Characterizing disproportionality in facility-level toxic releases in US manufacturing, 1998–2012

M Collins 1,4, S Pulver 2, D Hill 1 and B Manski 3

1 Department of Environmental Studies, SUNY College of Environmental Science and Forestry, Syracuse, United States of America
2 Environmental Studies Program, University of California, Santa Barbara, United States of America
3 Department of Sociology, University of California, Santa Barbara, United States of America
4 Author to whom any correspondence should be addressed.
E-mail: mbcollin@esf.edu

Keywords: disproportionality, US manufacturing, industrial pollution

Abstract

The relationship between economic activity and environmental pollution is a topic of extensive research. Although a proportional relationship between the two is often the default assumption, emerging scholarship suggests that polluting releases are disproportionally distributed across units of production. This paper examines if proportionality or disproportionality best characterizes the production of toxic pollution in US manufacturing from 1998 to 2012. Examining US Environmental Protection Agency Toxics Release Inventory data from over 25 000 facilities in 322 industries, we find consistently high levels of disproportionality across facility-level toxic releases within industries, even when controlling for facility size. Moreover, high levels of within industry disproportionality are remarkably stable over the fifteen-year study period. In other words, year by year a small handful of egregiously polluting facilities account for the vast majority of toxic releases within a given industry. Our findings suggest that disproportionality should be understood as the default pattern of pollution generation rather than an exceptional case and that policymakers should seek to reduce pollution via carefully considered targeting strategies rather than broad-stroke decision making.

1. Introduction

There exists a common, default presumption in environmental policy debates of a proportional relationship between economic activity and environmental degradation. More productive activity results in greater environmental harm (Hertwich et al. 2010). This presumption is grounded in the environmental impact models of the 1970s, which estimated environmental impact in relation to an additional unit of population or production (Ehrlich and Holdren 1971, Commoner et al. 1972, Meadows et al. 1972). However, even this early literature recognized the potential of technology to disrupt the proportional relationship, and evidence for a counter narrative of a disproportional relationship between productive activity and environmental harm is accumulating. This counter narrative contends that modern societies and economies are characterized by environmental disproportionalities in production and consumption, defined as ‘strikingly unequal patterns of privileged access to environmental rights and resources’ (Freudenburg 2005:89), and their counterpart, strikingly unequal patterns of exposure to environmental harm (Mohai and Saha 2015, Holifield et al. 2017).

Disproportionalities in exposure are well-established and widely accepted. The analysis of disproportionality in the production of environmental harm is a more recent research focus. Nevertheless, scholarship at multiple scales of analysis and using various measures of environmental degradation documents that the generation of environmental harm is not distributed equally across units but rather concentrated within a small group of egregious actors. At the national level, evidence for disproportionality exists in the differential per capita ecological footprints of societies with equivalent levels of prosperity (York et al. 2002, Chambers et al. 2014). Studies analyzing patterns in greenhouse gas (GHG) emissions pinpoint disproportionalities across nations, class groups, and...
households (Baer 2009, Kennedy et al 2014). Analyses of air pollution from vehicle emissions identify the disproportional contribution of high-emitting vehicles (Cadle et al 1997; Nguyen and Marshall 2018).

Our specific focus is on disproportionality in organizational contributions to environmental harm, adding a new perspective to the extensive research literature on business and the environment (Pulver and Manski 2021). Evidence for disproportionality in the production of toxic pollution across organizations emerges from studies examining toxic releases in the primary metals industry (Freudenburg 2005), the chemical industry (Ash et al 2009), the electric utility industry (Van Atten et al 2014, Prechel and Istvan 2016), and industries in the state of Wisconsin and the county of Milwaukee (Collins 2012). Likewise, research on organizational GHG emissions establishes disproportionality in GHG pollution as a key feature of energy production. For example, Heede (2014) shows that ‘63 percent of cumulative worldwide emissions of industrial CO₂ and methane between 1751 and 2010 [accrued] to 90 “carbon major” entities’. National electric power industries show various degrees of disproportionality in GHG emissions (Grant et al 2013), with consequences for levels of national carbon emissions (Jorgenson et al 2016).

Focusing on the US electric utility industry, Galli et al (2019) find disproportional patterns of GHG emissions at the level of parent companies but not individual facilities when normalizing for electricity produced. Alvarez et al’s (2018) analysis of methane emissions from the US oil and gas supply chain identifies abnormal operating conditions at a small number of facilities leading to a heavy tail in the distribution of emissions.

Extending these findings, other research embeds disproportionality in the production of pollution within larger social and environmental contexts. Nowak et al (2006) explicitly uses a social-environmental systems approach to analyze disproportional contributions to phosphorus loads to stream and lake systems by agricultural operators. Grant et al (2010) integrate pollution data, community data and production facility characteristics in their research. Likewise, Collins et al (2016) co-locate facilities that are outliers in toxic releases and low income communities of color, linking disproportionality in the production of toxic emissions to disproportional exposure. Focusing on air pollution, Tessum et al (2019) build an economy-wide model, linking emission sources, end uses and end users associated with fine particulate matter (PM_{2.5}) and highlight disproportionality in the racial-ethnic composition of groups producing PM_{2.5} pollution versus those subjected to exposure.

Cumulatively, the above research begins to provide evidence for a disproportionate relationship between economic activity and various forms of pollution. However, most studies are cross-sectional, focusing on disproportionality in the production of pollution at one point in time, and/or have a narrow focus, analyzing change over time in a single industry or sector. As a result, the extent of disproportionality in the production of environmental harm remains an open question. Is disproportionality the exception or the rule? Our effort herein is the first multi-industry, longitudinal investigation of the pattern that best characterizes the production of toxic pollution in the US economy. We analyze toxic release data from over 25 000 establishments from the US Environmental Protection Agency’s (EPA) Toxics Release Inventory (TRI) database, in over 300 industries in the US manufacturing sector over the fifteen-year period from 1998 to 2012. We find that high levels of disproportionality characterize facility-level toxic releases in the majority of industries. Moreover, high disproportionality is remarkably stable across time, despite overall declines in aggregate toxic emissions during the fifteen-year study period.

The findings have both scholarly and policy implications. First, the results indicate that scholarship on toxic pollution in the US needs to account for disproportionality in emissions across facilities. An underlying, disproportionate distribution suggests that research based on aggregate or average contributions to environmental degradation fundamentally mischaracterizes the actors and processes causing environmental harm. In particular, average impact estimates are likely to overestimate the environmental harm caused by the vast majority of actors and underestimate the impacts of the worst offenders. Rather, scholarship should focus on the group of organizations that pollute far more than their fair share as a distinct class of organizations, on the effects of disproportionality on society, and on the social structures that create and perpetuate disproportionality in the production of pollution. We support this new trajectory of environmental pollution research by providing metrics that quantify disproportionality in toxic releases by US manufacturing facilities across industries and time.

Our second key contribution is to advance scholarship using TRI data, with the goal of improving user guidance and expanding user access. Comprehensive, quantitative data about the environmental performance of facilities are a cornerstone for information-based environmental management, effective environmental regulation and citizen environmental action (Kareiva et al 2015). TRI offers one such data source, as one of the most comprehensive, longitudinal, facility-level environmental performance datasets in the US. In the social sciences, TRI data have been used to analyze both the drivers of industrial pollution and the distribution of toxic pollution across communities. The first analyzes pollution at the facility and firm levels. Pollution intensity has been found to correlate directly to both an organization’s ability and incentives to limit pollution. Management competence and the technology used in production processes (Streitwieser 1994), more complex organizational structures and lower overall profits (Prechel and Zheng 2012), firm size (Grant et al 2002), and community pressure...
and a firm’s desire to protect its reputation (Konar and Cohen 1997, Kraft et al 2011) all predict higher rates of pollution. Research analyzing patterns of environmental injustice also mobilize TRI data (Mohai and Saha 2015). Their primary unit of analysis is geographic, be it a neighborhood, city, census tract, etc (Auchincloss et al 2012). Our research offers a third application of TRI data, analyzing patterns of pollution over time at the industry level and contributes a novel dataset of quantifying disproportionality in toxic releases for the over 4 600 industry-years analyzed. We discuss the strengths and weaknesses of the TRI dataset, particularly issues related to industry-level, longitudinal analysis.

Finally, pervasive disproportionality in toxic releases also has direct implications for the regulation of pollution. Rather than designing regulations targeted at all facilities, a disproportionality approach suggests identifying and targeting egregious polluters. This approach has proved fruitful in attribution research in the climate policy community (Frumhoff et al 2015, Ekwurzel et al 2017) and in efforts to mitigate environmental injustice due to the co-location of egregious polluters and low income communities of color (Grant et al 2010, Collins et al 2016, Tessum et al 2019). Moreover, solutions that aim to reconfigure production at egregiously polluting facilities to meet industry averages, rather than closing such facilities, may generate less political opposition to regulation (Berry 2008, Collins 2012).

2. Data and methods

Our analysis calculates disproportionality in toxic emissions from over 25 000 facilities within 322 industries from 1998 to 2012. Toxic release data are sourced from the US EPA TRI program, which provides publicly available data about facility-level toxic emissions dating back to 1987. Reporting waste release type and amount to the TRI program is stipulated for every facility holding a hazardous waste generation permit that meets the following minimum requirements: (1) it be classified in a relevant industry per the North American Industry Classification System (NAICS), (2) it employ ten or more full-time employees, (3) it manufacture or process more than 25 000 pounds of a TRI-listed chemical or otherwise use more than 10 000 pounds of a listed chemical in a given year (US EPA 2014a).

Toxic emissions associated with a particular facility can be analyzed both as yearly hazardous waste totals in pounds (using the TRI) and as an aggregate yearly toxic burden, using the US EPA’s publicly available Risk-Screening Environmental Indicators (RSEI) program. Our analysis uses the latter. The RSEI program models TRI data to account for a chemical release’s toxicity according to its environmental fate and transport. This is especially important when comparing different chemical releases according to their potential environmental hazard, since all chemicals’ RSEI values are on the same scale. As with any model, RSEI does have some use limitations. For example, since RSEI relies exclusively on TRI-reported data to estimate release toxicity not all toxic chemicals or all sources of risk from environmental pollution are accounted for. Along these lines, a low RSEI score suggests a low potential concern from reported TRI releases, but other kinds of environmental risk may also be present. In addition, toxicity data is not available for all TRI chemicals, all exposure routes are not accounted for, and RSEI does not cover all chronic health effects. RSEI is focused most directly on human health toxicity and does not focus on environmental toxicity. Finally, RSEI does not produce risk estimates. Rather, RSEI scores are unitless and meaningful only for comparison to other RSEI scores (USEPA 2018b).

Our second transformation of the toxic emissions data is a size-normalization at the facility-level. If environmental harm is proportional to productive activity, then inequality across facility-level, size-normalized toxic contributions should disappear. We control for facility size using the number of employees at each facility, from the National Establishment Time-Series (NETS)—a proprietary database of facility-based business facts (Barnatchez et al 2017). Employment numbers are an imperfect measure of facility size as compared to production data. For example, more technologically advanced facilities may be both larger producers and smaller employers, inflating their per employee pollution contributions as compared to less technologically advanced facilities, with lower production and more employees. While this is a limitation of the analysis, employment data are the only available metric for facility size for the over 25 000 establishments included in the analysis. TRI and NETS data were merged using a TRI facility identification number (TRIFID)-Duns number crosswalk, that is available for purchase with the NETS data access. The initial merge was 92.1 percent successful, resulting in 524 720 matched facility-years. We did encounter one notable data irregularity—there were several unique TRIFIDs that were linked with the same Duns number. We researched these facilities individually and discovered that they were duplicate records for the same facility, as in every case the only variation in the record was the TRIFID. These duplicates were removed. Accounting for this deduplication effort, the final merge resulted in 519 929 unique facility-years and was 89.1 percent successful.

We calculate annual pollution disproportionality at the industry level, based on size-normalized facility-level RSEI contributions, using a Gini coefficient5, a...
common measure of the inequality among values in a frequency distribution (Yitzhaki and Schechtman 2013). Gini coefficients range from 0 (perfectly proportionate distribution) to 1 (perfectly disproportionate distribution). While there is some debate about the strengths and limitations of Gini as a metric of inequality (Gastwirth 2017), it is the most commonly used measure of inequality in the production of pollution literature. Industry membership is defined using the 2012 six-digit NAICS code, a field included within the TRI database. The first two digits of a NAICS code indicate economic sector. The third, fourth and fifth digits indicate economic subsector, industry group, and industry, respectively, while the sixth digit of a NAICS code indicates the national industry (Boettcher 1999). Analysis at the six-digit level is merited because it ensures that comparisons are being made between peer facilities, operating the same processes and having access to similar technologies.

As a second measure of disproportionality, we also calculate the proportion of facilities within an industry contributing 90 percent of the yearly toxic burden associated with that industry. For this calculation, facilities were first ranked by their toxic burden contribution per employee from largest to smallest. Second, both the individual proportional contribution and the cumulative proportional contribution for each facility was calculated. Third, we identified single facilities that accounted for 50 percent or more and 75 percent or more of the total hazard within an industry by year. We categorize these as egregious polluters.

Our study period extends from 1998 to 2012. Using TRI data longitudinally adds complexity, because the chemicals listed and reporting requirements have changed over time. TRI data were first published in 1987, but most analysts consider the first two years of data to be of limited quality (Kraft et al. 2011). Significant changes in TRI reporting requirements were instituted in the early 1990s and again in 1997. The 1997 change extended TRI reporting requirements to a range of new industries, including for example the highly polluting mining industry, which first reported toxic emissions in 1998. Moreover, the 1997 change created discontinuities in the number of facilities reporting in given industries, changing the disproportionality distribution in those industries across the 1997 to 1998 reporting years (US EPA 2014b). In addition to the 1997 changes, new chemicals were added to the TRI list between 1999/2000 and 2000/2001 (US EPA 2018a). While we selected our research timeframe to minimize these types of changes, some inconsistencies across years remain.

As with TRI reporting requirements, the NAICS system has also undergone revisions, with the 2012 categorization updating the 2007 classification. The net effect was an overall decrease from 1175 to 1065 listed industries, reflecting changes in the economy, including both a decline in the manufacturing sector and the addition of newly emerging industries. Specifically, the 2012 revision made noticeable changes to six of the 20 NAICS sectors. The changes most relevant to our study are within sectors 22 and 31–33. In sector 22, Other Electric Power Generation (NAICS 221119) was deleted and establishments were reclassified to reflect emergent clean energy work. As such five new 6-digit industries are included in the 2012 revision: Solar (NAICS 221114), Wind (NAICS 221115), Geothermal Electric (NAICS 221116), Biomass (NAICS 221117), and Other Electric Power Generation (NAICS 221118). In sectors 31 through 33, major changes caused the collapsing of detail—meaning that specific sectors were reclassified into more general ‘Other’ categories (NAICS Association 2018). A review of NAICS code changes over time at the facility level resulted in 9,166 shifts in NAICS codes, meaning that in these cases a facility reported releases under more than one NAICS code during the study timeframe. For these cases, the relevant 2012 six-digit NAICS code was applied.

Finally, in order to insure the robustness of our analysis, we made two relatively significant data cleaning decisions. First, although Gini coefficient calculations theoretically require a minimum of two cases, we exclude industry-years with fewer than five reporting facilities. This choice reflects a compromise between including in the analysis highly and consistently polluting industries with a small number of facilities, such as gold mining, and basing the disproportionality calculation on a meaningful number of underlying facilities. Second, we restrict our analysis to industries reporting in at least two-thirds of the reporting years (i.e. 10 out of 15 in the 1998–2012 period). These refinements result in a data set including 196,642 facility-year observations, representing 25,318 unique facilities from 322 industries. The majority of these fall in the manufacturing sector of the US economy, indicated by two-digit NAICS ranging from 31 to 33. Mining, Quarrying, and Oil and Gas Extraction (NAICS 21), Electric Utilities (NAICS 22), Wholesale Trade (NAICS 42), Waste Management and Remediation (NAICS 56), and Public Administration (NAICS 92) are also represented.

3. Results

3.1. Overall decline in RSEI hazard, 1998–2012

The overall context for our inquiry on disproportionality in facility-level toxic releases in US manufacturing is one of decline. Between 1998 and 2012, the total RSEI hazard generated by the facilities in our data set declined by 66 percent. This pattern was consistent across most industries. Approximately 74 percent of the industries analyzed saw a decline in RSEI hazard contribution between 1998 and 2012. Three effects contribute to causing such changes in aggregate pollution: scale effects, related to the overall growth or decline of an economy; composition effects, which
reflect changes in the industries contributing to an economy or economic sector; and technique effects, which describe changes in pollution per unit of production within an industry (Ochsner \textit{et al} 1995, Gallagher 2004). All three effects contributed to the overall decline in toxic releases in US manufacturing. First, US manufacturing declined between 1998 and 2012 (Pierce and Schott 2016). The total numbers of facilities in our data set reporting under the TRI program declined from the highest point of 14 692 in 2001 to 11 290 in 2012. In addition, overall employment in manufacturing declined from a high of 3 849 164 in 2001 to a low of 2 301 564 in 2012.

Second, the relative contribution of different industries to overall toxic releases also shifted, with some declining and others increasing. Finally, some of the decline in toxic pollution can be ascribed to facility-level pollution abatement actions, as evidenced by the handful of industries that saw increases in employment, coupled with overall declines in RSEI hazard (see figure 1).

### 3.2. Disproportionality—the rule, not the exception

Within the context of overall decline, disproportionality emerged as the characteristic pattern of toxic releases across facilities within US manufacturing industries throughout the study timeframe. Disproportionality in employee-normalized, facility-level RSEI toxicity was both remarkably high and stable over time. Of the 3 975 industry-years in the analysis, 44 percent are characterized by a disproportionality coefficient of above 0.9, and 87 percent are characterized by a disproportionality coefficient of 0.75 or above. Disproportionality ranged from a high of 0.995 in All Other Basic Inorganic Chemical Manufacturing (NAICS 325188) in 2001 to a low of 0.191 in Industrial Mold Manufacturing (NAICS 333511) in 1998. Averaging across the study period, 89 percent of industries are characterized by an disproportionality coefficient of 0.75 or above (figure 2).

These numbers are remarkable. For context, we calculated yearly disproportionality coefficients based on size-normalized facility-level toxic releases in US manufacturing as a whole, i.e. ignoring industry classifications. For this calculation, high disproportionality is to be expected, given the varied environmental footprints of distinct industries. Indeed, disproportionality in size-normalized facility-level toxic hazard in US manufacturing overall ranged from 0.991 in 1998 to 0.992 in 2012, with a low of 0.988 in 1999 and a high of 0.995 in 2002. Surprisingly, within-industry disproportionality comparing peer facilities is on par with manufacturing sector

![Figure 1](image-url)
disproportionality comparing unlike facilities. In other words, as comparisons are conducted among more (and more) homogenous groups, disproportionality remains high, if not higher, than it is across more diverse groups.

Analyzing the data longitudinally, we find that high within-industry disproportionality was stable over time—providing evidence that these patterns are longstanding, rather than recent. For the majority of industries, change in disproportionality was close to zero (figure 3). Comparing the years 2012 and 1998, about half (51 percent) of industries saw a decline and half (49 percent) saw an increase in disproportionality. A few industries did see significant changes in disproportionality between 1998 and 2012. The largest increase was 0.38 (Bitumous Coal Underground Mining, NAICS 212112 while the largest decrease was 0.49 (Carbon Black Manufacturing, NAICS 325182). Trends in changes in Gini coefficients over the entire timeseries within the 322 industries show results similar to the two end-year comparisons. We find that 54 percent of the 322 industries showed a declining trend in disproportionality over the study period, while 46 percent were characterized by an increasing trendline.

3.3. Polarization

The overall decline in RSEI hazard between 1998 and 2012 and the relative stability of industry-level disproportionality raises the question of where in the distribution of toxic pollution across facilities in US manufacturing reductions are being made. While Gini coefficients effectively indicate overall inequality across a distribution and how it changes over time, they do not distinguish between changes within a distribution in central tendency (location), density (shape) or some combination of both (Clementi et al 2017). In other words, does the stability in disproportionality over time mask changes in the extremes of the distribution? Drawing on metrics used to calculate polarization in income inequality (Morris et al 1994, Handcock and Morris 1998), we assessed polarization patterns in toxic releases in US manufacturing. An advantage of this method is the ability to distinguish between changes in comparative location, usually associated with shifts in the median, and changes in comparative shape, usually attributable to differences in variance and/or general asymmetries. We compared each successive year of log releases through 2012 to 1998 data, first location-matching values in subsequent distributions to values in the 1998 distribution. We then calculated changes in the median (to assess changes in location) and the three polarization indices (to assess changes in distributional shape).

We found a decrease in the distributional location, shown by a decrease in central tendency (figure 4). Notably, the additive decrease on the log scale is multiplicative on the real scale. We logged the non-normally, highly skewed data for interpretation purposes. We also calculated three shape indices—median relative polarization, lower relative polarization and upper relative polarization—which describe how the density of the distribution changes. These polarization indices range from −1 to 1, where negative values indicate greater polarization (comparative movement of density away from the center) and positive values indicate less polarization (comparative movement of density toward the center of the distribution). All three polarization indices were close to zero (figure 4). In
combination, these metrics indicate that reductions in toxic hazard between 1998 and 2012 were effected across the entire range of facilities, rather than being concentrated among the most polluting facilities.

3.4. Egregious polluters

The above patterns of disproportionality are the consequence of a small number of facilities, polluting far more than their peers. Figure 5 shows a histogram of the proportion of facilities, by industry and averaged across study years, that generates 90 percent of the industry’s total hazardous emissions. For 81 industries, 10 percent or fewer facilities generated 90 percent of industry-wide hazardous emissions on average and for 210 industries, 20 percent or fewer facilities generate 90 percent of the total industry

Figure 3. This is a histogram of average yearly change in Gini coefficient, measuring disproportionality in facility-level size-normalized toxic hazard by industry (defined by 6-digit NAICS). Change is measured as the slope of the trendline that characterizes the relationship between industry-level Gini coefficient and study year. It shows that most industries were characterized by little change in the notably high levels of disproportionality between 1998 and 2012.

Figure 4. This figure shows density functions for log transformed relative hazard generation distributions in 1998 (solid line) and 2012 (dashed line). Upon visual comparison there are two notable findings. First, the shape of these distributions are similar, reinforcing the null finding for the median, upper, and lower relative polarization metrics. Second, the location of the two distributions is different; the entire distribution in 2012 shifted down, reflecting the general decrease in toxic hazard in US manufacturing over the study timeframe.
hazard. While there were industries that show less extreme patterns of disproportionality, they were relatively few in number. There were no industries where 90 percent of facilities generated 90 percent of emissions; what might be considered the proportional case. The most disproportionally distributed industries in terms of pollution production were: All Other Basic Inorganic Chemical Manufacturing (NAICS 325188) where, on average, 1.0 percent of facilities accounted for 90 percent of the industry’s aggregate toxic burden; and All Other Plastics Product Manufacturing (NAICS 326199) where, on average, 2.0 percent of facilities accounted for 90 percent of the industry’s aggregate toxic burden. Conversely the least disproportionally distributed industries in terms of pollution production were: Carbon Black Manufacturing (NAICS 325182) where, on average, 58.4 percent of facilities generated 90 percent of industry-wide toxic releases; and Cut Stock, Resawing Lumber and Planing (NAICS 321912) where, on average, 56.4 percent of facilities generated 90 percent of industry-wide toxic emissions.

Focusing on individual facilities, of the over 25,000 establishments analyzed, 1116 facilities (4.4 percent) can be characterized as egregious polluters, defined as contributing 50 percent or more to their industry’s total yearly hazard. Only 31 facilities meet this threshold for at least ten of the fifteen study years. Setting the threshold for egregious pollution as a facility contributing 75 percent of more to their industry’s total yearly hazard results in 643 unique facilities (2.5 percent) and 21 consistent egregious polluters.

4. Conclusion

Debates over the relationship between economic activity and environmental harm default to a presumption of a proportional relationship between the two, due to both the continued influence of early environmental impact models and compelling historical evidence of economic growth matched by growth in environmental harm. Even aggregate toxic release data seem to support the proportional assumption. As the US manufacturing industry declined, so did toxic emissions. However, a closer analysis reveals that these aggregate patterns obscure the key feature of toxic releases in US manufacturing between 1998 and 2012, namely that a small number of facilities in each industry accounted for the vast majority of the toxic burden created by that industry. Of the 342 industries analyzed, 90 percent are characterized by an average disproportionality coefficient of 0.75 or above, and 39 percent are characterized by an average disproportionality coefficient of above 0.9. Ours is the first study to establish disproportionality in the production of toxic pollution across both a wide range of industries and over an extended period of time, and the results clearly challenge the presumption of proportionality between economic activity and environmental harm.

The findings have both scholarly and policy implications for the nature of pollution generation, for the role of targeting in environmental regulatory decision-making, and for efforts by environmental advocates. The scale and scope of EPA’s TRI data provide an exceptional resource for analyzing the relationship between economic activity and environmental harm,
because these data are available at the facility-level and provide a continuous record dating back to 1988. To scholars, we confirm that high disproportionality is a stable feature the yearly toxic contribution of facilities within most manufacturing industries over the fifteen-year study period, despite overall declines in toxic emissions. Our research indicates that scholarship on toxic pollution in the US needs to account for disproportionality in emissions. As such, the use of quantile functions, medians, and other transformations are important methodological decisions. The data also demonstrate the existence of egregious polluters, as a class of organizations distinct from peer facilities in their industries, raising questions about the organizational and social structures that account for their emergence and persistence. Third, our analysis suggests that toxic pollution can be studied not only at the level of individual facilities but also as a relational phenomenon across facilities within industries. It raises questions regarding the causes of differences in disproportionality across industries and across time. Finally, we see opportunities to comparatively analyze patterns of disproportionality across types of pollution. GHG pollution associated with electricity production facilities is proportional to facility size (Galli Robertson and Collins 2019), while methane releases show patterns of disproportionality similar to toxic emissions (Alvarez et al. 2018).

The research also has implications for environmental policy and politics. The high disproportionality of toxic hazard across establishments within industries suggests that a targeted policy approach, focused on a handful of egregious polluters, could dramatically reduce overall pollution burdens. Given the challenges of compliance with and enforcement of environmental regulations, data-driven targeting of egregious polluters within industry groups, and ideally within industries, could be an effective way to deploy limited resources. Inspections represent the central approach used by the US EPA’s Office of Enforcement and Compliance Assurance to compel regulatory compliance. In this era of tight budgets and constant pressures to contain government spending, it is essential that such limited resources be used cost-effectively. Moreover, having available the list of egregious polluters within industries can support environmental advocacy. For example, those who specifically work to enforce the Clean Air Act via legal actions might save valuable effort by pre-screening the landscape of polluters for egregious actors who could become the focus of advocacy action. It should be noted that the temporal volatility of the facility-level egregiousness classification complicates implications for policy and advocacy. An approach targeting disproportionate pollutants would have to be carefully designed.

In conclusion, the extremeness of disproportionality in toxic releases upends standard assumptions about how to limit the harm generated by industrial production. A majority of pollution can be mitigated simply by causing the small number of egregious polluters to pollute at average levels for their particular industry. In short, acknowledging disproportionality in relation to society’s impact on the environment is an opportunity for theoretical development, a research agenda and a call to action centered on identifying and targeting egregious polluters.

Data availability statement

No new data were created or analyzed in this study.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1534976. This work was also supported by the National Socio-Environmental Synthesis Center (SESYNC) under funding from the National Science Foundation DBI-1639145. The authors thank the two anonymous reviewers for providing constructive critiques of the manuscript. MC and SP jointly planned the research and wrote the manuscript. MC, SP, and DH conducted the data analysis. BM assisted with data quality control. The authors declare no competing interests. The authors jointly dedicate this publication to William Freudenburg, whose research inspired the analysis.

ORCID iDs

M Collins @ https://orcid.org/0000-0003-2286-8575
SPulver @ https://orcid.org/0000-0001-6988-5976
D Hill @ https://orcid.org/0000-0002-4073-6093
B Manski @ https://orcid.org/0000-0001-7879-3838

References

Alvarez R A, Zavala-Araiza D, Lyon D R, Allen D T, Barkley Z R, Brandt A R, Davis K J, Herndon S C, Jacob D J and Karion A 2018 Assessment of methane emissions from the US oil and gas supply chain Science 361 186–8
Ash M, Boyce J K, Chang G, Manuel P, Scoggins J and Tran J 2009 Justice in the Air: Tracking Toxic Pollution from Americas’ Companies to our States, Cities, and Neighborhoods (Amherst, MA: Political Economy Research Institute, University of Massachusetts)
Auchincloss A H, Gebreab S Y, Mair C and Drez Roux A V 2012 A review of spatial methods in epidemiology, 2000–2010 Annu. Rev. Public Health 33 107–22
Baer P 2009 Equity in climate–economy scenarios: the importance of subnational income distribution Environ. Res. Lett. 4 411
Barnatchez K, Crane L D and Decker R 2017 An assessment of the national establishment time series (NETS) database FEDS Working paper No. 2017–110
Berry L M 2008 Inequality in the creation of environmental harm: looking for answers from within Res. Soc. Probl. Public Policy 15 239–65
Boettcher J 1999 Challenges and opportunities presented by NAICS J. Bus. Finance Librariansh. 5 3–13
Cadle S H, Mulawa P A, Ball J, Donase C, Weibel A, Sagebiel J C, Knapp K T and Snow R 1997 Particulate emission rates from in-use high-emitting vehicles recruited in Orange County, California Environ. Sci. Technol. 31 3405–12
