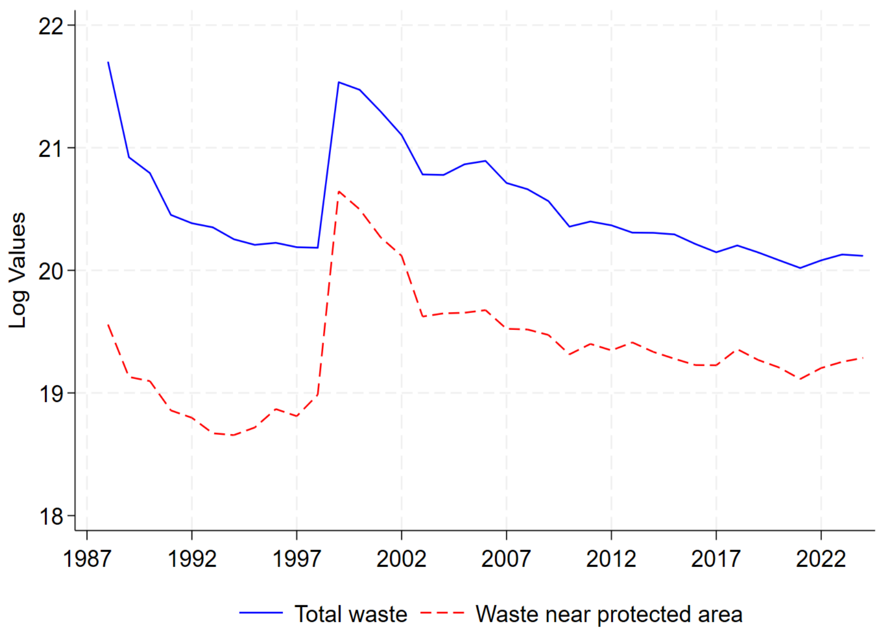


Figure 5: Growth in built-up area surrounding the first reporting date of the TRI facility

Figure 5 presents the year over year log growth in built up area in a 500 meter area around the Toxics Release Inventory facility coordinates. This is calculated for facilities in the contiguous United States. Spatial data for built-up area is obtained from the National Land Cover Database fractional impervious product which is available from 1985-2023. Time zero is the year that the facility first reports to the TRI program.

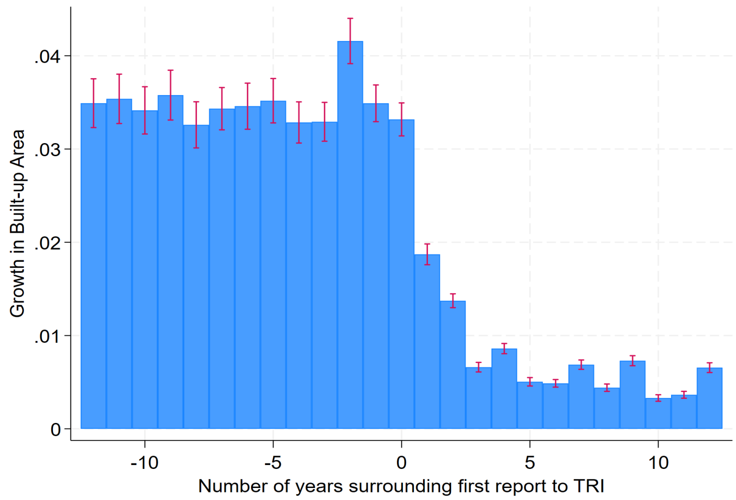


Figure 6: Difference in reporting year and estimated construction year

Figure 6 presents the difference in the first reporting year of a TRI facility and the estimated construction year using a jump detection algorithm. The jump detection algorithm classifies the construction date of a facility as first time the year over year growth of the built up area in the 500 m radius of the coordinates provided by to TRI is over 0.473. This number, 0.473, is the average year over year growth of built up area of all facilities in the sample. Time zero is the first reporting date of the facility to the TRI.

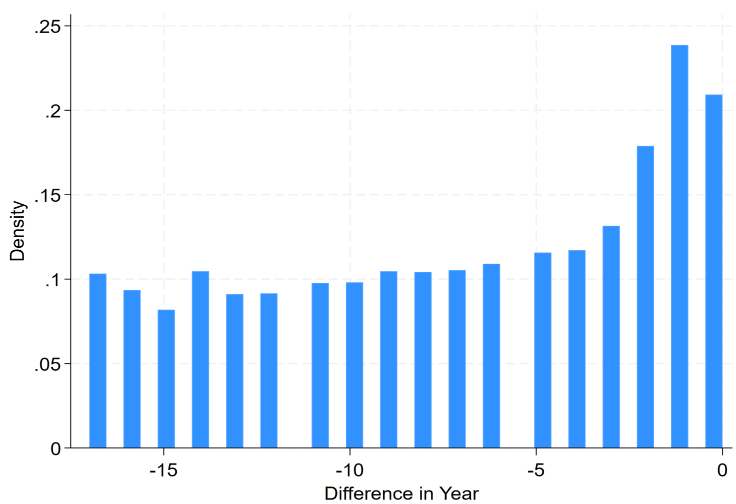


Figure 7: Growth in built-up area using the construction date of a TRI facility

Figure 7 presents the year over year log growth in built up area in a 500 meter area around the TRI facility coordinates. The firms here are first filtered using a jump detection algorithm. The jump detection algorithm classifies the construction date of a facility as first time the year over year growth of the built up area in the 500 m radius of the coordinates provided by to TRI is over 0.473. Time zero is therefore the identified construction date of the

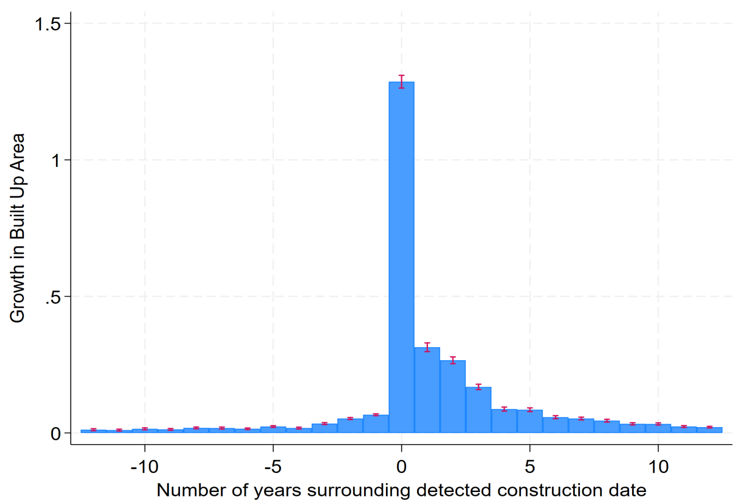


Figure 8: Coefficient on interaction term for lead at various distances from facilities

Figure 8 presents estimates of interaction terms that measure the difference in the average lead concentration in water bodies closer to and farther away from TRI facilities. The dependent variable is the average lead measurement in a year at a monitoring station according to the Water Quality Portal. The specification interacts two indicator variables, *Construct* and *Distance*. Construct is a 0/1 variable that equals 1 if the construction date is between [0,4] years of the date identified using the jump-detection algorithm, and zero if between two to five years prior [-2,-5]. Distance is an indicator equal to zero if a lead monitoring station is farther away from a facility and one if it is closer to a facility. For example, an x-axis value of 4 indicates that distance equals zero if monitoring stations are 5-30 km away from a facility and one if the station is 500 m - 4 km away. Statistically significant positive interaction terms suggest facilities closer to water bodies show higher lead concentrations. Estimates are double clustered at the facility-monitoring station level.

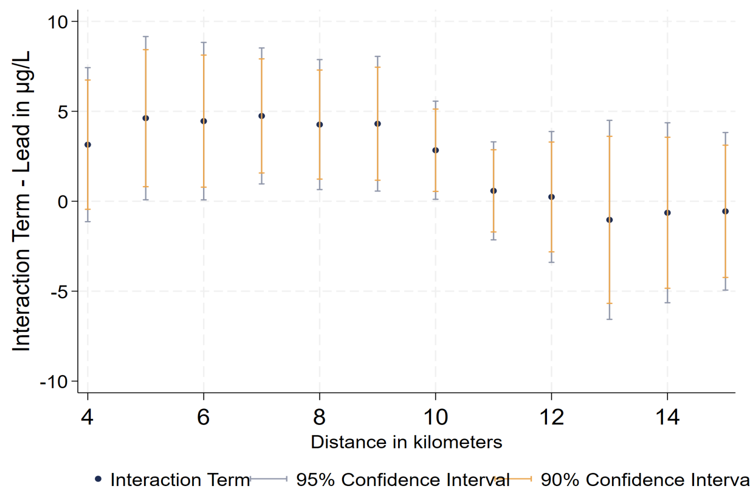


Figure 9: Stacked Difference-in-Difference estimate of lead concentration after construction of facility

Figure 9 presents an event-plot of coefficients from a stacked difference-in-differences estimation measuring the difference in lead concentration in bodies of water within 500 m to 8 km of a facility compared to that found 9-30 km away after a facility is constructed. Treated monitoring stations are those closer when the facility is constructed and untreated are those stations that have yet to be treated or between 9-30 km away. The base year is the year before the construction date of the facility as identified by the jump detection algorithm. Confidence intervals at the 95% level.

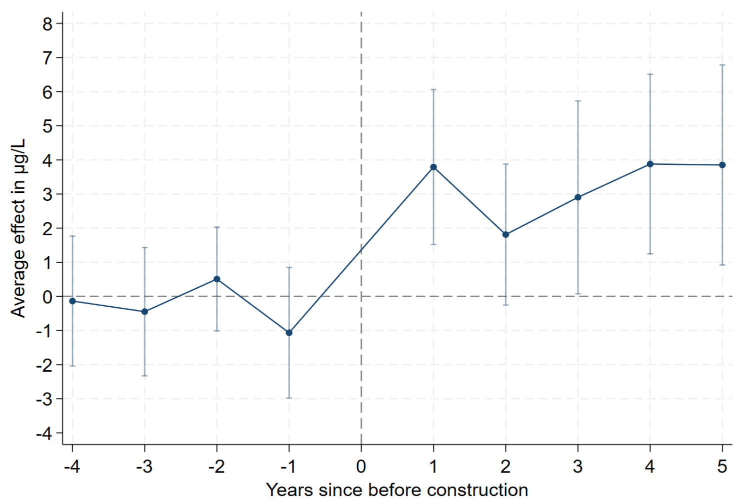


Figure 10: Coefficient on interaction term for animal population loss at various distances from facilities

Figure 10 presents estimates of interaction terms that measure the difference in the animal population counts closer to and farther away from TRI facilities. The dependent variable is the effort corrected animal counts at a survey site according to the PREDICTS database (Hudson et al., 2017). The specification interacts two indicator variables, *Construct* and *Distance*. Construct is a 0/1 variable that equals 1 if the construction date is between [0,4] years of the date identified using the jump-detection algorithm, and zero if between two to five years prior [-2,-5]. Distance is an indicator equal to zero if a survey site is farther away from a facility and one if it is closer to a facility. For example, an x-axis value of 6 indicates that distance equals zero if a animal survey site is 7-30 km away from a facility and one if the site is 500 m - 6 km away. Statistically significant negative interaction terms suggest survey sites closer to facilities have lower animal population counts than those farther away. Standard errors are based on clusters at the facility level.

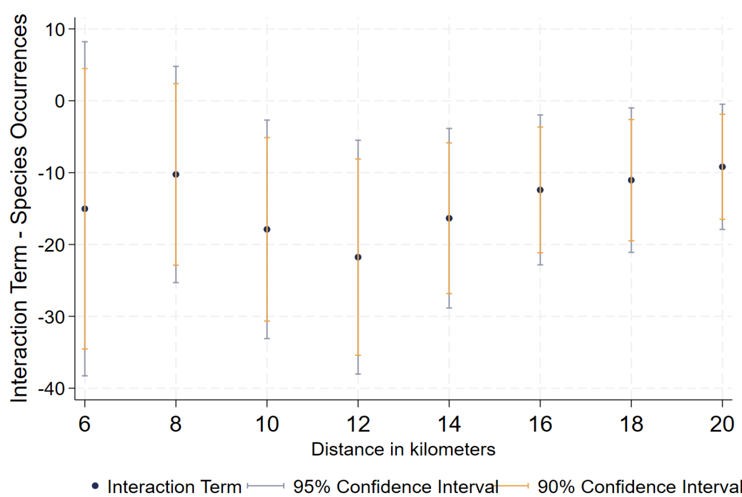
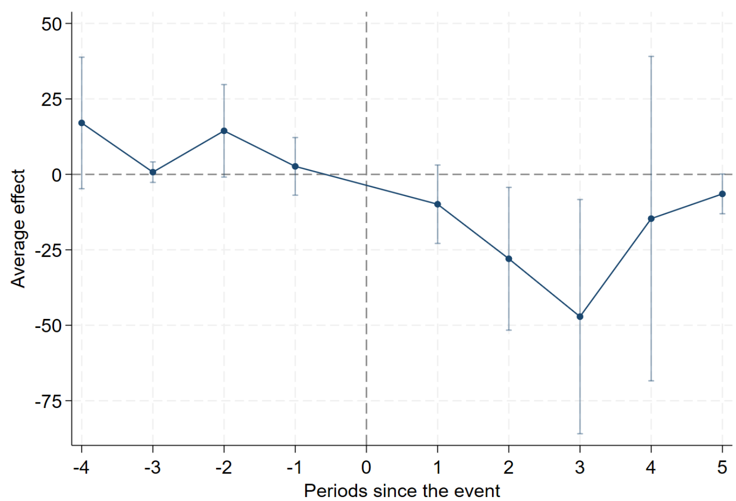


Figure 11: Stacked Difference-in-Difference estimate of animal population after construction of facility

Figure 11 presents an event-plot of coefficients from a stacked difference-in-differences estimation measuring the difference in animals in survey sites within 500 m to 12 km of a facility compared to survey sits found 13-30 km away after a facility is constructed. The dependent variable is the effort corrected animal counts at a survey site according to the PREDICTS database (Hudson et al., 2017). Treated sites are those closer when the facility is constructed and untreated are those survey sites that have yet to be treated or between 13-30 km away. The base year is the year before the construction date of the facility as identified by the jump detection algorithm. Confidence intervals at the 9



7 Tables

Table 1: Summary Statistics

	Mean	SD	P25	P50	P75	N
Number of Cases	1.15	(0.71)	1.00	1.00	1.00	15,320
Penalty, sqrt	815.55	(4,280.94)	31.62	114.89	291.55	15,320
Releases, sqrt	331.22	(762.56)	16.53	106.39	341.43	15,320
Houses	53,233.79	(86,546.30)	7,635.00	26,302.00	68,442.00	15,286
Minority Pop, Perc	39.92	(16.14)	26.67	36.68	49.98	15,210
Pop Density	1,824.80	(3,015.25)	240.58	884.11	2,294.72	15,320
Voluntary Disclosure	0.21	(0.40)	0.00	0.00	0.00	15,320
Biodiversity Footprint	0.26	(0.35)	0.00	0.00	0.51	218,914
Emissions Intensity	14.35	(27.49)	1.68	6.06	16.84	218,914
Monthly Returns	1.26	(12.96)	(5.01)	0.85	6.74	218,914
Assets, log	7.31	(1.96)	5.91	7.30	8.68	218,914
Tobin's Q	1.67	(1.01)	1.13	1.40	1.90	218,019
Book to Market	0.57	(5.57)	0.32	0.52	0.81	218,580
Leverage	0.28	(0.18)	0.16	0.27	0.38	218,914
Return on Equity	(0.00)	(3.46)	0.01	0.01	0.02	218,547
Plant, Prop, Equip	0.32	(0.18)	0.18	0.28	0.43	218,854
Vertebrate Population	55.48	(78.66)	17.00	25.00	64.00	10,670
Voluntary Reporting	0.21	(0.40)	0.00	0.00	0.00	282
Avg Built Up, 500 m - 12 km	18.98	(11.82)	9.37	18.11	26.82	302,454
Avg Built Up, 12 km - 30 km	15.34	(11.98)	7.46	14.26	20.82	302,454
Facility Built Up, YoY ln Growth	0.02	(0.15)	0.00	0.00	0.00	1,214,751
Facilities, Lead Matched						4,543
Monitoring Stations, Facilities						5,941
Facilities, Survey Sites Matched						1,141
Survey Sites, Facilities Matched						1,485
Ecoregion Types						47
Landuse Types						10
Survey Studies						25
Firms						1,373
Facilities Near Protected Areas						1,853
Other Facilities						8,271

Table 2: Difference-in-Difference estimates for lead concentration when a facility is constructed

Table 2 presents estimates of interaction terms that measure the difference in the average lead concentration in water bodies closer to and farther away from TRI facilities. The dependent variable is the average lead measurement in a year at a monitoring station according to the Water Quality Portal. The specification interacts two indicator variables, *Construct* and *Distance*. Construct is a 0/1 variable that equals 1 if the construction date is between [0,4] years of the date identified using the jump-detection algorithm, and zero if between two to five years prior [-2,-5]. Distance is an indicator equal to zero if a lead monitoring station is farther away from a facility and one if it is closer to a facility. For example, in column (1), a value of 4 indicates that distance equals zero if monitoring stations are 5-30 km away from a facility and one if the station is 500 m - 4 km away. The average built up area is calculated for both the nearer and further area from the Annual National Land Cover Database. Statistically significant positive interaction terms suggest facilities closer to water bodies show higher lead concentrations. There are facility by monitoring station fixed effects that subsume the distance variable. Year fixed effects are also included. Standard errors, in parentheses, are double clustered at the facility-monitoring station level. Significance at the 1%, 5%, and 10% levels is denoted by ***, **, *.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	[4, 5-30]	[5, 6-30]	[6, 7-30]	[7, 8-30]	[8, 9-30]	[9, 10-30]	[10, 11-30]	[11, 12-30]	[12, 13-30]	[13, 14-30]	[14, 15-30]	[15, 16-30]
Construct=1 x Distance=1	3.145	4.619**	4.454**	4.742**	4.262**	4.309**	2.835**	0.578	0.239	-1.035	-0.641	-0.561
	(2.184)	(2.315)	(2.232)	(1.929)	(1.843)	(1.910)	(1.392)	(1.389)	(1.855)	(2.822)	(2.552)	(2.235)
Construct=1	-1.754	-1.824	-1.930	-1.993	-2.178	-2.149	-1.331	-1.635	-0.746	-1.238	-1.442	-1.392
	(1.781)	(1.786)	(1.821)	(1.847)	(1.892)	(1.917)	(1.463)	(1.685)	(1.307)	(1.447)	(1.452)	(1.488)
Built-up Area, Near	-0.274*	-0.349*	-0.394*	-0.418*	-0.347	-0.391*	-0.411*	-0.395	-0.364	-0.388	-0.219	-0.356
	(0.150)	(0.184)	(0.212)	(0.218)	(0.221)	(0.228)	(0.238)	(0.241)	(0.234)	(0.292)	(0.235)	(0.312)
Built-up Area, Far	0.338	0.448	0.511	0.568	0.527	0.531	0.547	0.559	0.470	0.406	0.247	0.323
	(0.467)	(0.486)	(0.522)	(0.533)	(0.539)	(0.551)	(0.538)	(0.491)	(0.512)	(0.591)	(0.525)	(0.560)
FacilityxStation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R Squared	0.212	0.212	0.212	0.213	0.212	0.212	0.221	0.215	0.217	0.213	0.211	0.207
Observations	22793	22738	22668	22743	22645	22576	22554	22523	22426	22489	22212	22321

Table 3: Stacked Difference-in-Difference estimate of lead concentration after construction of facility

Table 5 presents the results from a stacked difference-in-differences estimation measuring the difference in lead concentration in bodies of water within 500 m to 8 km of a facility compared to that found 9-30 km away after a facility is constructed. Treated monitoring stations are those closer when the facility is constructed and untreated are those stations that have yet to be treated or between 9-30 km away. The base year is the year before the construction date of the facility as identified by the jump detection algorithm. The average built up area is calculated for both the nearer and further area from the Annual National Land Cover Database. Statistically significant positive interaction terms suggest facilities closer to water bodies show higher lead concentrations. There are facility by monitoring station fixed effects that subsume the distance variable. Year fixed effects are also included. Standard errors, in parentheses, are double clustered at the facility-monitoring station level. Significance at the 1%, 5%, and 10% levels is denoted by ***, **, *.

	(1)
Pre 1 Years	-0.138
	(0.972)
Pre 2 Years	-0.447
	(0.959)
Pre 3 Years	0.509
	(0.776)
Pre 4 Years	-1.066
	(0.977)
Post 1 Years	3.792***
	(1.158)
Post 2 Years	1.812*
	(1.054)
Post 3 Years	2.904**
	(1.442)
Post 4 Years	3.878***
	(1.343)
Post 5 Years	3.852**
	(1.496)
Built-up Area, Near	-0.464***
	(0.028)
Built-up Area, Far	0.580***
	(0.064)
FacilityxStation	Yes
Year	Yes
Adj R Squared	0.178
Observations	823314

Table 4: Difference-in-Difference estimates for animal loss when a facility is constructed

Table 4 presents estimates of interaction terms that measure the difference in the animal population counts closer to and farther away from TRI facilities. The dependent variable is the effort corrected animal counts at a survey site according to the PREDICTS database (Hudson et al., 2017). The specification interacts two indicator variables, *Construct* and *Distance*. Construct is a 0/1 variable that equals 1 if the construction date is between [0,4] years of the date identified using the jump-detection algorithm, and zero if between two to five years prior [-2,-5]. Distance is an indicator equal to zero if a survey site is farther away from a facility and one if it is closer to a facility. For example, a value of 6 indicates that distance equals zero if a animal survey site is 7-30 km away from a facility and one if the site is 500 m - 6 km away. I include facility, year, survey study, land use, and ecoregion fixed effects. The average built up area is calculated for both the nearer and further area from the Annual National Land Cover Database. Statistically significant negative interaction terms suggest survey sites closer to facilities have lower animal population counts than those farther away. Standard errors are clustered at the facility level. Significance at the 1%, 5%, and 10% levels is denoted by ***, **, *.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	[6, 7-30]	[8, 9-30]	[10, 11-30]	[12, 13-30]	[14, 15-30]	[16, 17-30]	[18, 19-30]	[20, 21-30]
Construct=1	14.400*	9.445	11.487	33.909***	21.927**	20.403**	16.799*	13.073
	(7.983)	(8.894)	(9.578)	(12.791)	(9.695)	(9.427)	(9.428)	(8.882)
Distance=1	-1.035	-0.034	3.748	6.728*	7.859*	5.144	5.877*	5.989*
	(7.872)	(5.541)	(3.191)	(3.836)	(4.243)	(3.230)	(3.390)	(3.386)
Construct=1 x Distance=1	-15.030	-10.251	-17.891**	-21.762***	-16.347**	-12.405**	-11.045**	-9.190**
	(11.860)	(7.677)	(7.759)	(8.301)	(6.376)	(5.321)	(5.128)	(4.444)
Built-up Area, Near	-9.309***	-9.915***	-9.660***	-18.410***	-16.432***	-15.257***	-13.801***	-7.662
	(2.051)	(2.750)	(2.850)	(4.106)	(4.305)	(4.455)	(4.561)	(4.885)
Built-up Area, Far	34.317*	42.947**	51.965**	29.184	42.650*	39.818*	42.108*	25.459
	(18.979)	(20.683)	(20.390)	(22.090)	(21.878)	(21.792)	(23.729)	(19.011)
Facility	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Study	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Land Use	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ecoregion	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R Squared	0.774	0.769	0.780	0.786	0.774	0.771	0.773	0.769
Observations	2505	2469	2450	2343	2450	2473	2469	2435

Table 5: Stacked Difference-in-Difference estimate of animal count after construction of a facility

Table 5 presents estimates from a stacked difference-in-differences estimation measuring the difference in animals in survey sites within 500 m to 12 km of a facility compared to survey sits found 13-30 km away after a facility is constructed. The dependent variable is the effort corrected animal counts at a survey site according to the PREDICTS database (Hudson et al., 2017). Treated sites are those closer when the facility is constructed and untreated are those survey sites that have yet to be treated or between 13-30 km away. The base year is the year before the construction date of the facility as identified by the jump detection algorithm. I include facility, year, survey study, land use, and ecoregion fixed effects. The average built up area is calculated for both the nearer and further area from the Annual National Land Cover Database. Statistically significant negative interaction terms suggest survey sites closer to facilities have lower animal population counts than those farther away. Standard errors are clustered at the facility level. Significance at the 1%, 5%, and 10% levels is denoted by ***, **, *.

	(1)
Pre 1 Years	17.041
	(11.115)
Pre 2 Years	0.738
	(1.727)
Pre 3 Years	14.427*
	(7.820)
Pre 4 Years	2.644
	(4.871)
Post 1 Years	-9.892
	(6.628)
Post 2 Years	-27.965**
	(12.066)
Post 3 Years	-47.134**
	(19.798)
Post 4 Years	-14.658
	(27.422)
Post 5 Years	-6.472*
	(3.355)
Built-up Area, Far	31.564***
	(5.979)
Built-up Area, Near	-4.644***
	(1.199)
Facility	Yes
Year	Yes
Study	Yes
Land Use	Yes
Ecoregion	Yes
Adj R Squared	0.847
Observations	63990

Table 6: Fama Macbeth Regressions

Table 5 presents risk premiums and standard errors of Fama Macbeth regressions with the firm level biodiversity footprint and emissions intensity. The biodiversity footprint of a firm is calculated as the square root of total emissions from a firm's emissions divided by all emissions according to the Toxics Release Inventory dataset. Fama and French (2015) five factors are included along. Various firm level control variables are included such as log assets—these are lagged values from one quarter ago. Significance is denoted at the 1%, 5%, and 10% levels by ***, **, *.

	(1)	(2)
Biodiversity Footprint	-0.000	-0.000
	(0.000)	(0.000)
Emissions Intensity	0.002	0.000
	(0.002)	(0.001)
MktRF	0.379	0.350
	(0.260)	(0.264)
SMB	0.100	0.104
	(0.156)	(0.153)
HML	-0.094	-0.115
	(0.199)	(0.202)
RMW	-0.092	-0.064
	(0.139)	(0.142)
CMA	-0.086	-0.147
	(0.125)	(0.127)
Assets, log	0.003*	0.008**
	(0.002)	(0.003)
Tobin's Q	0.000	0.003
	(0.003)	(0.003)
Book to Market	-0.001	-0.002
	(0.001)	(0.001)
Leverage	0.000	0.001**
	(0.000)	(0.000)
Return on Equity	0.000	0.000
	(0.000)	(0.000)
Plant, Property, Equipment	0.000	0.000
	(0.000)	(0.001)
Industry FEs	No	Yes
Average R Squared	0.129	0.151
Observations	216756	216367

Table 7: Event Study Surrounding Kunming Declaration

Table 5 presents regressions documenting the stock price reactions to the Kunming Declaration which was signed October 13, 2021. Biodiversity footprint and emissions intensity are at the firm level. The biodiversity footprint of a firm is calculated as the square root of total emissions from a firm's emissions divided by all emissions according to the Toxics Release Inventory dataset. Various event windows are used around this date. Post is a indicator variable equal to 1 if time is equal October 13, 2021 plus additional days above the column. It is equal to zero for the days before. Abnormal returns are returns in excess of the market, small-minus-big, and high-minus-low factors. Significance is denoted at the 1%, 5%, and 10% levels by ***, **, *.

	(1)	(2)	(3)	(4)	(5)	(6)
	[-7,+6]	[-6,+5]	[-5,+4]	[-4,+3]	[-3,+2]	[-2,+1]
Biodiversity Footprint	0.069	-0.036	-0.071	-0.280**	-0.187	-0.292*
	(0.092)	(0.095)	(0.104)	(0.122)	(0.127)	(0.158)
Post × Biodiversity Footprint	0.026	0.202	0.137	0.335*	0.309	0.287
	(0.125)	(0.132)	(0.148)	(0.171)	(0.195)	(0.234)
Emissions Intensity	-0.001	0.002	0.004	0.006	0.006	0.015
	(0.004)	(0.005)	(0.005)	(0.007)	(0.008)	(0.013)
Post × Emissions Intensity	0.001	-0.003	-0.006	-0.005	-0.005	0.001
	(0.007)	(0.008)	(0.009)	(0.009)	(0.014)	(0.019)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Day	Yes	Yes	Yes	Yes	Yes	Yes
Adj R Squared	0.026	0.032	0.035	0.039	0.030	0.012
Observations	5390	4620	3850	3080	2310	1540

Table 8: Government monitoring and facility enforcement costs

Table 8 presents regressions documenting the relationship between the total federal penalties and the square root of toxic chemicals released by a facility. The sample comprises all enforcement cases reported in the ICIS-FE&C database by the EPA. The variable "Protected" equals 1 if a facility is within 12 km of a protected area and 0 otherwise. Control variables include the number of houses, minority population percentage, and population density within a 5 km radius, sourced from the 2010 U.S. census. Significance is denoted at the 1%, 5%, and 10% levels by ***, **, *.

	(1) $ Costs	(2) $ Costs	(3) $ Costs
Releases, sqrt	0.686***	0.658***	0.626***
	(0.095)	(0.097)	(0.096)
Protected=1	231.273***	173.071**	187.120**
	(81.145)	(83.273)	(80.424)
Protected=1 × Releases, sqrt	-0.499***	-0.467**	-0.415**
	(0.181)	(0.188)	(0.175)
Houses		-0.000	-0.001
		(0.002)	(0.002)
Minority Pop, Perc		-1.609	-1.091
		(2.018)	(2.357)
Pop Density		-0.014	0.005
		(0.048)	(0.058)
IndustryxYear	Yes	Yes	Yes
State	No	No	Yes
Adj R Squared	0.298	0.304	0.309
Observations	15252	15143	15143

Table 9: Facility self reporting violations and emissions near protected areas

Table 9 presents results from linear probability models documenting the relationship between the square root of toxic chemicals released by a facility and whether a violation was self-reported by the facility. The sample comprises all enforcement cases reported in the ICIS-FE&C database by the EPA. The variable "Protected" equals 1 if a facility is within 12 km of a protected area and 0 otherwise. Control variables include the number of houses, minority population percentage, and population density within a 5 km radius, sourced from the 2010 U.S. census. Significance is denoted at the 1%, 5%, and 10% levels by ***, **, *.

	(1)	(2)	(3)
Releases, sqrt	-0.004***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)
Protected	-2.122**	-3.288***	-3.105***
	(0.924)	(0.956)	(0.956)
Protected × Releases, sqrt	0.002***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)
Houses		0.000	0.000
		(0.000)	(0.000)
Minority Pop, Perc		-0.065***	-0.076***
		(0.022)	(0.025)
Pop Density		-0.001	-0.001
		(0.001)	(0.001)
IndustryxYear	No	Yes	Yes
State	No	No	Yes
Adj R Squared	0.158	0.161	0.176
Observations	15252	15143	15143

Table 10: Relationship Between Self-Reporting Biodiversity in 10-Ks and Emissions in Biodiverse Areas

Table 10 presents results from linear probability models documenting the relationship between metrics from Giglio et al. (2023) and the biodiversity footprint of the firm. The biodiversity footprint of a firm is calculated as the square root of total emissions from a firm's emissions divided by all emissions according to the Toxics Release Inventory dataset. Various firm level control variables are included such as log assets—these are lagged values from one year ago. Significance is denoted at the 1%, 5%, and 10% levels by ***, **, *.

	(1) Aggregate	(2) Regulation	(3) Negative
Biodiversity Footprint	-0.006	-0.006	0.005
	(-0.292)	(-0.329)	(0.274)
Emissions Intensity	0.001**	0.001**	0.001*
	(2.463)	(2.302)	(1.755)
L.Assets, log	0.004	0.004	-0.005
	(0.713)	(0.724)	(-0.592)
L.Tobin's Q	0.002	0.000	0.000
	(1.405)	(0.279)	(0.192)
L.Return on Equity	-0.000	-0.002	0.002
	(-0.002)	(-0.147)	(0.064)
L.Book to Market	0.010*	0.007*	0.006
	(1.832)	(1.675)	(0.570)
L.Leverage	-0.017	-0.022	0.025
	(-0.738)	(-1.120)	(0.916)
L.Plant, Property Equipment	0.076	0.050	0.115
	(1.098)	(0.903)	(1.278)
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
ADJ R Squared	0.518	0.527	0.400
Observations	29027	29027	29027

Table 11: Relationship between voluntary reporting of biodiversity impact in the carbon disclosure project and biodiversity footprint

Table 11 presents results from linear probability models documenting the relationship between firms mentioning they have a biodiversity impact and their biodiversity footprint. The biodiversity footprint of a firm is calculated as the square root of total emissions from a firm's emissions divided by all emissions according to the Toxics Release Inventory dataset. Each column represents a different reporting year for the TRI dataset. Various firm level control variables at a yearly frequency are included such as log assets—these are lagged values from one year ago. Significance is denoted at the 1%, 5%, and 10% levels by ***, **, *. The constant is suppressed.

	(1) 2020	(2) 2021	(3) 2022	(4) 2023
Biodiversity Footprint	-0.011	0.011	0.004	0.042
	(-0.169)	(0.162)	(0.060)	(0.586)
Emissions Intensity	0.001	0.001	-0.001	-0.001
	(0.343)	(0.145)	(-0.344)	(-0.252)
L.Assets, log	0.057**	0.057**	0.059**	0.056**
	(2.511)	(2.369)	(2.506)	(2.346)
L.Tobin's Q	0.013	0.002	0.009	0.013
	(0.491)	(0.084)	(0.417)	(0.349)
L.Return on Equity	0.295	0.588*	-3.180**	-0.394
	(0.699)	(1.792)	(-2.548)	(-0.416)
L.Book to Market	0.138	0.036	0.243	0.085
	(1.150)	(0.413)	(1.276)	(0.670)
L.Leverage	0.083	0.019	0.125	-0.036
	(0.367)	(0.077)	(0.481)	(-0.150)
L.Plant, Property Equipment	0.365*	0.352*	0.468**	0.327
	(1.775)	(1.728)	(2.215)	(1.601)
Industry FE	Yes	Yes	Yes	Yes
ADJ R Squared	0.186	0.178	0.224	0.153
Observations	200	195	191	189

A Appendix A

A.1 Firm level biodiversity footprint

First, I develop a measure of a firm's toxic footprint and contrast this with an emissions intensity measure (Hsu et al., 2023). I define the toxic footprint of a firm as the square root of the emissions released within a 15-km radius of a protected area divided by the total emissions reported. I define emissions intensity as the total emissions reported by the firm divided by the total assets of a firm. I show in Figure 12 the time series correlation between the two metrics. Generally, I find a low correlation coefficient between emission intensity and toxic footprint suggesting that the toxic footprint is differentiated from the one used in (Hsu et al., 2023).

Figure 12: Correlation coefficient of biodiversity footprint and emissions intensity

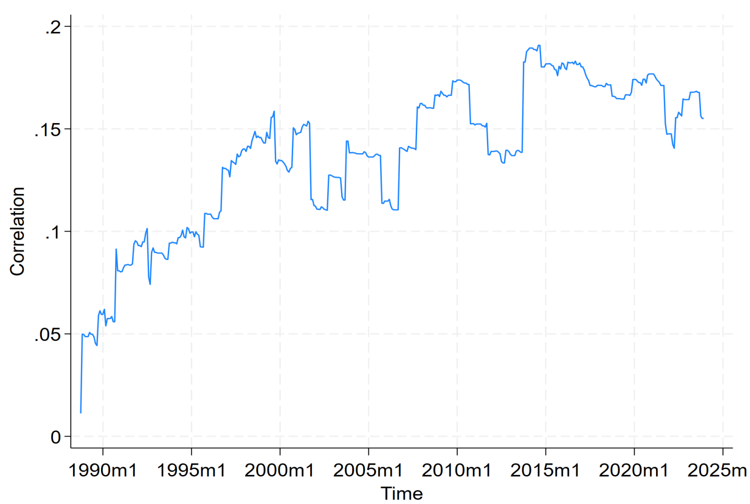


Figure 12 presents correlation coefficients of

Figure 13 presents a scatter plot of the mean, standardized values of the toxic footprint and emissions intensity metrics of the Fama-French 48 industries. Generally, most industries have fairly low values of emission intensities and footprints. Some, however, are standouts. Gold, mining, and coal mining sectors have relatively high values of either toxic footprints or emissions intensity. Another standout is the "Other" industry which includes Steam & air conditioning supplies, Irrigation systems, and Sanitary services. This industry has a large emission intensity but low toxic footprints.

Figure 13: Industries with high emissions and toxic footprints

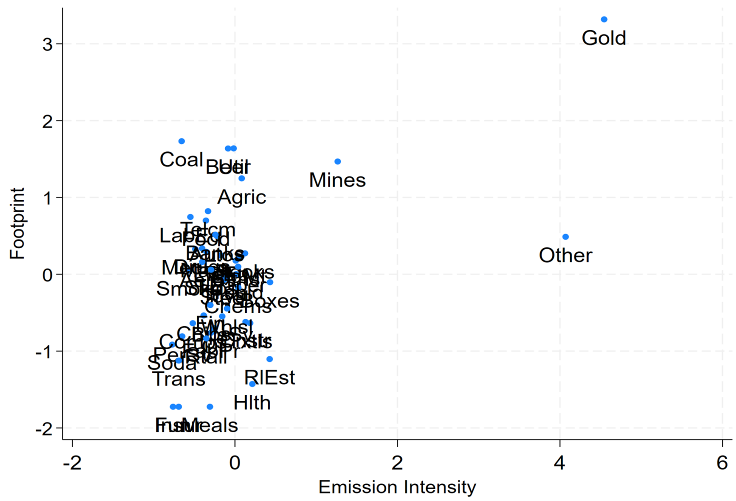


Figure 13 presents the mean, standardized values of emission intensity and toxic footprint in each Fama-French 48 industry sector.