Environmental Justice and the Clean Water Act: Implications for Economic Analyses of Clean Water Regulations*

Tihitina Andarge[1], Yongjie Ji[2], Bonnie L. Keeler[3], David A. Keiser[1], and Conor McKenzie[3]
[1]UMass Amherst, [2]Iowa State University, [3]University of Minnesota

May 19, 2023

Abstract

Since President Clinton's 1994 Executive Order 12898, federal agencies have been required to conduct environmental justice (EJ) analyses of federal rules and regulations. More recently, the Biden Administration has instituted several major efforts to reform regulatory review and promote a more equitable distribution of environmental benefits and burdens. This paper seeks to understand how prior guidelines have been implemented in federal regulatory reviews related to the Clean Water Act and provide a baseline for future studies of the distributional effects of clean water regulations. We reviewed 18 regulatory impact assessments relating to the Clean Water Act conducted since 1992. Only five of these studies conducted a quantitative analysis of distributional impacts and none of the 18 rules were determined to have disproportionately adverse effects on low-income or minority communities. Anticipating that future regulatory review will require more comprehensive distributional analyses, we combine national data on the location of all regulated point sources of water pollution with demographic characteristics to develop a baseline assessment of the distribution of water pollution facilities. Overall, we find that discharge locations tend to be located in areas that are poorer, have a higher White population share, and have less education. We find that rurality partly explains this pattern. The top 40% of census block groups in terms of rural population share contain almost all water pollution discharge locations. We conclude with a discussion of the policy implications of these analyses and suggestions for future work.

JEL Codes: Q50; Q52; Q53; Q56; Q58
Keywords: Environmental Justice, Clean Water Act, Surface Water Pollution

*We thank William Wheeler at EPA for helpful discussions, particularly related to the use of EPA datasets. We also thank Tatyana Deryugina, Matthew Kotchen, seminar participants at UMass Amherst, and Jonathan Adler and participants at the Clean Water Act at 50: An Interdisciplinary Symposium at Case Western Reserve University for helpful comments and suggestions.

1. Introduction

Environmental Justice (EJ) has featured prominently in the Biden Administration's policy initiatives. Two major efforts, the Justice40 Initiative and the White House's memo on Modernizing Regulatory Review, seek to greatly expand the federal government's role in promoting equity as it relates to environmental improvements and regulations (White House 2021, EPA 2022a).[1] The Justice40 Initiative seeks to direct 40 percent of the overall benefits of certain environmental federal initiatives to disadvantaged communities. The Modernizing Regulatory Review memo directs federal agencies to provide concrete suggestions for improving how regulatory review is performed with an eye on advancing social and racial equity while promoting regulations that promote traditional goals of economic growth and safeguarding public health and safety. Both initiatives were announced on President Biden's first day in office.

A large EJ literature supports the need to provide more equitable environmental policies and programs (Lee 2002, Mohai et al. 2009). Numerous studies have documented how low-income populations and communities of color in the U.S. are more likely to face greater exposure to air pollution (Wang et al. 2022, Tessum et al. 2021, Colmer et al. 2020), extreme heat (Benz and Burney 2021), flood risks (Tate et al. 2021) and hazardous waste facilities (Bullard et al. 2008). In the water sector, researchers have found disparities in the affordability and quality of drinking water (Mueller and Gasteyer 2021, Balazs et al. 2014), access to safe and reliable water distribution services (Dietz and Meehan 2019), and the enforcement of and compliance with clean water regulations (Konisky et al. 2021, Mueller and Gasteyer 2021).

In contrast to these other environmental stressors, only a handful of studies have investigated the distribution of facilities that discharge surface water pollution in the U.S. across social and demographic characteristics, and these have been at the state or regional scale (Wilson et al 2002, Son et al. 2021, Liévanos 2017). As a result, there is limited information on how the benefits of federal regulations that target major sources of water pollution are distributed. Since 1970, the U.S. has spent more on surface water pollution control programs than on any other environmental initiative (Keiser and Shapiro 2019). The economic impacts of these investments remain poorly quantified (Keiser et al. 2019), raising questions about both equity and efficiency implications of federal water quality policies.

This paper combines a qualitative assessment of how agencies have attempted to assess ex ante the distributional impacts of proposed Clean Water Act rules and regulations with a quantitative

[1] The Justice40 Initiative is part of a broader Executive Order 14008. In addition, the Biden Administration has promoted several related efforts, including 2021's Executive Order (EO) 13985 that promotes "Racial Equity and Underserved Communities Through The Federal Government", an additional EO in February 2023 on "Further Advancing Racial Equity and Support for Underserved Communities Through The Federal Government", and numerous other EOs on Diversity, Equity, Inclusion, and Accessibility (see for example, https://www.usaid.gov/equity/executive-orders-deia).

analysis of the distribution of polluting facilities in federal datasets. We conduct the first analysis to better understand prior EJ efforts as they relate to federal surface water quality regulations, which often vary by industry. We conduct the second analysis for two reasons. First, our analysis highlights particular industries where EJ concerns may be most prevalent, and thus, require more detailed focus in future industry specific rules. Second, prior EJ analyses performed by the EPA focus on the distributional consequences of proposed rules, not necessarily distributional differences in surface water pollution sources. Thus, our analysis complements these prior efforts to provide a more comprehensive picture related to surface water pollution.

We proceed by first characterizing how the US Environmental Protection Agency (EPA) has implemented federal guidance under Executive Order 12898 - Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations (1994) - to assess the EJ impacts of surface water pollution regulations. We searched government databases and reports from the Office of Internal and Regulatory Affairs (OIRA) for economic analyses associated with the Clean Water Act, compiling a dataset of 18 economic analyses of major water quality rules going back to 1992. For each economic analysis in our dataset, we reviewed how the agency evaluated any potential justice or distributional impacts. Agencies are required to assess equity and distributional considerations as mandated in Executive Order 12898, even if rules will have only positive or uniform effects on water quality. In our review, only five studies attempted to quantify distributional impacts and no rule was determined to have disproportionately high or adverse effects on low-income or minority communities. Those rules which did not perform quantitative analysis made their determinations based on the rules scope and effect. For example, a determination of no distributional effects was justified by including a statement that the rule was likely to have a limited effect on water quality or lead to general improvements in water quality that would not disproportionately burden certain communities.

The second step in our analysis assesses the distribution of industrial and municipal point source polluters across a range of demographic characteristics. Effluent standards for point source polluters are a cornerstone of the Clean Water Act, and thus a large fraction of economic analyses of the Act have focused on these emitters. To explore how the potential distribution of water quality benefits vary with point sources, we compile three main categories of data: 1) the location and operation status of nearly 700K point source polluters from 1990 to 2022, 2) socioeconomic information of residents living within close proximity to these polluters, and 3) information on the type and amount of pollution from these sources. We capture the relationship between the number of pollution discharge locations and demographic data using pseudo-Lorenz curves, calculate related Gini coefficients, and employ cross-sectional models to assess relationships between point source variables and demographic characteristics. Although data limitations prevent us from examining the economic damages associated with this pollution, our analysis provides one of the first national pictures of how surface water pollution sources vary with demographics and socioeconomic characteristics.

Our results show that facility outfall locations, or discharge points, tend to be located in areas where a greater fraction of the population is White, living below the poverty threshold, and without a college degree. Over time, we find narrowing differences in the distribution of water pollution outfalls by race, income, and education between census block groups with and without permitted facilities. We also find that most outfalls are located in rural areas; the top 40% of census block groups in terms of rural population share contain nearly all outfalls. When we examine the distribution of outfalls within rural areas, we find less unevenness in the distribution of outfalls suggesting that rurality may partly explain the overall pattern that we observe across all facilities.

We find that the presence of outfalls varies across industrial sectors. Pollution outfalls from industrial sectors such as mining, manufacturing, and wholesale trade are more likely to be located in areas with higher poverty and lower levels of education. We observe a similar pattern when we focus on industries within manufacturing that have more toxic discharges. When conditioning on rurality, we observe similar patterns across industries for poverty and education, though the distributions are more even. For some industries, we find different results for race. For example, in rural areas, we find that outfalls from manufacturing and transportation and communications tend to be concentrated in areas where a greater share of the population is non-White. In rural areas, we also find that outfalls from facilities with more toxic discharges are located disproportionately in census block groups with higher non-White population shares. These findings suggest that when we look across the US as a whole, water pollution outfalls tend to be located in areas with higher poverty and lower levels of education, but also in areas with a greater share of the population that is White. However, when we focus more narrowly on the rural areas of the country where most outfalls are located, the relationships with poverty and education remain qualitatively similar, but there are some important differences with respect to the share of the population that is non-White.

Overall, our findings reaffirm the need to consider distributional consequences of water pollution regulations and highlights particular industries where additional attention may be warranted. We conclude the study with recommendations to guide future assessments.

2. EJ Benefits in US EPA Analyses and Context within EJ Literature

2.1. Literature

Executive Order (EO) 12898 - Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations, was signed into law by President Clinton on February 11, 1994 with the following stated purpose:

> *To the greatest extent practicable and permitted by law . . . each Federal agency shall make achieving environmental justice part of its mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of its programs, policies, and activities on minority populations and low-income populations in the United States and its territories.*

In addition to EO 12898, federal guidance on benefit-cost analysis also permits consideration of distributional impacts. The Office of Management and Budget's (OMB's) guidance to Federal agencies on the development of regulatory analysis known as Circular A-4 (2003) includes the following guidance for agencies in assessing distributional effects of proposed rules or regulations:

> *Where distributive effects are thought to be important, the effects of various regulatory alternatives should be described quantitatively to the extent possible, including the magnitude, likelihood, and severity of impacts on particular groups…Your analysis should also present information on the streams of benefits and costs over time in order to provide a basis for assessing intertemporal distributional consequences, particularly where intergenerational effects are concerned.*

Together, both 12898 and Circular A-4 affirm the importance of assessing distributional impacts of regulatory policies. Circular A-4 provides slightly greater methodological detail, noting that a quantitative analysis should be used where possible and assessments should include information on the magnitude, likelihood, and severity of impacts over time. Both guidance documents stop short of prescribing specific methodologies, leaving agencies to make their own determination of affected populations, appropriate comparison groups, and whether any observed disparities count as "disproportionately high and adverse".

Previous scholars have reviewed the implementation of federal guidelines for EJ analysis across multiple federal agencies. Vajjhala, Van Epps, and Szambelan (2008) found large gaps in the information required for effective analysis of potentially differential impacts on minority and low-income populations. A 2002 report by the National Environmental Justice Advisory Council (NEJAC) found that implementation of EO 12898 varied greatly across agencies with no consistent framework applied across agencies or regulations (NEJAC, 2020). A USEPA

Inspector General report released 10 years after the signing of EO 12898 found that the agency had yet to consistently integrate EJ into its day-to-day operations, had failed to take the critical initial steps to define the populations covered by EO 12898, and to develop criteria for determining disproportionate impacts (USEPA 2004).

Clearly there is a perceived disconnect between the mandate to consider EJ impacts in federal regulatory analysis and the implementation of this guidance in practice. Earlier reviews looked broadly across multiple agencies, focused on identifying the frequency of key terms such as "disproportionately high", "distributional consequences", and "achieving environmental justice" but not the actual methodologies or substance of any distributional analysis, and did not include more recent regulatory impact assessments in their review. Here, we build on these past studies with an explicit focus on regulations under the Clean Water Act.

2.2. Methods

We investigated rules and regulations associated with the CWA. Agencies are tasked with promulgating regulations that will implement statutes. Regulations may comprise multiple individual rules. We compiled all major rules and regulations related to the CWA by searching public reports, including the Reports to Congress on the Costs and Benefits of Federal Regulations (1995–2009), Costs of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities (2010–2014), and the Annual Reports to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance with the Unfunded Mandates Reform Act (2015–2020). These reports summarize significant regulatory activities for Congress and therefore should identify and report on any new or modified rule or regulation. We further supplemented this review through conversations with EPA staff economists. The development of regulations is governed by a rule-making process, including a notice of proposed rule-making and a public comment period. Agencies may produce analyses at each stage of the rule-making process. Here we focused only on final rules (Appendix Table 1) but we included each rule's supporting technical, economic, and environmental analysis documents.

For each of the 18 Clean Water rules in our database, we document the approach that EPA took to assess any potential distributional impacts of the rule or regulation. As noted above, EPA is required to assess if programs, policies, and activities will have "disproportionately high and adverse human health or environmental effects" on minority populations and low-income populations in order to comply with EO 12898. We reviewed each regulatory assessment for text related to EO 12898 as well as supporting materials or other analyses conducted as part of EPA assessments of the benefits and costs of water quality regulations.

Appendix Table 1 lists the 18 rules related to the Clean Water Act. For each rule, we examine the demographic characteristics used by EPA to assess disproportionate impacts, the spatial scale of the analysis including the selection of affected population and reference populations, the specific types of water quality benefits assessed, and the final determination of any EJ impacts. In many cases, the text of the rule contained little information on the actual analytical procedures used to determine distributional impacts and required us to look through supporting documentation in search of methodological details.

2.3. Results

Of the 18 rules in our dataset, no rule was deemed by EPA analysts to have disproportionately high or adverse human health or environmental effects on minority or low-income populations. In the majority of assessments, this determination was made because the proposed rule or regulation would be uniformly applied to all facilities or that regulations would improve environmental quality for all residents.

For example, the Effluent Limitations Guidelines and New Source Performance Standards for the Meat and Poultry Products (2004) includes the following brief text regarding environmental justice impacts under EO 12898:

> *EPA has determined that this rulemaking will not have a disproportionate effect on minority or low-income communities because the technology-based effluent limitations guidelines are uniformly applied nationally irrespective of geographic location. The final regulation will reduce the negative effects of meat and poultry products industry waste in our nation's waters to benefit all of society, including minority and low-income communities. The cost impacts of the rule should likewise not disproportionately affect low-income communities given the relatively low economic impacts of the rule.*

Five of the rules included a quantitative analysis of distributional impacts (Table 1). In the 2015 Effluent Limitations Guidelines and Standards for Steam Electric Power rule (RIN 2040–AF77), the EPA evaluated the demographic characteristics of individuals living in proximity to steam electric facilities, individuals served by public water systems downstream from steam electric facilities, and populations exposed to steam electric power facility wastewater through consumption of recreationally caught fish. EPA found disparities between the affected population and state or national averages and determined that regulatory options that increase pollutant exposure compared to the baseline may disproportionately affect minority and low-income communities. However, the estimated changes in exposure between the baseline and regulatory options were small and EPA determined that these small changes in risk did not meet the criteria of disproportionately high and adverse effects.

The majority of rules (13 out of 18) included some mention of race and income as the demographic variables of interest. A few studies also considered indigeneity, national origin, effects on children, and impacts on subsistence anglers. However, demographics, as well as affected and comparative populations, were not consistent across analyses and the choices made regarding which variables and populations to include were not entirely transparent. A few rules compared the demographics of the counties containing affected facilities with state averages. Other rules used a proximity analysis to estimate downstream populations. In the 2004 Final Regulations to Establish Requirements for Cooling Water Intake Structures at Phase II Existing Facilities, EPA calculated the poverty rate and the percentage of the population classified as non-White for populations living within a 50-mile radius of each of the 543 in-scope facilities and compared these rates with national averages. A 2014 rule on cooling water intake structures also used a 50-mile radius from a regulated facility to assess affected individuals, and included any anglers who live outside of the 50-mile facility buffer but within a 50-mile radius of the river segments, or river reaches, nearest to the facilities.

For the rules that did assess potential benefits of regulation (or forgone benefits) the most commonly assessed benefits were impacts on subsistence fishing, cancer risks from exposure to toxic chemicals, and general health impacts. Beyond subsistence fishing for tribal communities, the analyses did not mention potential impacts on cultural values, recreational values, or other non-material benefits of clean water.

In summary, in our review of the implementation of EO 12898 in rules and regulations related to surface water pollution under the Clean Water Act, we observed inconsistency in the scale and scope of analysis of distributional impacts. A majority of rules did not conduct a quantitative analysis of impacts, and no rules were determined to have "disproportionately high and adverse impacts." These findings are consistent with other recent publications investigating the consideration of distributional and equity considerations in federal regulatory review across a broad range of agencies and policy domains (Robinson et al. 2016, Cecot and Hahn 2022, Revesz and Yi 2022).

3. Environmental Justice (EJ) Analysis of Point Source Polluters - Methods and Data

As noted above, President Biden has made EJ a priority of his administration. Indeed, proposed updates to the guidance documentation for analyzing the costs and benefits of regulations known as Circular A-4, were recently released by the Office of Management and Budget (White House 2023). The revised guidance includes an expanded section on proper procedures for conducting distributional analyses, including the importance of placing any proposed rules or regulations in context based on an assessment of baseline distributions of environmental benefits or burdens. The OMB guidance also states that it is "not sufficient for your analysis to merely state that the chosen alternative does not make relevant groups worse off; it is important to analyze and describe the benefits and costs of different regulatory alternatives for different groups." Given the limited nature of EJ analyses in prior regulatory reviews under the CWA, we present a baseline assessment of the distribution of regulated water pollution facilities that could inform future applications of distributional analyses of clean water rules or regulations.

3.1. Methods

Our EJ analysis is descriptive in nature. We do not intend to describe causal relationships between socioeconomic characteristics and pollution. Rather, as an important first step, we seek to describe how the location of pollution outfalls varies with socioeconomic characteristics and demographics and leave important questions of causality to future studies.[2]

Our analysis implements two empirical approaches. First, we construct pseudo-Lorenz curves that provide a visual representation of the relationship between water pollution sources and several measures of demographics and socioeconomic characteristics.[3] These curves further help us describe how these relationships have changed over time. We call these pseudo-Lorenz curves as Lorenz curves traditionally focus on income distributions. However, we find that this representation enables an easy means to visualize how water pollution sources are distributed within society. Our primary analyses focus on the number of water pollution outfalls, or discharge points, from regulated water pollution dischargers. We focus on facilities and their outfalls given that this information is more consistently reported over space and time than other measures. In supplementary analyses, we examine how outfalls may (imperfectly) reflect the quantity and toxicity of pollution. We link each outfall to its census block group given that this is the finest spatial unit for which we have available demographic information.

To construct our pseudo-Lorenz curves, we first rank the communities (2010 census block groups) according to a particular socioeconomic variable (i.e., a measure of race, income, education, etc.). If there is more than one census block group with the same value, we rank those census block groups with the lowest number of outfalls (or other measure of pollution) first. This ranking forms the variable for our x-axis. We normalize the x-axis ranking from 0 to 1. Once we have ranked census block groups, we calculate the cumulative distribution value of outfalls (or other measure of pollution) from 0 to 1. To form our pseudo-Lorenz curves, we plot the pair of calculated x and y values for each census block group. Additionally, for each pseudo-Lorenz curve, we compute the Gini coefficient by calculating the area between the 45-degree line and the pseudo-Lorenz curve.

[2] In a similar vein, except for conditioning on rurality, all of our analyses are unconditional, meaning that we examine the pure correlation between pollution outfalls and demographic information without controlling for other factors. We believe this provides important information on the current distribution of outfalls across space, but does not attempt to explain why the variation in outfalls arises or whether this distribution is unequal when conditioning on other factors. For example, in the air pollution literature, some studies include factors such as population density, employment patterns, and land use to explain observed pollution exposure patterns (Mennis 2003, Ash and Boyce 2018). We leave these important questions with respect to surface water pollution to future work.

[3] Lorenz curves or variations of Lorenz curves have been used in the environmental justice literature to depict the distribution of a given outcome along demographic and socioeconomic lines (Mehta et al. 2014, Mohammed et al. 2021, Sheriff 2023).

In addition to these curves, we estimate complementary cross-sectional models that quantify the magnitude of the relationship between the presence of outfalls and these demographic and socioeconomic characteristics by estimating the following equation:

$$y_i = \alpha + \beta D_i + \varepsilon_i \tag{1}$$

where y_i is an indicator for whether an outfall is present in census block group i, D_i is one of the demographic or socioeconomic characteristics we consider, and ε_i is the error term. In supplementary analyses, we replace y_i with other measures of water pollution that we discuss below. To account for potential correlation in our error term within geographic areas, we cluster standard errors at the county level.

3.2. Data

Our analysis uses three primary sources of data. Two of these sources provide information on the location and discharge of pollution from point sources in the U.S. The third source provides information on residents within close proximity to these sources.

Point Source Location and Operating Status

We combine several data sets from US EPA's Enforcement and Compliance History Online (ECHO) database and the Integrated Compliance Information System - National Pollution Discharge Elimination System (ICIS-NPDES) to construct a data set of outfalls active during the 1990 - 2022 period. We use the permit issue dates to pinpoint when a facility first becomes active.[4] If a permit is missing an issue date, we drop that observation. To determine when a facility becomes inactive, we use the permit termination date if the permit's status is "terminated" and the maximum of the termination, retirement, and expiration dates if the permit's status is "not needed." If a permit should have an end date but has missing values for the termination, retirement, and expiration dates, then we drop that observation. This provides us with information on 761,905 of 814,130 facilities in the U.S.

We obtain the geographic coordinates for a facility's outfall(s) from ICIS-NPDES Discharge Points. We use the coordinates to geolocate each outfall inside a census block group using the 2010 Tiger/Line Block Group shapefile. If an observation is missing outfall coordinates, we instead use the facility coordinates, which we supplement from ICIS-NPDES Facilities and the Facility Registry Service (FRS), and assume it has a single outfall at those coordinates. If a permit is missing both the outfall and facility coordinates, then we drop that observation. Additionally, we drop facilities outside of the conterminous US and off-shore facilities as we are

[4] Specifically, we define the date of first activity as the minimum of the original issue date and issue date.

not able to locate them inside census block groups. After this process, we are left with information on 687,788 facilities (863,511 outfalls).[5]

To determine which industrial division a facility belongs to, we use information on the industrial codes reported with the permits. We place facilities into 11 broad industrial categories based on their two-digit Standard Industrial Classification (SIC) codes: (1) Agriculture, Forestry, and Fishing, (2) Mining, (3) Construction, (4) Manufacturing, (5) Transportation and Communications Services, (6) Electric, Gas, and Sanitary Services, (7) Wholesale Trade, (8) Retail Trade, (9) Finance, Insurance, and Real Estate, (10) Services, and (11) Public Administration (see Appendix Table 2 for descriptions of each industrial division). We place a facility in one of these categories if any one of its two-digit SIC codes belong to that category.[6] Of the 687,788 facilities (863,511 outfalls), we are able to identify the industrial category for 385,613 facilities (551,072 outfalls). Since a sizable portion of facilities in our sample are missing industrial codes, focusing only on facilities with information on their SIC codes may lead to sample selection bias if missing values are nonrandom. Therefore, we conduct analyses on both the full sample and the subsample containing facilities with SIC information. Additionally, we identify publicly owned treatment works (POTWs) using a facility type indicator associated with each permit and remove these facilities from the 11 industrial categories. We examine these facilities separately given their major point of emphasis within the Clean Water Act. There are 18,168 POTWs with 29,289 outfalls in our sample. Lastly, we identify major dischargers in the pooled sample and in each industry.[7]

Effluent Data

In addition to the location, operating status, and industrial division of facilities, we obtain estimates of flow from ICIS-NPDES Discharge Monitoring Reports (DMRs). DMRs must be regularly submitted by facilities with permits that require sampling and monitoring, which tend to be large point sources, standard industrial dischargers, and POTWs.[8] The DMRs contain

[5] Dropping facilities without start and end dates and geographic coordinates may lead to sample selection bias if incompleteness of this information is nonrandom.

[6] Facilities may belong to multiple industrial sectors. Thus, a given facility may belong to more than one of these 11 categories. We place facilities into industrial categories largely along the lines of the SIC divisions. However, we split the Transportation, Communications, Electric, Gas, and Sanitary Services industrial division into two separate sub-divisions given the importance of the Electric, Gas, and Sanitary Services as major sources of water pollution.

[7] US EPA typically designates major sources as POTWs with a total design flow greater than 1 millions of gallons per day (MGD), industrial sources with a score greater than 80 on the NPDES Permit Rating Worksheet, or sources designated as “major” by the regulator. Note, the NPDES Permit Rating Worksheet takes into account six factors to determine the score: toxicity potential, flow, amount of conventional pollutants that would be discharged, public health impact, characteristics of the receiving stream and potential for violation of water quality standards, and proximity to coastal waters.

[8] Most large point sources, but not all, are required to submit DMRs. Most standard industrial dischargers (i.e., those that discharge directly to surface waters), POTWs, and major facilities in the municipal stormwater subprogram regularly submit DMRs. Some industrial stormwater facilities, such as those regulated under EPA's Multi-Sector General Permit, regularly submit DMRs. In general, most construction stormwater facilities and non-major municipal stormwater facilities are not required to submit DMRs. A small percentage of these facilities may have to submit DMRs due to a violation and subsequent enforcement action. CAFOs submit DMRs quite irregularly (e.g.,

outfall-level information on flow and the quantity and concentration of the discharged pollutants. Of 687,788 facilities, DMRs are available for 68,608. We use these data to determine the amount of flow, in millions of gallons per day (MGD), discharged from an outfall as an imperfect proxy for the quantity of pollution recognizing that higher flow does not necessarily correspond to higher pollution levels. There are several types of measures, called statistical base codes in ICIS-NPDES, for flow (e.g., 1 day geometric mean, 12 day average, 12 month average, etc.). We focus on monthly averages as this is one of the most frequently reported types of flow measures. Facilities calculate monthly averages by taking the average of all flow measurements at each outfall during a calendar month and include it in their DMRs.[9] We take the mean of monthly average flow at the census block group level. As a separate check on this measure of flow, we use the volume and percentile of wastewater discharge estimated by the Climate and Economic Justice Screen Tool.[10]

Socioeconomic Data

Our socioeconomic and demographic data are from the 1990, 2000, 2010 censuses, and 5-year American Community Survey (ACS) estimates of 2013 and 2019. Since the census block group level demographic information from ACS is based on the 2010 census boundary, we normalize all demographics to 2010 census block groups using the Integrated Public Use Microdata Series National Historical Geographic Information System project (IPUMS NHGIS).[11] In cases where there are no harmonized data series for our variables of interest, we use crosswalk matrices between 1990, 2000, and 2010 census boundaries to construct our own harmonized series. The socioeconomic variables we use are the share of the population that is non-White, the share of the population that is Black or African American, the share of the population that is Hispanic, the share of individuals with income lower than the federal poverty line, the share of the population aged 25 and above without a college degree or higher, and the share of population that is in a rural area. Given the number of possible variables, we limit our main analysis of pseudo-Lorenz curves to the share of the population that is non-White, the share of the population below the poverty line, the share of the population without a college degree or higher, and the share of the population living in a rural area. This enables us to focus on measures that reflect some racial differences, income differences, education differences, and urban/rural differences. In the regression analyses, we further explore differences across additional measures of race and ethnicity (share Black or African American, share Hispanic).

Matching Point Sources to Demographic and Socioeconomic Data

after a major storm event). For more information, see section 2 of EPA's Economic Analysis of the National Pollution Discharge Elimination System (NPDES) Electronic Reporting Final Rule.

[9] The number of measurements required may vary across permits.

[10] We downloaded version 1.0 from the "Downloads" section of the following website: https://screeningtool.geoplatform.gov/en/#3/33.47/-97.5.

[11] The IPUMS NHGIS website provides additional information (https://www.nhgis.org).

The availability of demographic and socioeconomic data at the census block group level is somewhat limited temporally which prevents us from constructing an annual panel data set of matched outfalls and demographics. For this reason, we match outfalls to demographic and socioeconomic data as follows: (1) outfalls active during 1990 - 1999 are matched to 1990 census data, (2) outfalls active during 2000 - 2012 are matched to 2000 census data, (3) outfalls active during 2013 - 2018 are matched to 2013 ACS data, and (4) outfalls active during 2019 - 2022 are matched to 2019 ACS data.[12]

Additional Sources of Data

The Toxic Release Inventory (TRI) database contains facilities that emit toxic chemicals with highly adverse health and environmental impacts. To account for potential toxicity-related differences in the distribution of outfalls along socioeconomic characteristics, we separately examine water pollution facilities that also appear on the TRI. Additionally, as the presence of an outfall need not necessarily result in poor water quality, we perform additional analyses using measures that capture some ambient water quality conditions. Section 303(d) of the CWA requires states to construct a list of impaired waters which do not meet state-established water quality standards. We obtain information on the total impaired stream length (km) with any listed CWA 303(d) impairment from the EPA's EnviroAtlas database.[13] As this information is provided at the subwatershed level, we aggregate the census block group level socioeconomic variables to the subwatershed level.[14] Then, we construct the pseudo-Lorenz curves by ranking subwatersheds rather than census block groups.

Data Summary

Figure 1 displays the location of all outfalls used in our analysis. Our data provide broad coverage of water pollution facilities across the US. Table 2 provides summary statistics from our most recent time period. In this period, there are over 600K outfalls at active facilities in approximately half of all census block groups. We have data on the industrial code for approximately two-thirds of these outfalls. Of these records, construction, manufacturing, mining, and electric, gas, and sanitary services have the most outfalls. In these summary statistics, we see that census block groups with facilities across all industries tend to have a lower non-White population than census block groups without permitted facilities. Compared to census block groups without permitted facilities, the share without a college degree in census block groups with permitted facilities is slightly higher for about half of the industries and substantially

[12] Unfortunately, the 2010 and 2020 censuses do not provide poverty and education data at the census block group level which prevents us from using decades starting with census years to construct our panel. As a result, we rely on ACS data from 2013 and 2019 and adjust our time periods accordingly.

[13] EnviroAtlas is an online tool developed by the USEPA that provides geospatial data and other information on the nation's ecosystems and their services. It allows users to explore and analyze environmental and socio-economic factors at various scales. We obtained impaired waterway length from the September 2021 version of EnviroAtlas. For more information about downloading the dataset, please visit the following website: https://www.epa.gov/enviroatlas/forms/enviroatlas-data-download.

[14] The subwatershed is identified by the 12 digit hydrologic unit code (HUC 12).

higher for the other half. The share below the poverty line is higher in some industries and lower in others compared to the average for the conterminous US (CONUS). The biggest difference we see in these summary statistics is the share of the population that is in a rural area. In our data, census block groups with facilities are more likely to be classified as "rural" relative to CONUS averages.

4. EJ Analysis of Point Source Polluters - Results

4.1. Pseudo-Lorenz Curves and Gini coefficients

We begin by reporting the pseudo-Lorenz curves for the count of outfalls. Figures 2 and 3 display these results. We present results with the share of the population that is non-White, the share of the population below the federal poverty line, the share of the population without at least a college degree, and the share of the population living in rural areas. Each figure displays how the distribution of outfalls has changed over time, from the 1990-1999 to 2019-2022 time periods. Gini coefficients are reported in parentheses in the legend, next to the line for the respective time period. The 45-degree line is a reference point for an equal distribution of outfalls. The way in which we define our indicator variables is such that any curve to the right of this 45-degree line suggests that these outfalls are disproportionately located in areas that are historically underserved or potentially disadvantaged. Similarly, a positive Gini coefficient indicates that outfalls are disproportionately located in such areas; a larger Gini coefficient suggests a greater concentration of outfalls in those areas.

In Figure 2, we summarize these curves across all industrial sectors for each indicator. As described in Section 3.1, we first rank census block groups according to the particular demographic variable of interest. A normalized ranking from 0 to 1 serves as our x-variable. We then graph the cumulative distribution value of outfalls of Y on the vertical axis. For example, Figure 2, Panel b) shows that in the 1990-1999 time period, the first 40 percent of census block groups ranked in terms of poverty (i.e., the richest 40 percent of census block groups) contain approximately 20 percent of the total number of outfalls. The figure also shows that the poorest 20 percent of census block groups (moving from 0.8 to 1.0 on the x-axis) contain approximately 30 percent of the total number of outfalls. We find that census block groups with greater numbers of outfalls tend to be located disproportionately in areas with a higher share of the population that is White, in areas with a greater share of the population that is below the poverty line, and in areas with a greater share of the population that does not have at least a college education. However, the variation in outfalls has smoothed out over time for race, poverty, and education. The curves for the 2013 - 2018 period (the green line) and the curves for the 2019 - 2022 period (orange line) are both very close to the 45-degree line. Note, part of this change may be due to improved data reporting over time. In particular, the 2015 Electronic Reporting Rule drastically changed state reporting requirements to EPA for certain facilities potentially improving data quality.[15] Thus, while we continue to discuss changes over time, we emphasize that these changes may be attributable to changes in data quality rather than changes in the distribution of outfalls.

Outfalls tend to be located disproportionately in areas with a higher rural population share. In fact, in the first 50 percent of census block groups in terms of rural population share (i.e., the least rural 50 percent of census block groups), there are zero outfalls which is why the pseudo-Lorenz curves in Figure 2(d) are horizontal at the lower end of the support. Contrary to the findings for race, poverty, and education, the tendency for outfalls to be located disproportionately in rural areas has persisted over time. In all time periods, an overwhelming number of outfalls are located in the top 40 percent of census block groups in terms of rural population share. Given this finding, we further examine the distribution of outfalls across race, poverty, and education among the top 40 percent of census block groups in terms of rural population share. Figure 3 presents pseudo-Lorenz curves for these three indicators, conditioning on rurality. One notable finding is that the pseudo-Lorenz curves for the share of the population that is non-White is closer to the 45-degree line (i.e., the Gini coefficients are closer to 0) indicating that rurality may partly drive the pattern we observe when including all census block groups. In fact, during the 2013 - 2018 and 2019 - 2022 periods, the distribution of outfalls is almost completely even in terms of the share of the population that is non-White. We observe a similar pattern for the share in poverty and share without a college degree, though the difference between the Gini coefficients for all census block groups and more rural census block groups is less stark during earlier periods.

While Figures 2 and 3 provide an overall snapshot of all outfalls, they may mask industry-specific heterogeneity which may be important as the type, toxicity, and quantity of discharges varies across industries. For example, there is substantial heterogeneity in the amount of TRI chemicals discharged across industries with electric, gas, and sanitary services; services, finance, insurance, and real estate; and manufacturing accounting for most TRI discharges. Similarly, when examining discharges of CWA priority pollutants, a list of 126 pollutants that are deemed to be toxic, facilities in manufacturing or electric, gas, and sanitary services are responsible for most of the discharges. As industrial patterns in the toxicity and amount of discharge exist, it may also be the case that the relationship between outfalls and demographic and socioeconomic characteristics is heterogeneous across industries. We also explore heterogeneity by industry to help inform future rules or regulations that may target potential industries.

[15] Part of this change may be due to improved data reporting over time. In particular, the 2013 - 2018 and 2019 - 2022 time periods coincide with the implementation of the 2015 Electronic Reporting (eReporting) Rule by EPA. This rule changed state reporting requirements to EPA for non-major facilities. Prior to this rule, states were required to report only basic facility information about non-major facilities to EPA. Some states did, however, voluntarily report on non-majors even before the implementation of the rule. Given that the universe of permits overwhelmingly consists of non-major facilities, the implementation of this rule potentially improves overall data quality.

In Figures 4 - 7, we display bar graphs depicting Gini coefficients across different industrial sectors for our four main demographic and socioeconomic variables. Each time period is represented by the same color as in previous figures. For each time period, the width of the bar corresponds to the magnitude of the Gini coefficient with positive Gini coefficients appearing to the right and negative Gini coefficients appearing to the left of 0, which is delineated by a vertical dashed line. In each bar graph, we order industries by the average size of the Gini coefficient so that industries most concentrated in areas with a higher share of traditionally underserved populations appear first. In these figures, we also show the Gini coefficient across all industries as a useful point of comparison. As a reminder, this category includes outfalls regardless of whether there is a reported industrial classification.

We first examine how outfalls are distributed across industries by the share of the population that is non-White (Figure 4). As with Figure 2, we see a similar pattern across most industrial classifications that outfalls tend to be located in areas with a greater fraction of the population that is White (i.e., negative Gini coefficients). We focus on (1) construction, (2) electric, gas, and sanitary services, (3) manufacturing, (4) mining, (5) public wastewater treatment, and (6) wholesale trade given the large number of these facilities (Table 2). Three of these industrial categories, mining; public wastewater treatment; and electric, gas, and sanitary services, show a large number of outfalls in areas with a larger share of the population that is White, even more so than across all sectors. Contrary to the Gini coefficients computed using all sectors, the distribution for these industries has remained almost unchanged over time.

Turning to Figure 5, we examine how these outfalls vary by industrial classification and by the share of the population below the poverty line. Many industries follow a similar pattern that we see in the average across all sectors. However, mining, wholesale trade, and manufacturing in particular have more outfalls in areas with higher poverty. Unlike the overall picture, certain industries such as mining, public wastewater treatment, and wholesale trade have changed relatively little over the study period. In contrast, industrial divisions such as finance, insurance, and real estate; construction; public administration; and services exhibit changes over time and are even slightly concentrated in lower poverty areas during later time periods. A similar story appears in Figure 6 where we display these results by industrial sector for the share of the population without a college degree. Similar to the overall picture, outfalls are located predominantly in areas with a larger share of the population without a college degree. This appears to largely be driven by outfalls from mining; wholesale trade; agriculture, forestry, and fishing; manufacturing; public wastewater treatment, and electric, gas, and sanitary services. This pattern has largely persisted over time for these industrial divisions. These results highlight these industries as potential focal points for future, or more in-depth, EJ analyses by the EPA and academic researchers. For the remaining industrial divisions, the distribution of outfalls has become more evenly distributed over time. Interestingly, for construction, outfalls are slightly concentrated in more educated areas during the latest two time periods.

Lastly, we examine how outfalls are distributed by the rural population share. Similar to the overall results, we find that outfalls are heavily concentrated in areas with a higher rural population share for all industries and this pattern is highly persistent over time, regardless of industry (Figure 7). This prompts us to consider the distribution of outfalls across race, poverty, and education within more rural areas. In Appendix Figures 1 - 3, we present bar graphs with Gini coefficients for the distribution of outfalls across dimensions of race, poverty, and education in the top 40 percent of census block groups in terms of rural population share. Generally, even among rural census block groups, outfalls tend to be located in areas that are predominantly White, though the Gini coefficients are much smaller in magnitude indicating a relatively more even distribution (Appendix Figure 1). For manufacturing and transportation and communications, however, the Gini coefficients are positive but small indicating that outfalls from those industrial divisions somewhat tend to be located in areas with a higher share of the population that is non-White. For poverty and education, the overall industrial patterns largely hold though the Gini coefficients are somewhat smaller (Appendix Figures 2 and 3). For three of the industries, construction; finance, insurance, and real estate; and public administration, outfalls are slightly concentrated in areas with lower poverty and a lower share of the population without a college degree. Interestingly, for mining and agriculture, forestry, and fishing, the Gini coefficients for the share in poverty are larger when conditioning on rurality. These results underscore the need for analyses that separately examine different industries; the distribution of outfalls across demographics and socioeconomics is heterogeneous.

As we demonstrate, there are some differences in locational patterns across industries. This is important as certain industries tend to discharge more toxic pollutants with greater potential harm to the surrounding community. The literature identifies two major industrial groups within the manufacturing industrial division as having a higher potential for harm: (1) Chemicals and Allied Products and (2) Petroleum Refining and Related Industries (Liévanos et al. 2017). In Appendix Figures 4 and 5, we present pseudo-Lorenz curves for the distribution of outfalls from these two industries for our four main demographic and socioeconomic variables. In general, the patterns are similar to what we observe for all industries. The results only diverge for the distribution of outfalls from Chemicals and Allied Products along racial lines; the pseudo-Lorenz curves are very close to the 45-degree line indicating a relatively even distribution. Once we condition on the top 40 percent of census block groups in terms of rural population share, however, outfalls from this industry are located disproportionally in areas with a higher share of the population that is non-White (Appendix Figure 6). For Petroleum Refining and Related Industries, however, we find that outfalls are more evenly distributed across racial lines once we condition on rurality (Appendix Figure 7). Again, this highlights the importance of performing industry-specific analyses as ignoring these differences may mask potential EJ concerns.

We turn next to other measures of the presence of pollution. In Figure 8, we display bar graphs depicting Gini coefficients for monthly average flow, outfalls at facilities deemed “Major” by the USEPA, the number of permitted dischargers that also appear on the TRI, and the total length of impaired waterways within a HUC12 region.[16] Panel (a) presents the results for the non-White population share, panel (b) for the share below the poverty line, and panel (c) for the share without a college degree. When using monthly average flow as our measure of pollution, the results appear similar to the number of outfalls, with more equal distribution related to poverty. The results for “Major” facilities are largely consistent with the overall number of outfalls. Outfalls from facilities that appear on the TRI, however, are less prevalent in areas with a greater share of the population that is White compared to outfalls overall. In terms of poverty and education, consistent with the overall number of outfalls, outfalls from facilities that appear on the TRI are more concentrated in areas with higher poverty and lower education though this pattern is more persistent over time for this measure of pollution. Lastly, we examine the total length of impaired waterways within a subwatershed. Here, we find a departure from the prior results. This measurement of impaired waterways appears very evenly distributed across these demographic variables. One possibility is that richer and more educated areas are able to direct attention and funding towards listing polluted areas as a first step towards remediation. However, we advise caution in relying too heavily on these impairment results as they reflect state processes that do not necessarily capture the ambient status of all waterways.[17]

When we examine these other measures of the presence of pollution within the top 40% of census block groups in terms of rural population share, some of the patterns in terms of race, poverty, and education change (Appendix Figure 8). The distribution of monthly average flow in terms of race, poverty, and education is more even within more rural census block groups compared to all census block groups. The distribution of outfalls from major facilities is more even in terms of race, poverty, and education though the Gini coefficients for race are now positive but small. Departing from previous results, the distribution of outfalls from facilities on the TRI, is more concentrated in areas with a larger non-White population share. This is consistent with our findings for outfalls from Chemicals and Allied Products. With respect to poverty and education, outfalls from facilities on the TRI within the most rural census block groups are more evenly distributed compared to all census block groups. Lastly, even within rural census block groups, impaired waterways are evenly distributed. Again, we interpret the impairment results with caution.

[16] We follow the procedures in one EPA document to convert demographic information from block groups to each HUC12
(https://www.epa.gov/system/files/documents/2022-03/demographics-indicator-reference-sheet-20220306.pdf).
Since we do not have information on which year’s impaired status was used in the EnviroAtlas database, we only construct the pseudo-Lorenz curve with 2019 demographic information.

[17] While EPA guidelines contain a list of information that must be considered, the specific framework for assessing water quality widely differs across states (National Research Council 2001). Thus, a given waterbody may be deemed impaired by one state but not impaired by another due to differences in their assessment framework.

We also investigated the distribution of Wastewater Discharge as collected by USEPA and compiled in the Climate and Economic Justice Screen Tool database. One would expect that the number of outfalls would roughly correspond to the amount of wastewater discharge within the same census block group. Appendix Figure 9 plots a pseudo-Lorenz curve for the number of outfalls for all facilities ranked by the percentile of wastewater discharge. Here, we see that wastewater discharge corresponds nearly one to one with the number of outfalls.[18]

4.2 Regression Results

We complement the analysis in Section 4.1 with regression results that examine how facility outfalls are correlated with socioeconomic characteristics. We examine the association between outfalls and the share of the population below the poverty line as well as the share of the population without a college degree or higher. We split out our variable that captures one measure of race (share non-White) to examine the correlations between the number of outfalls and more specific measures of race and ethnicity such as share Black and share Hispanic. We also add the rural population share to further examine how water pollution outfalls are distributed across the country.

Figure 9 summarizes our results across each of these six variables from regressions using all four time periods. These plots display coefficient estimates (blue dots) and 95% confidence intervals for each industrial classification for each particular variable of interest. No other controls are included in this specification, which allows us to examine the cross-sectional variation in outfalls across the entire U.S. Consistent with the pseudo-Lorenz curves, we find that outfalls are concentrated in areas of the country with a higher share of the population that is White, below the poverty line, and without a college education. We also find that outfalls are less likely to be in areas with a higher share of the population that is Black or Hispanic. When we examine the variable for rural population share, we find a strong association with the presence of outfalls across all industries. The results are similar when we include state fixed effects (Appendix Figure 10), which suggests these associations hold within states as well as across states.

As with the pseudo-Lorenz curves, we examine correlations with other potential indicators of the quantity and toxicity from these outfalls. Appendix Figures 11 and 12 show results for outfalls at major facilities and the monthly average flow at outfalls, respectively. The results for outfalls at major facilities are qualitatively similar as those using the number of outfalls, though there is arguably less variation across demographics. The magnitude of the differences is also much smaller. The results for monthly average flow are also qualitatively similar to the total number of outfalls, though flow is very strongly correlated with rural population share and lower education levels.

[18] Note, it may be the case that discharge data is not always available or completely reported to the federal government.

5. Discussion

Our analysis contributes to an understanding of the EJ implications of federal regulatory policy. Decades of individual studies and meta-analyses have demonstrated statistically significant relationships between race and many types of environmental hazards (Ringquist 2005, Mohai and Saha 2007). However, the majority of these studies focus on air pollution and hazardous waste disposal, with relatively few studies focusing on the distribution of point-source water pollution.

We aimed to address this gap through a two-pronged approach. First, we conducted a qualitative assessment of the implementation of EO 12898 in Clean Water Act rules and associated economic analyses by EPA. Second, we used data on the locations of permitted point-source facilities to evaluate the distribution of outfalls across different demographic variables of interest. Our goal was to describe the content and quality of current EJ analyses in clean water rules, evaluate alternative approaches to assess disparities in the location of water pollution sources, and inform future analyses of EJ in proposed rules or regulations designed to protect or restore water quality.

In the review of existing water quality rules, we come to similar conclusions as previous assessments of the implementation of EO 12898 in federal regulatory review. Geltman and Jovanovic (2016) and Banzhaf (2011) have strongly suggested the need for more rigorous analysis. After reviewing all rules since 1992, we found that EPA never determined a clean water regulation to have disproportionately high and adverse impacts on low-income or minority communities. While these findings could be correct, a lack of quantitative analysis within these reviews may leave one skeptical about the lack of EJ concerns. Further, the fact that a rule may not lead to adverse *changes* on low-income or minority communities does not necessarily imply that *current* EJ concerns are not important. In this regard, we also observed that EPA did not include publicly available data on baseline pollution exposure in their analyses. For example, in the rule evaluating the 2003 Effluent Limitation Guidelines and Standards for Concentrated Animal Feeding Operations, EPA determined the rule would have no disproportionate effect on minority or low-income communities. In this case, EPA could have cited research on the distribution of CAFOs, which have been shown to be disproportionately located in minority communities and low-income communities (Wilson et al. 2004, Son et al. 2021). However, no further analysis or research was conducted based on the justification that the rule would "benefit all of society."[19]

In our assessment of the baseline distribution of water polluting facilities in the U.S., we found that water pollution sources are more likely to be located in areas with a larger share of the

[19] Guidance on best practices for conducting distributional analyses of regulations is beyond the scope of this paper. However, we point readers to Ando et al. (2023), Lienke et al. (2021), and Banzhaf et al. (2019) as resources for best practices in assessing equity and distributional impacts of federal policies.

population that is White, below the poverty line, and without a college education. Overall, this pattern holds for most industries though the extent of the disproportionate siting is heterogeneous. In later time periods, we observe more even distributions across these characteristics and across most industries. We also found that block groups with the greatest number of outfalls are more likely to be rural. Within the most rural census block groups, however, the demographic and socioeconomic patterns are a bit more mixed. For the share of the population that is non-White, depending on industry and toxicity, the distribution remains the same, becomes more evenly distributed, or even slightly concentrated in areas with higher non-White population shares. Of note, facilities discharging more toxic pollutants tend to be located in rural areas with higher non-White population shares. In general, the concentration of outfalls in areas with higher poverty and lower education is slightly lower within rural areas, regardless of industry or toxicity of discharges. Overall, our results suggest that rurality partly drives the observed patterns for the other demographic and socioeconomic variables. However, we reiterate that our analysis does not address *causal* reasons for the location of these outfalls, but rather documents how they vary across space and socioeconomic characteristics.

We present a few takeaways from our assessment of the baseline distribution of polluting facilities that may be helpful for future EJ analyses. We found disparities across educational attainment, with greater numbers of pollution outfalls in census block groups with lower levels of education. EO 12898 only requires analysts to assess impacts on "minority and low-income populations in the US." Our analysis suggests that education may be an important factor to consider in future distributional assessments, especially as education attainment may be related to awareness of environmental hazards (Meyer 2015). Additionally, the heterogeneity we observed across industries highlights the importance of performing industry-specific EJ analyses; our findings revealed greater EJ concerns for certain industries such as mining, public wastewater treatment, and manufacturing.

Relatedly, not all outfalls pose equal risks to adjacent populations and examining outfalls with more harmful discharges in combination with those with relatively benign discharges may mask EJ concerns. Our analysis focusing on facilities with more toxic discharges is an imperfect exercise as it uses a facility's industrial codes and appearance on the TRI to capture the toxicity of discharges; a more refined approach that uses the type, quantity, and concentration of the discharged pollutants could provide a more accurate picture of the distribution of water pollution (see Liévanos et al. 2017 for an example of an approach that accounts for toxicity). To this end, recent agency investments such as the U.S. EPA's Risk-Screening Environmental Indicators (RSEI) database and the P2 EJ Facility Mapping Tool could increase the ease of future EJ analyses and the identification of toxic discharges.[20] Future work using more finely defined industrial categories focusing on specific pollutants may also improve our understanding of the distribution of water polluting facilities. Lastly, improved data collection on polluting facilities would greatly facilitate EJ analyses. The data sets we use are missing industrial codes for a substantial fraction of facilities which could alter our industry-specific conclusions.

Limitations and Caveats

We acknowledge several limitations to our analysis that may affect the interpretation of our findings. A challenge in conducting distributional assessments is enumerating the affected population and associated baseline group. In our analysis we assumed the impacted population was all households within a census block group containing a permitted facility. Some census block groups with high concentrations of polluting facilities may not have demographic data associated with them because they are in unpopulated industrial areas. These census block groups may be located adjacent to populated block groups, but in our analysis would not be identified as impacted by the number of facilities or outfalls in adjacent spatial units. An alternative would be to assess the affected population based on a proximity analysis (e.g. Mohai et al. 2009 assessed demographic characteristics within 1 mile of polluting facilities). A further limitation is that pollution outfalls do not necessarily represent pollution exposure. The movement of water pollutants through surface water and groundwater is complex and requires more data-intensive hydrologic modeling in order to link outfalls with concentrations, transport of pollutants downstream, and exposure of communities to pollution via direct or indirect consumption or water contact recreation.

We also acknowledge that in some cases the siting of water treatment facilities may indicate environmental improvement, not degradation, if the alternative was release of untreated pollution into adjacent waterways. Lastly, the context in which permits are issued matters for the overall interpretation of the distribution of water pollution sources. For example, a permit issued for construction activities may signify local development whereas a permit issued for chemical manufacturing may suggest more harmful polluting activities.

Inconsistent methodologies across distributional analyses also make it difficult to compare our results with previous studies. We know that selection of affected communities, reference population, spatial unit of analysis, assumptions about exposure and health impacts, and the ability to control for other contributing factors have been found to impact the conclusions of past EJ studies (Anderton et al. 1996, Mohai et al. 2009). As others have noted (Keeler et al. 2012,

[20] The U.S. EPA's Risk-Screening Environmental Indicators (RSEI) database links potential chemical releases from facilities on the Toxic Release Inventory to surface water "flowlines" up to 300 kilometers downstream from a facility. The RSEI method also attempts to link pollution to exposure via pathways of drinking water and recreational and subsidence fish consumption (US EPA 2023). The P2 EJ Facility Mapping Tool that allows users to identify industrial facilities located in or adjacent to communities with EJ concerns, including facilities included in the Toxic Release Inventory (TRI) and Resource Conservation and Recovery Act (RCRA).

Keiser, Kling, and Shapiro 2019) water quality-related benefits remain difficult to quantify and monetize, making it challenging to assess net benefits of proposed rules or regulations. Welfare impacts of exposure to water pollution can be moderated or exacerbated by infrastructure, adoption of avoidance behaviors, preexisting health conditions, and baseline exposure to other contaminants among other factors.

Our analysis offers limited insight into how changes in water pollution will affect other types of valued benefits, including recreation, cultural resources, non-use benefits or other aspects of human wellbeing. Water quality-related benefits, including the destruction of culturally-valued species, loss of access to ceremonial springs, mercury contamination of fish, and polluted beaches and swimming places can have a particular significance for EJ communities (EJCW 2005), but were beyond the scope of this analysis. We also focus solely on water pollution sources. This may miss important cumulative impacts of exposure to other types of pollution and other stressors that communities face. Indeed, EPA notes that cumulative impacts are an important area for future focus (EPA 2022b).

Future Directions

Our retrospective analysis of the implementation of EO 12898 suggests that more work is needed to come to a shared definition of what constitutes a “disproportionately high and adverse human health or environmental effect” as it relates to water quality. Consistency across methodologies and their application will allow for more systematic assessment of impacts of proposed water quality rules or regulations. Agency analysis aside, there is also no agreed-upon methodology in the academic literature to assess disproportionality of environmental benefits and burdens. The draft updated Circular A-4 guidance from OMB calls for increased investment in distributional analysis but stops short of prescribing a standardized approach. Access to publicly available data on the location of polluting facilities and outfalls opens opportunities for analysis that could greatly improve on past assessments of water quality rules and regulations. Here, we demonstrate how analysts can use these data, along with sociodemographic information, and other environmental variables to assess potential distributional effects of changes that affect permitted facilities. Future analyses could assess the sensitivity of assumptions about the appropriate spatial unit to determine affected population, focus on specific industries or regulated contaminants, assess compliance records of regulated facilities, or link outfall data with toxicity information to understand relative levels of harm from water pollution. Any or all of these approaches would represent progress relative to past implementation of environmental justice analyses in clean water rules.

References

Anderton, D. L. 1996. “Methodological issues in the spatiotemporal analysis of environmental equity.” *Social Science Quarterly,* 77(3), 508-515. http://www.jstor.org/stable/42863497

Ando. A., Awokuse, T.O., Chan, N.W., González-Ramírez, J., Gulati, S., Interis, M.G., Jacobson, S., Manning, D.T. and Stolper, S. 2023. “Environmental and Natural Resource Economics and Systemic Racism.” (Working Paper 23-06). *Resources for the Future.* https://media.rff.org/documents/WP_23-06.pdf

Ash, M., Boyce, J. K. 2018. “Racial disparities in pollution exposure and employment at US industrial facilities.” *Proceedings of the National Academy of Sciences of the United States of America* 115(42), 10636-10641. https://doi.org/10.1073/pnas.1721640115

Balazs C., Ray, I. 2014. “The Drinking Water Disparities Framework: On the Origins and Persistence of Inequities in Exposure.” *American Journal of Public Health* 104(4), 603-611. https://doi.org/10.2105/AJPH.2013.301664

Banzhaf, H. 2011. “Regulatory Impact Analyses of Environmental Justice Effects.” *Journal of Land Use & Environmental Law 27*(1), 1-30.

Banzhaf, H., Ma, L., Timmins, C. 2019. “Environmental Justice: Establishing Causal Relationships.” *Annual Review of Resource Economics* 11(1), 377-398.

Benz, S. A., Burney, J. A. 2021. “Widespread Race and Class Disparities in Surface Urban Heat Extremes across the United States.” *Earth's Future*, 9(7), e2021EF002016. https://doi.org/10.1029/2021EF002016

Buckhoy, N. 2015. “Environmental Justice for Whom?: A Social Construction Framework Analysis of Executive Order 12898.” *Environmental Justice* 8(5), 157-164. http://doi.org/10.1089/env.2015.0014

Bullard, R. D., Mohai, P., Saha, R., & Wright, B. 2008. “Toxic wastes and race at twenty: Why race still matters after all of these years.” *Environmental Law* 38(1), 371.

Cecot, C., Hahn, R. W. 2022. “Incorporating Equity and Justice Concerns in Regulation.” *George Mason Law & Economics Research Paper,* 22-19. https://doi.org/10.1111/rego.12508

Colmer, J., Hardman, I., Shimshack, J., Voorhies, J., 2020. “Disparities in PM2.5 air pollution in the United States.” *Science* 369(6503), 575-578. https://doi.org/10.1126/science.aaz9353

Dalzell. 2012. “Linking water quality and human well-being for improved assessment and valuation of ecosystem services.” *The Proceedings of the National Academy of Sciences.* 109(45), 18619-18624. https://doi.org/10.1073/pnas.1215991109

Deitz, S., Meehan, K. 2019. “Plumbing Poverty: Mapping Hot Spots of Racial and Geographic Inequality in U.S. Household Water Insecurity.” *Annals of the American Association of Geographers,* 109(4), 1092-1109. https://doi.org/10.1080/24694452.2018.1530587

Environmental Justice Coalition for Water (EJCW). 2005. “Thirsty for Justice: A People’s Blueprint for California Water.” *Oakland, CA: Environmental Justice Coalition for Water.*

Exec. Order No. 12,898, 32 CFR § 651.17. 1994. Federal Actions To Address Environmental Justice in Minority Populations and Low-Income Populations. https://www.epa.gov/sites/default/files/2015-02/documents/exec_order_12898.pdf

Geltman, E.G., Gill, G., Jovanovic, M., 2016. “Beyond Baby Steps: An Empirical Study of the Impact of Environmental Justice Executive Order 12898.” *Fam Community Health* 39(3),143-50. https://doi.org/10.1097/fch.0000000000000113

Keeler, B.L., S. Polasky, K. Brauman, K. Johnson, J. Finlay, A. O’Neill, K. Kovacs, B. 2012. “Linking water quality and well-being for improved assessment and valuation of ecosystem services.” *Environmental Science* 109(45), 18619-18624. https://doi.org/10.1073/pnas.1215991109

Keiser, D. A., Kling, C. L., Shapiro, J. S. 2019. “The low but uncertain measured benefits of US water quality policy.” *The Proceedings of the National Academy of Sciences*, 116(12), 5262-5269. https://doi.org/10.1073/pnas.1802870115

Keiser, D.A. and J.S. Shapiro. 2019. “U.S. Water Pollution Regulation Over the Last Half Century: Burning Waters to Crystal Springs?” *Journal of Economic Perspectives*, 33(4): 51-75. https://doi.org/10.1257/jep.33.4.51

Konisky, D.M., Reenock, C. and Conley, S., 2021. “Environmental Injustice in Clean Water Act Enforcement: Racial and Income Disparities in Inspection Time.” *Environmental Research Letters*, *16*(8), 084020. https://doi.org/10.1088/1748-9326/ac1225

Lee, C. 2002. “Environmental Justice: Building a Unified Vision of Health and the Environment.” *Environmental Health Perspectives*, 110(suppl 2), 141-144. https://doi.org/10.1289/ehp.02110s2141

Lienke, J., Paul, I., Sarinsky, M., Ünel, B., Varela, A. V. 2021. “Making Regulations Fair: How Cost-Benefit Analysis Can Promote Equity and Advance Environmental Justice.” *Institute for Policy Integrity.*

Liévanos, R. S. 2017. “Sociospatial dimensions of water injustice: The distribution of surface water toxic releases in California’s Bay-Delta.” *Sociological Perspectives 60*(3), 575-599. https://doi.org/10.1177/0731121416648935

Mack EA, Wrase S. 2017. “Correction: A Burgeoning Crisis? A Nationwide Assessment of the Geography of Water Affordability in the United States.” *PLOS ONE* 12(4), e0176645. https://doi.org/10.1371/journal.pone.0176645

Manson, S., Schroeder, J., Riper, D.W., Kugler, T., Ruggles, S. 2022. “IPUMS National Historical Geographic Information System: Version 17.0 Minneapolis, Mn.” *Integrated Public Use Microdata Series.* http://doi.org/10.18128/D050.V17.0

McDonald, Y. J., Jones, N. E. 2018. “Drinking Water Violations and Environmental Justice in the United States, 2011–2015.” *American Journal of Public Health* 108(10), 1401–1407. https://doi.org/10.2105/AJPH.2018.304621

Mehta V. K., Goswami, R., Kemp-Benedict E., Muddu, S. Malghan, D. 2014. “Metabolic Urbanism and Environmental Justice: The Water Conundrum in Bangalore, India.” *Environmental Justice* (7)5, 130-137. http://doi.org/10.1089/env.2014.0021

Mennis, J. L. 2003. “The Distribution and Enforcement of Air Polluting Facilities in New Jersey.” *The Professional Geographer* 57(3), 411-422. https://doi.org/10.1111/j.0033-0124.2005.00487.x

Meyer, A. (2015). “Does education increase pro-environmental behavior? Evidence from Europe.” *Ecological Economics* 116, 108-121. https://doi.org/10.1016/j.ecolecon.2015.04.018

Miller, D. 2005. “Methods for Evaluating Environmental Justice — Approaches to Implementing U.S. Executive Order 12898.” *Beyond Benefit Cost Analysis.* https://doi.org/10.4324/9781351162685

Mohai, P., Saha, R. 2007.. “Racial inequality in the distribution of hazardous waste: A national-level reassessment.” *Social Problems* 54(3), 343-370. https://doi.org/10.1525/sp.2007.54.3.343

Mohai, P., Lantz, P. M., Morenoff, J., House, J. S., Mero, R. P. 2009. “Racial and Socioeconomic Disparities in Residential Proximity to Polluting Industrial Facilities: Evidence From the Americans' Changing Lives Study.” *American Journal of Public Health* 99(3), S649-S656. https://doi.org/10.2105/ajph.2007.131383

Mohammed, M., Dukku, S. J., Babanyara, Y. Y., Muhammad, I., Mohammed, J. K. 2021. “A Comparative Analysis of Environmental Justice Between Urban Neighbourhoods.” *Journal of*

Inclusive Cities and Built Environment 1(2), 27-40. https://doi.org/10.54030/2788-564X/2021/v1i2a4.

Mueller, J.T., Gasteyer, S. 2021. "The widespread and unjust drinking water and clean water crisis in the United States." *Nature Communications* 12(1), 3544. https://doi.org/10.1038/s41467-021-23898-z

National Environmental Justice Advisory Council (NEJAC). 2002. "EJ 2020 Action Agenda." *Environmental Protection Agency.* Online at: https://www.epa.gov/sites/default/files/2017-10/documents/ej_2020_action_agenda.pdf

National Research Council, 2001. *Assessing the TMDL approach to water quality management*. National Academies Press.

O'Neil, SG. 2007. "Superfund: Evaluating the Impact of Executive Order 12898." *Environmental Health Perspectives.* 115(7), 1087-1093. https://doi.org/10.1289/ehp.9903.

Revesz, R. L., Yi, S. 2022. "Distributional Consequences and Regulatory Analysis." *Environmental Law 52(1)*, 53-98. https://www.jstor.org/stable/48657962

Ringquist, E. J. 2005. "Assessing Evidence of Environmental Inequities: A Meta‐analysis." *Journal of Policy Analysis and Management* 24(2), 223-247. https://doi.org/10.1002/pam.20088

Robinson, L. A., Hammitt, J. K., Zeckhauser R. J. 2016. "Attention to Distribution in U.S. Regulatory Analyses." *Review of Environmental Economics and Policy* 10(2), 308-328.

Sheriff, G., Maguire, K. 2013. "Ranking Distributions of Environmental Outcomes Across Population Groups." (Working Paper 13-04) *National Center for Environmental Economics*. http://dx.doi.org/10.22004/ag.econ.280914

Sheriff, G. 2023. "California's GHG cap and trade program and the equity of air toxic releases." *Journal of the Association of Environmental and Resource Economists*. Forthcoming.

Son, J.Y., Muenich, R.L., Schaffer-Smith, D., Miranda, M.L., Bell, M.L. 2021. "Distribution of Environmental Justice Metrics for Exposure to CAFOs in North Carolina, USA." *Environmental Research* 195, 110862. https://doi.org/10.1016/j.envres.2021.110862

Tate, E., Rahman, M. A., Emrich, C. T., Sampson, C. C. 2021. Flood exposure and social vulnerability in the United States. *Natural Hazards* 106(1), 1-23. https://doi.org/10.1007/s11069-020-04470-2

Tessum et al., 2021. "$PM_{2.5}$ Polluters Disproportionately and Systemically Affect People of Color in the United States." *Science Policy* 7(18). https://doi.org/10.1126/sciadv.abf4491

USEPA (US Environmental Protection Agency). 2004. "Effluent Limitations Guidelines and New Source Performance Standards for the Meat and Poultry Products Point Source Category." Online at: https://www.federalregister.gov/documents/2004/09/08/04-12017/effluent-limitations-guidelines-and-new-source-performance-standards-for-the-meat-and-poultry

— 2004. "EPA Needs to Consistently Implement the Intent of the Executive Order on Environmental Justice." *Report No. 2004-P-00007*. Online at: https://www.epa.gov/office-inspector-general/report-epa-needs-consistently-implement-intent-executive-order

— 2009. "Oil Pollution Prevention; Spill Prevention, Control, and Countermeasure (SPCC) Rule—Amendments." Online at: https://www.federalregister.gov/documents/2009/11/13/E9-27156/oil-pollution-prevention-spill-prevention-control-and-countermeasure-spcc-rule-amendments

— 2010. "Water Quality Standards; Establishment of Numeric Criteria for Priority Toxic Pollutants; States' Compliance." Online at: https://www.federalregister.gov/citation/57-FR-60848

— 2011. "Oil Pollution Prevention; Spill Prevention, Control, and Countermeasure (SPCC) Rule-Amendments for Milk and Milk Product Containers." Online at: https://www.federalregister.gov/documents/2011/04/18/2011-9288/oil-pollution-prevention-spill-prevention-control-and-countermeasure-spcc-rule-amendments-for-milk

— 2015. "Effluent Limitations Guidelines and Standards for the Steam Electric Power Generating Point Source Category." Online at: https://www.federalregister.gov/documents/2019/11/22/2019-24686/effluent-limitations-guidelines-and-standards-for-the-steam-electric-power-generating-point-source

— 2022a. "Justice40 at EPA." Online at: https://www.epa.gov/environmentaljustice/justice40-epa

— 2022b. "Cumulative Impacts: Recommendations for ORD Research." Online at: https://www.epa.gov/system/files/documents/2022-01/ord-cumulative-impacts-white-paper_externalreviewdraft-_508-tagged_0.pdf

— 2023. "EPA's Risk-Screening Environmental Indicators (RSEI) Methodology: RSEI Version 2.3.11". Online at: https://www.epa.gov/system/files/documents/2023-04/rsei-methodology-document-v2311-March2023.pdf

Vajjhala, S. P., Szambelan, S., Van Epps, A. 2008. "Integrating EJ into Federal Policies and Programs: Examining the Role of Regulatory Impact Analyses and Environmental Impact Statements." Resources for the Future Discussion Paper No. 04-45. https://dx.doi.org/10.2139/ssrn.1299062

Wang, Y., Apte, J.S., Marshall, J. 2022. "Location-Specific Strategies for Eliminating US National Racial-Ethnic PM2.5 Exposure Inequality." *Environmental Sciences* 119(44), e2205548119. https://doi.org/10.1073/pnas.2205548119

White House. 2021. "Modernizing Regulatory Review." Online at: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/modernizing-regulatory-review/

White House. 2023. "Circular A-4 Draft for Public Review". Online at: https://www.whitehouse.gov/wp-content/uploads/2023/04/DraftCircularA-4.pdf

Wilson, S.M., Howell, F., Wing, S., Sobsey, M. 2002. "Environmental injustice and the Mississippi Hog Industry." *Environmental Health Perspectives* 110, 195-201. https://doi.org/10.1289/ehp.02110s2195

Figures and Tables

Table 1. Summary of Quantitative CWA Environmental Justice Assessments (1992-2019).

Rule	Demographic Characteristics Evaluated	Spatial Scale of Analysis	Benefits Assessed	Determination
National Emission Standards for Hazardous Air Pollutants for Source Category: Pulp and Paper Production; Effluent Limitations Guidelines, Pretreatment Standards, and New Source Performance Standards: Pulp, Paper, and Paperboard Category (1998)	Race, income, indigeneity	EPA analyzed subsistence anglers fishing in the vicinity of bleached kraft mills from the consumption of dioxin-contaminated fish. EPA also examined county-level race and income data to assess whether bleached kraft mills have a disproportionate effect on minority and low-income populations.	Subsistence fishing, cancer risks. price increases due to increased compliance costs	EPA expects the final rule to reduce substantially the cancer risks to tribal populations.
Effluent Limitations Guidelines and New Source Performance Standards for the Metal Products and Machinery Point Source Category (2003)	Race, national origin, income level, indigeneity	EPA assessed counties traversed by water receiving discharges from 32 sample MP&M facilities and compared them to state averages.	Subsistence fishing, cancer risks, systemic health risk	EPA expects that the rule will neither promote nor discourage environmental justice.
National Pollutant Discharge Elimination System—Final Regulations to Establish Requirements for Cooling Water Intake Structures at Phase II Existing Facilities (2004)	Race, income	EPA analyzed demographics of communities within 50 miles radius of affected facilities compared to national averages.	Subsistence fishing	All populations, including minority and low-income populations, would benefit from improved environmental conditions as a result of this rule.
National Pollutant Discharge Elimination System—Final Regulations To Establish Requirements for Cooling Water Intake Structures at Existing Facilities and Amend Requirements at Phase I Facilities (2014)	Race, income	EPA defined the benefit population as (1) all individuals who live within a 50-mile radius of the facilities and (2) any additional anglers who live outside of the 50-mile facility buffer but within a 50-mile radius of the river segments, or river reaches, nearest to the facilities. EPA compared this to the general state population.	Subsistence fishing	EPA expects that this final rule will help to preserve the health of aquatic ecosystems near regulated facilities, EPA expects that all populations, including minority and low-income populations, will benefit from improved environmental conditions.
Effluent Limitations Guidelines and Standards for the Steam Electric Power Generating Point Source Category (2015)	Race, income	EPA assessed affected communities within 50 miles of steam electric power plants and compared them to state averages.	Subsistence fishing, cancer risks, IQ decrements, systemic health risks	EPA's analysis finds very small changes in exposure between the baseline and regulatory options, amounting to very small changes in risk for this population.

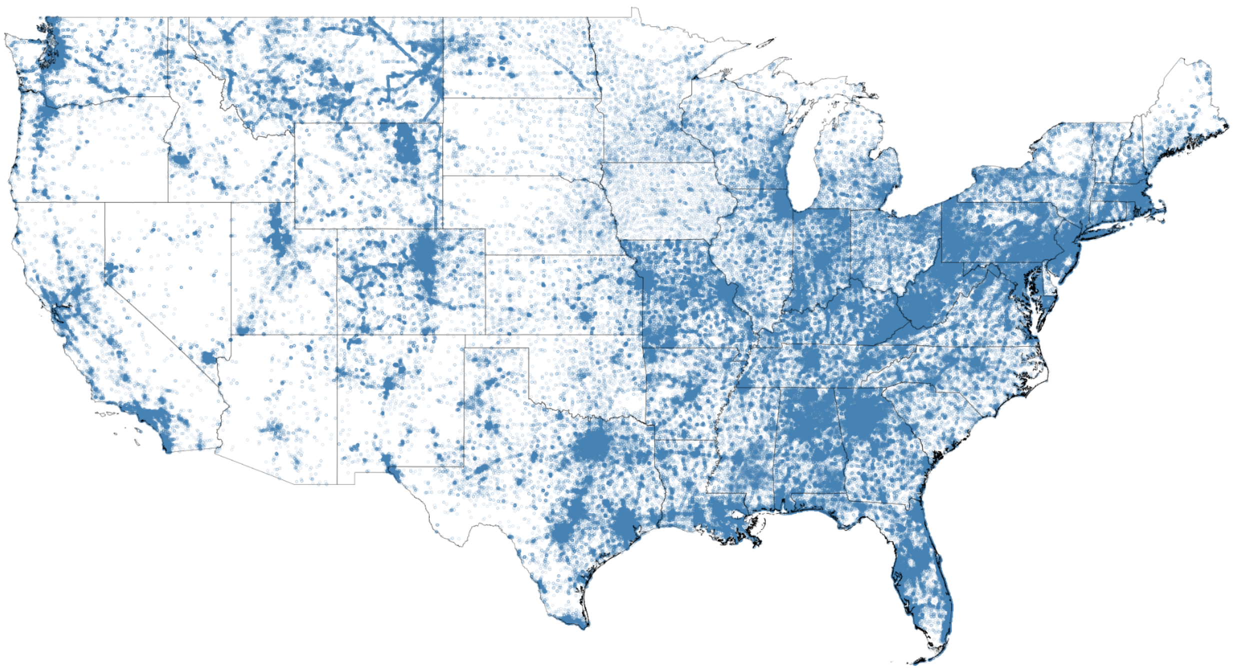
Notes: This table provides a summary of the five quantitative EJ analyses of the water pollution rules in our data set (1992-2019). It does not include the 13 water pollution rules which did not contain quantitative EJ analyses.

Table 2. Number of Outfalls by Industrial Classifications and Corresponding Demographics (2019-2022).

Industry	No. of Outfalls	No. of CBGs	Non-White	Non-College	Poverty	Rural Population
Agriculture, Forestry, Fishing	9,403	3,963	11%	78%	12%	83%
Construction	91,375	20,490	20%	71%	14%	34%
Electric, Gas, Sanitary Services	45,640	17,115	17%	75%	14%	57%
Finance, Insurance, Real Estate	12,799	5,304	17%	72%	12%	49%
Manufacturing	90,311	26,377	22%	77%	15%	43%
Mining	64,841	9,747	15%	79%	14%	73%
Public Administration	12,949	6,059	19%	72%	14%	38%
Retail Trade	3,891	2,602	22%	76%	15%	44%
Wholesale Trade	20,902	8,906	24%	79%	17%	39%
Services	13,232	7,814	17%	72%	13%	54%
Transportation and Communications	31,012	11,155	25%	75%	16%	30%
POTWs	26,417	16,370	15%	77%	14%	63%
All Industries - with industrial code	422,772	70,487	20%	74%	14%	44%
All Industries	632,609	97,643	21%	72%	14%	37%
All CBGs (CONUS Average)	632,609	216,330	27%	70%	15%	23%

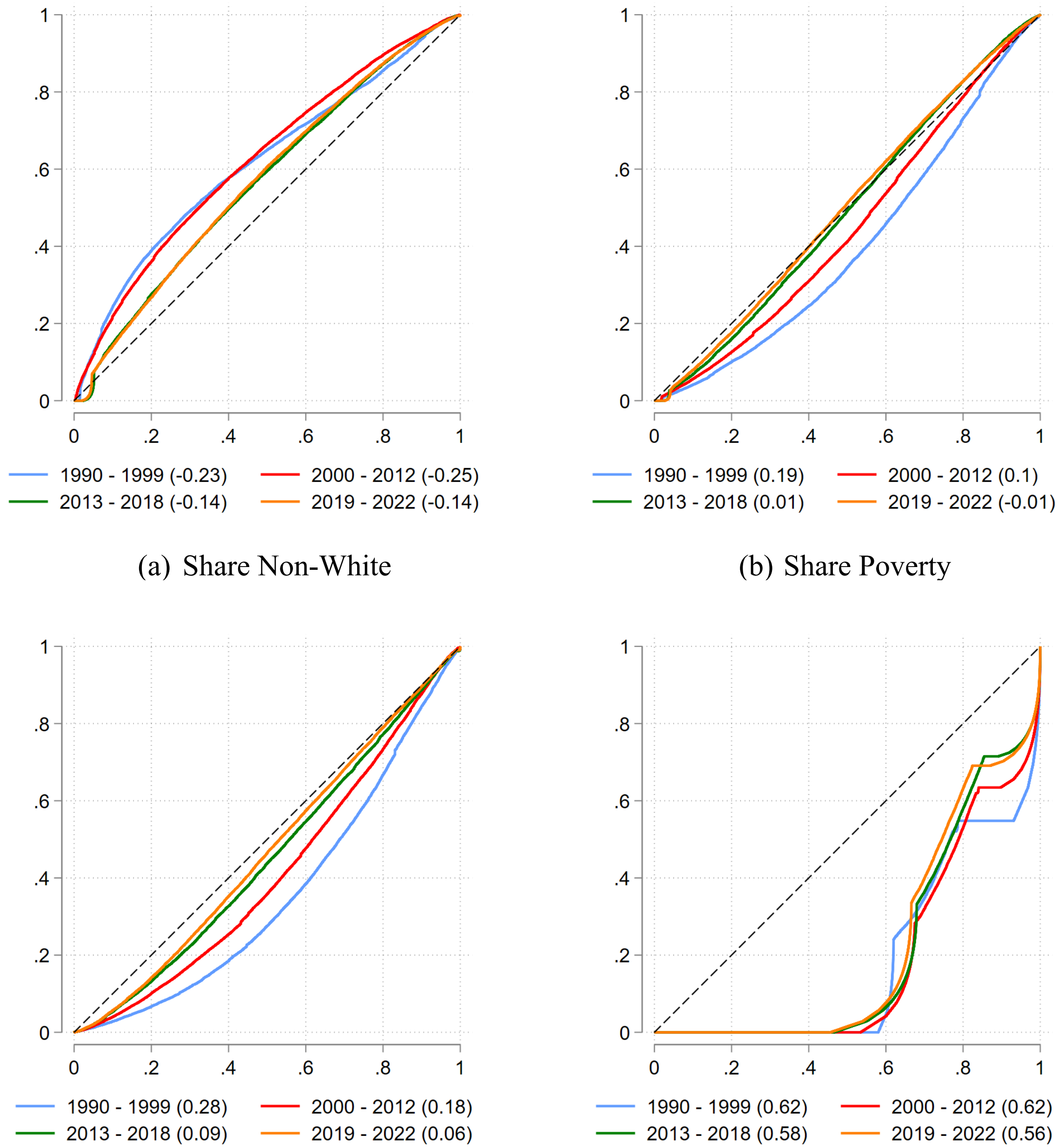
Notes: This table provides summary statistics of the total number of outfalls and the distribution of outfalls by industrial classification for the 2019-2022 period. The table also shows corresponding census block group demographic information for all census block groups and by industrial classification. "All Industries - with industrial code" summarizes these statistics for outfalls that correspond to facilities with at least one industrial code. "All Industries" summarizes these statistics for all outfalls, regardless of the availability of the industrial code. "All CBGs" summarizes these statistics for all census block groups in the conterminous US (CONUS) for comparison purposes.

Figure 1. Map of Outfalls from Active Facilities (1990 - 2022).



Notes: This map depicts outfalls from facilities active at any point during the 1990-2022 period.

Figure 2. Pseudo-Lorenz Curves for the Count of All Outfalls by Race, Poverty, Education, and Rurality.



(a) Share Non-White

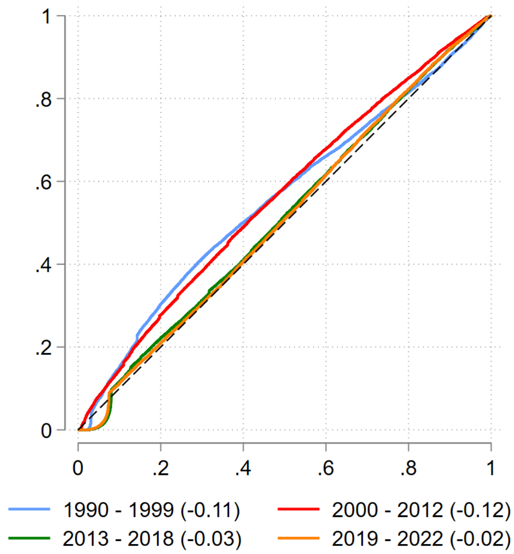
(b) Share Poverty

(c) Share Non-College

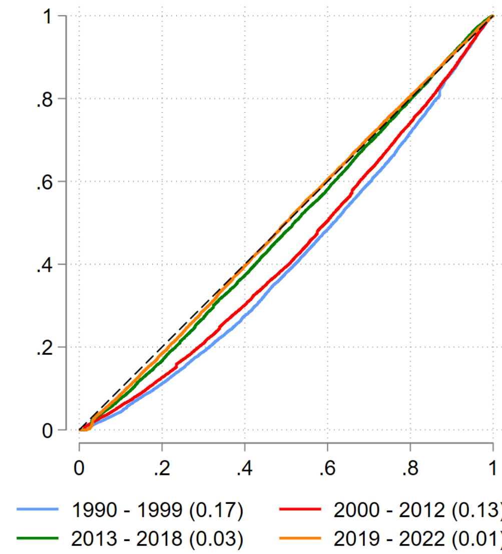
(d) Rural Population Share

Notes: This figure displays pseudo-Lorenz curves for the count of all outfalls across all industrial sectors by the share of the population non-White (Figure 2a), share of the population below the poverty line (Figure 2b), the share of the population without a college education or higher (Figure 2c), and the share of the population living in rural areas (Figure 2d). The 45-degree line represents equal distribution. Each figure shows how these counts change over time from the 1990-1999 period (blue line) to the 2019-2022 period (orange line). Gini coefficients for each period are in parentheses in the legend.

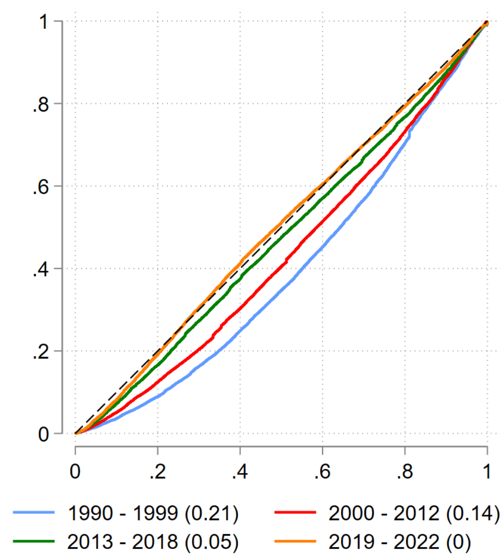
Figure 3. Pseudo-Lorenz Curves for the Count of All Outfalls by Race, Poverty, Education for the Top 40% of Census Block Groups in Terms of Rural Population Share.



(a) Share Non-White



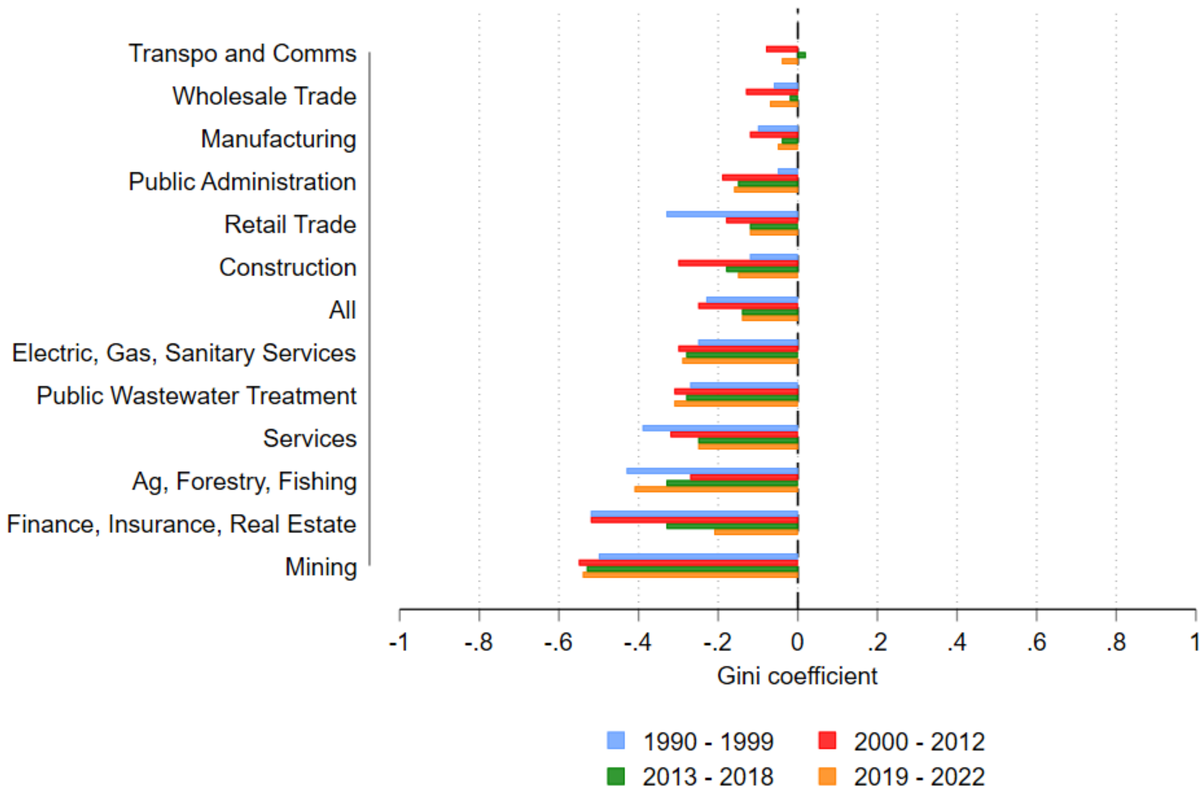
(b) Share Poverty



(c) Share Non-College

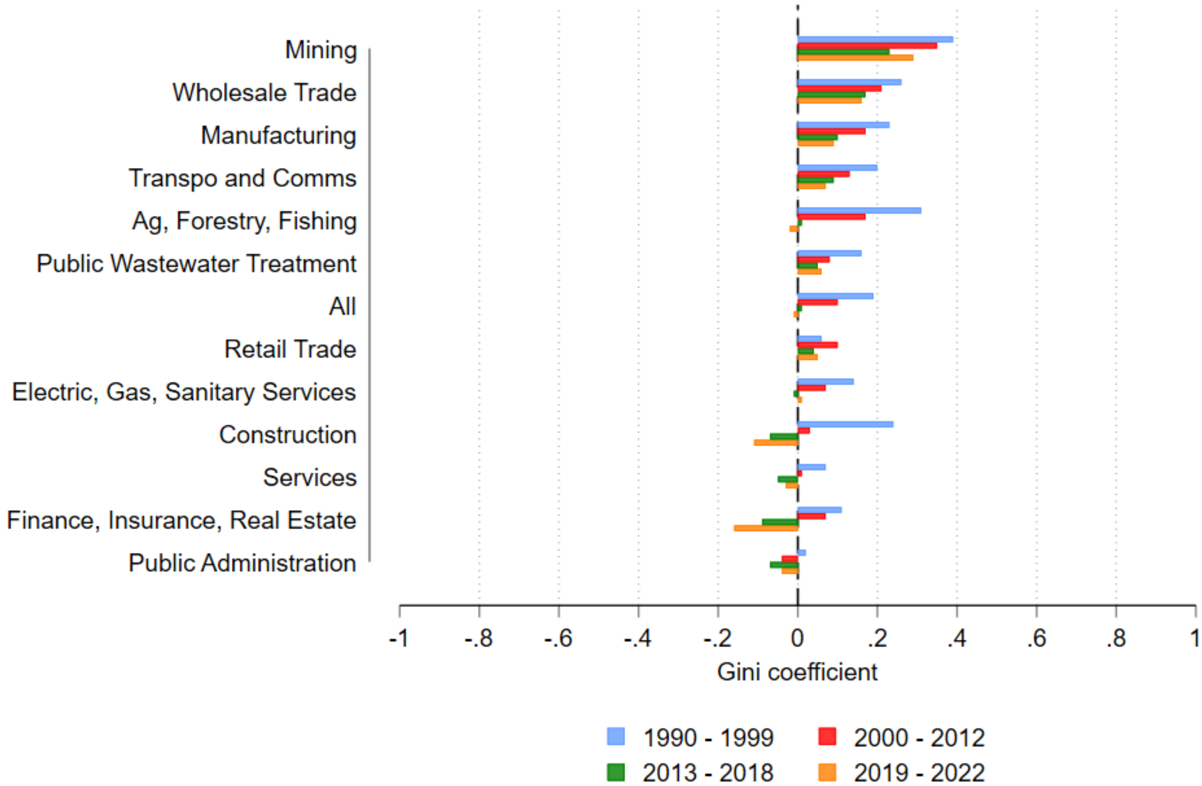
Notes: This figure displays pseudo-Lorenz curves for the count of all outfalls across all industrial sectors by the share of the population non-White (Figure 3a), share of the population below the poverty line (Figure 3b), and the share of the population without a college education or higher (Figure 3c). The 45-degree line represents equal distribution. Each figure shows how these counts change over time from the 1990-1999 period (blue line) to the 2019-2022 period (orange line). Gini coefficients for each period are in parentheses in the legend. We restrict the sample to the top 40% of census block groups in terms of rural population share.

Figure 4. Gini Coefficients for Distribution of Outfalls by Share Non-White and by Industry.



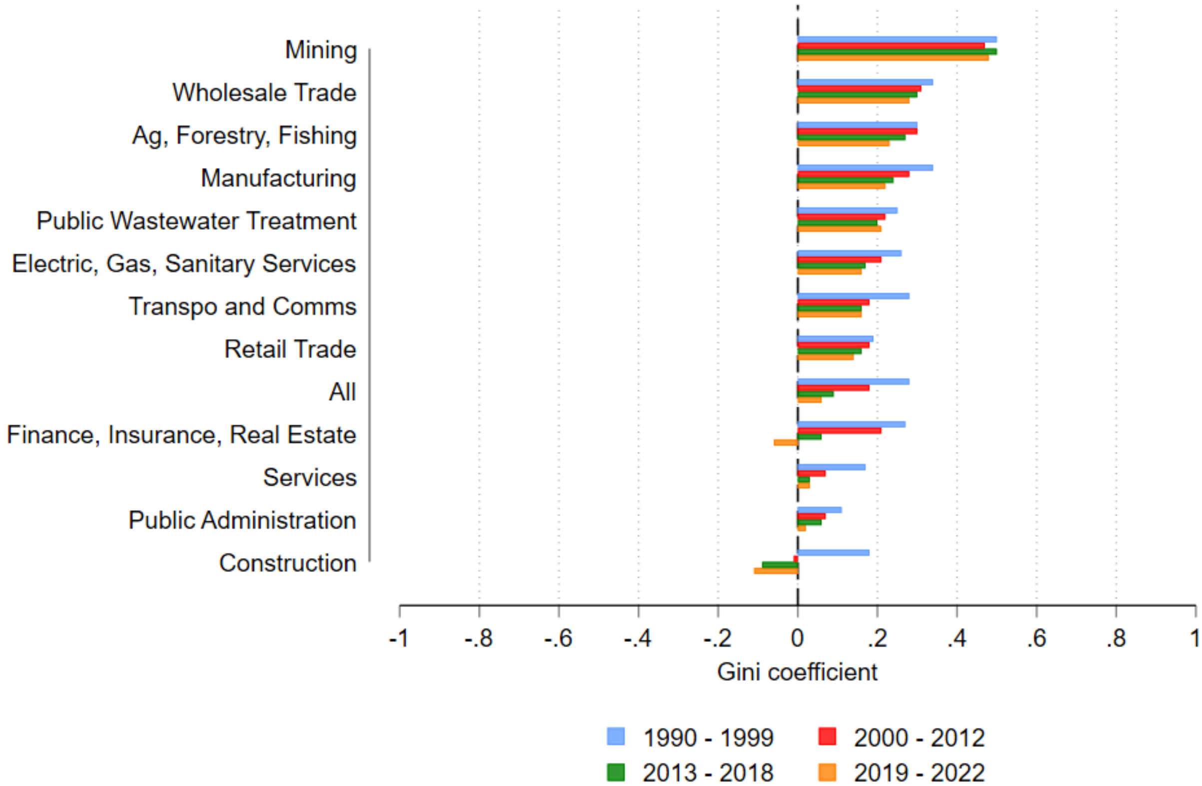
Notes: This graph presents Gini coefficients by industry for the share of the population that is non-White. The category "All" depicts Gini coefficients across all facilities, regardless of the availability of the industrial code. Blue bars are for the 1990 - 1999 period, red bars are for the 2000 - 2012 period, green bars are for the 2013 - 2018 period, and orange bars are for the 2019 - 2022 period.

Figure 5. Gini Coefficients for Distribution of Outfalls by Poverty and by Industry.



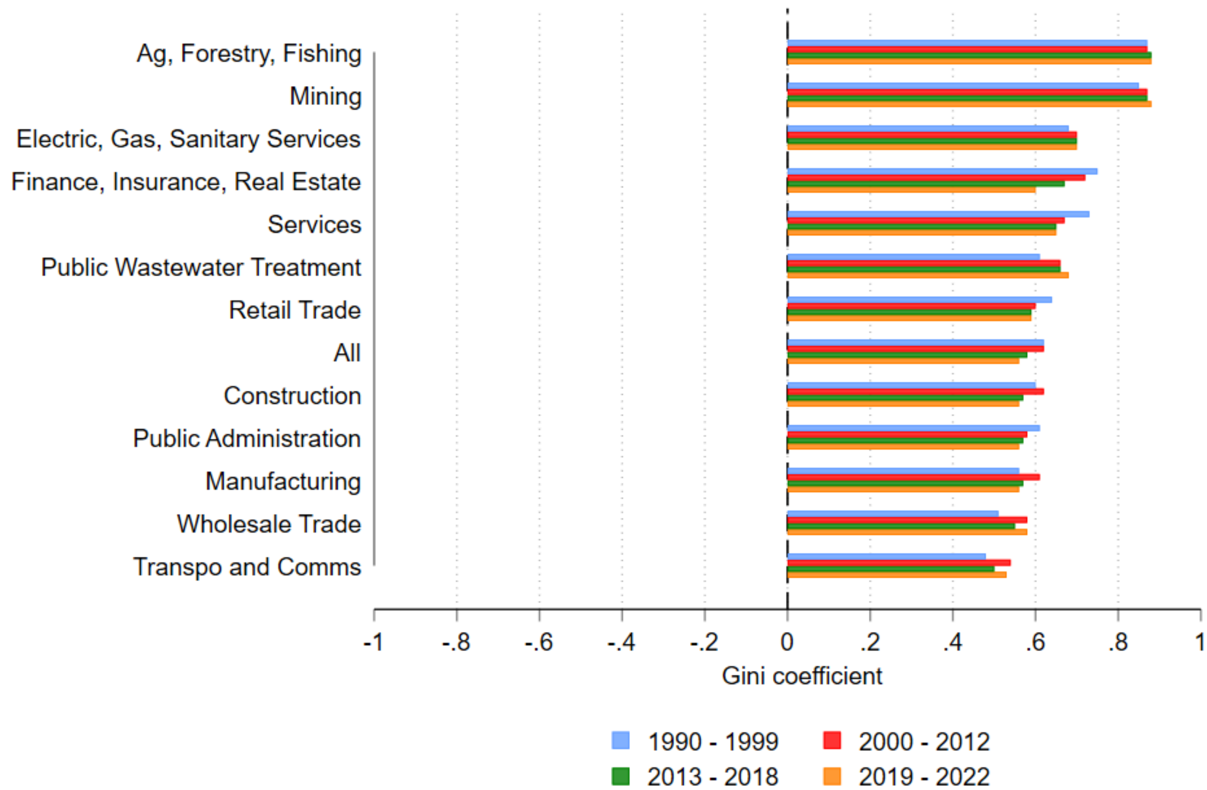
Notes: This graph presents Gini coefficients by industry for the share of the population below the poverty line. The category "All" depicts Gini coefficients across all facilities, regardless of the availability of the industrial code. Blue bars are for the 1990 - 1999 period, red bars are for the 2000 - 2012 period, green bars are for the 2013 - 2018 period, and orange bars are for the 2019 - 2022 period.

Figure 6. Gini Coefficients for Distribution of Outfalls by Education and by Industry.



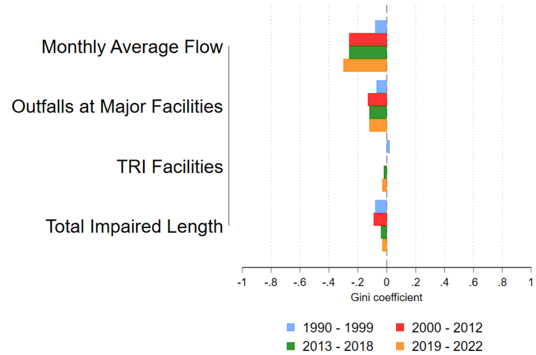
Notes: This graph presents Gini coefficients by industry for the share of the population without a college degree or higher. The category "All" depicts Gini coefficients across all facilities, regardless of the availability of the industrial code. Blue bars are for the 1990 - 1999 period, red bars are for the 2000 - 2012 period, green bars are for the 2013 - 2018 period, and orange bars are for the 2019 - 2022 period.

Figure 7. Gini Coefficients for Distribution of Outfalls by Rural Population Share and by Industry.

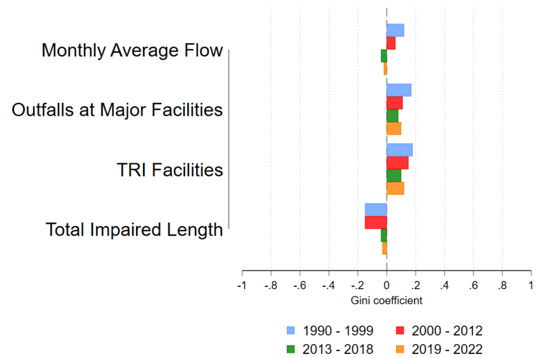


Notes: This graph presents Gini coefficients by industry for the rural population share. The category "All" depicts Gini coefficients across all facilities, regardless of the availability of the industrial code. Blue bars are for the 1990 - 1999 period, red bars are for the 2000 - 2012 period, green bars are for the 2013 - 2018 period, and orange bars are for the 2019 - 2022 period.

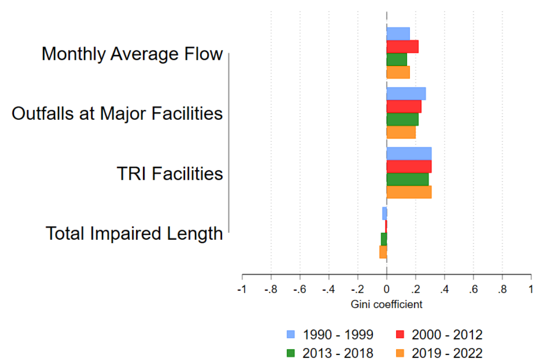
Figure 8. Gini Coefficients for Other Measures of Pollution.



(a) Share Non-White



(b) Share Poverty



(c) Share Non-College

Notes: This figure displays bar graphs depicting Gini coefficients using the following measures: (1) monthly average flow, (2) outfalls at facilities deemed "Major" by USEPA, (3) the number of facilities on the TRI with a water pollution discharge permit, and (4) the total impaired waterway length. For each measure of pollution, we examine the distribution across the share of the population non-White (Figure 8a), the share of the population below the poverty line (Figure 8b), and the share of the population without a college education or higher (Figure 8c).

Figure 9. Cross-Sectional Results Between Presence of Outfalls and Demographics.

Notes: This figure displays results from cross-sectional regressions of an indicator for the presence of outfalls in a census block group versus a given measure of demographics. The outcome variable for the category "All" is an indicator for the presence of an outfall in a census block group, regardless of industrial classification. This category includes facilities with and without SIC information. For the remaining categories, we use industry-specific indicators for the presence of an outfall in a census block group as the outcome variable. Coefficient estimates are shown in blue dots and 95% confidence intervals are shown by the corresponding lines. Standard errors are clustered at the county level. Results are grouped by demographic variable and by industrial classification. The sample includes all four time periods.

Appendix

Appendix Table 1. List of Regulatory Impact Assessments (RIAs) in Data Set.

Date	Title	ID #
1992	Water Quality Standards; Establishment of Numeric Criteria for Priority Toxic Pollutants; States' Compliance	40 CFR Part 131
1993	Oil and Gas Extraction Point Source Category; Offshore Subcategory Effluent Limitations Guidelines and New Source Performance Standards	RIN 2040-AA12
1995	Water Quality Standards for Surface Waters of the Sacramento River, San Joaquin River, and San Francisco Bay and Delta of the State of California	60 FR 4664
1998	National Emission Standards for Hazardous Air Pollutants for Source Category: Pulp and Paper Production; Effluent Limitations Guidelines, Pretreatment Standards, and New Source Performance Standards: Pulp, Paper, and Paperboard Category	RIN 2040–AB53
1999	National Pollutant Discharge Elimination System—Regulations for Revision of the Water Pollution Control Program Addressing Storm Water Discharges	RIN 2040–AC82
2003	Effluent Limitations Guidelines and New Source Performance Standards for the Metal Products and Machinery Point Source Category	RIN 2040-AB79
2003	National Pollutant Discharge Elimination System Permit Regulation and Effluent Limitation Guidelines and Standards for Concentrated Animal Feeding Operations (CAFOs)	RIN 2040–AD19
2004	Effluent Limitations Guidelines and New Source Performance Standards for the Meat and Poultry Products Point Source Category	RIN 2040–AD56
2004	National Pollutant Discharge Elimination System—Final Regulations to Establish Requirements for Cooling Water Intake Structures at Phase II Existing Facilities	RIN 2040–AD62
2006	Oil Pollution Prevention; Spill Prevention, Control, and Countermeasure Plan Requirements— Amendments	RIN 2050–AG23
2009	Construction and Development Effluent Guidelines	RIN 2040-AE91
2009	Oil Pollution Prevention; Spill Prevention, Control, and Countermeasure (SPCC) Rule— Amendments	RIN 2050–AG16
2010	Water Quality Standards for the State of Florida's Lakes and Flowing Waters	RIN 2040–AF11
2011	Oil Pollution Prevention; Spill Prevention, Control, and Countermeasure (SPCC) Rule— Amendments for Milk and Milk Product Containers	RIN 2050–AG50
2014	National Pollutant Discharge Elimination System—Final Regulations To Establish Requirements for Cooling Water Intake Structures at Existing Facilities and Amend Requirements at Phase I Facilities	RIN 2040–AE95
2015	Effluent Limitations Guidelines and Standards for the Steam Electric Power Generating Point Source Category	RIN 2040–AF77
2015	Clean Water Rule: Definition of 'Waters of the United States''	RIN 2040–AF30
2019	Definition of "Waters of the United States" - Recodification of Preexisting Rule	RIN 2040-AF74

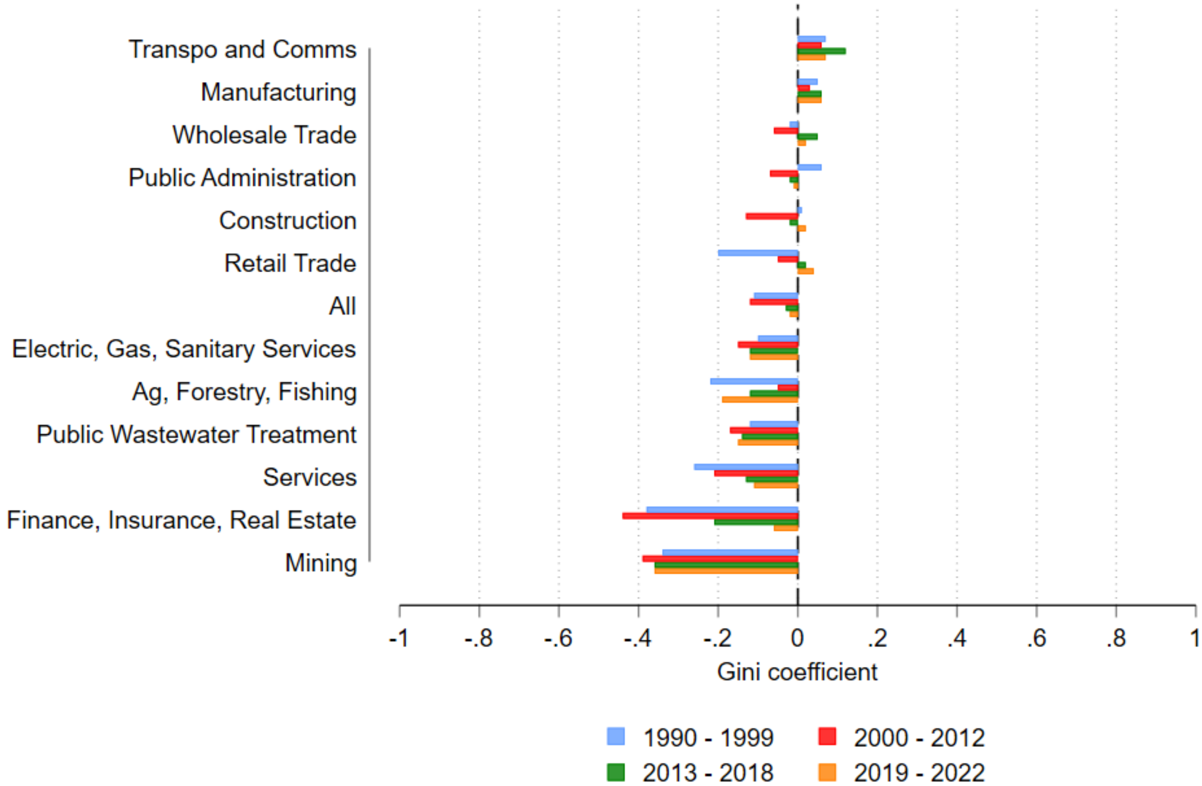
Appendix Table 2. Definitions of Industrial Divisions in Data Set.

Industrial Division	Major Group (Two Digit SIC Code)	
A: Agriculture, Forestry, and Fishing	01:	Agricultural Production Crops
	02:	Agriculture Production Livestock and Animal Specialties
	07:	Agricultural Services
	08:	Forestry
	09:	Fishing, Hunting, and Trapping
B: Mining	10:	Metal Mining
	12:	Coal Mining
	13:	Oil and Gas Extraction
	14:	Mining and Quarrying of Nonmetallic Minerals, Except Fuels
C: Construction	15:	Building Construction General Contractors and Operative Builders
	16:	Heavy Construction other than Building Construction Contractors
	17:	Construction Special Trade Contractors
D: Manufacturing	20:	Food and Kindred Products
	21:	Tobacco Products
	22:	Textile Mill Products
	23:	Apparel and other Finished Products Made from Fabrics and Similar Materials
	24:	Lumber and Wood Products, Except Furniture
	25:	Furniture and Fixtures
	26:	Paper and Allied Products
	27:	Printing, Publishing, and Allied Industries
	28:	Chemicals and Allied Products
	29:	Petroleum Refining and Related Industries
	30:	Rubber and Miscellaneous Plastics Products
	31:	Leather and Leather Products
	32:	Stone, Clay, Glass, and Concrete Products
	33:	Primary Metal Industries
	34:	Fabricated Metal Products, Except Machinery and Transportation Equipment
	35:	Industrial and Commercial Machinery and Computer Equipment
	36:	Electronic and Other Electrical Equipment and Components, Except Computer Equipment
	37:	Transportation Equipment
	38:	Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods; Watches and Clocks
	39:	Miscellaneous Manufacturing Industries
E1: Transportation and Communications	40:	Railroad Transportation
	41:	Local and Suburban Transit and Interurban Highway Passenger Transportation
	42:	Motor Freight Transportation and Warehousing
	43:	United States Postal Service
	44:	Water Transportation
	45:	Transportation by Air
	46:	Pipelines, Except Natural Gas
	47:	Transportation Services
	48:	Communications

E2: Electric, Gas, and Sanitary Services	49:	Electric, Gas, and Sanitary Services
F: Wholesale Trade	50:	Wholesale Trade-durable Goods
	51:	Wholesale Trade-non-durable Goods
G: Retail Trade	52:	Building Materials, Hardware, Garden Supply, and Mobile Home Dealers
	53:	General Merchandise Stores
	54:	Food Stores
	55:	Automotive Dealers and Gasoline Service Stations
	56:	Apparel and Accessory Stores
	57:	Home Furniture, Furnishings, and Equipment Stores
	58:	Eating and Drinking Places
	59:	Miscellaneous Retail
H: Finance, Insurance, and Real Estate	60:	Depository Institutions
	61:	Non-depository Credit Institutions
	62:	Security and Commodity Brokers, Dealers, Exchanges, and Services
	63:	Insurance Carriers
	64:	Insurance Agents, Brokers, and Service
	65:	Real Estate
	67:	Holding and Other Investment Offices
I: Services	70:	Hotels, Rooming Houses, Camps, and Other Lodging Places
	72:	Personal Services
	73:	Business Services
	75:	Automotive Repair, Services, and Parking
	76:	Miscellaneous Repair Services
	78:	Motion Pictures
	79:	Amusement and Recreation Services
	80:	Health Services
	81:	Legal Services
	82:	Educational Services
	83:	Social Services
	84:	Museums, Art Galleries, and Botanical and Zoological Gardens
	86:	Membership Organizations
	87:	Engineering, Accounting, Research, Management, and Related Services
	88:	Private Households
	89:	Miscellaneous Services
J: Public Administration	91:	Executive, Legislative, and General Government, Except Finance
	92:	Justice, Public Order, and Safety
	93:	Public Finance, Taxation, and Monetary Policy
	94:	Administration of Human Resource Programs
	95:	Administration of Environmental Quality and Housing Programs
	96:	Administration of Economic Programs
	97:	National Security and International Affairs
	99:	Nonclassifiable Establishments

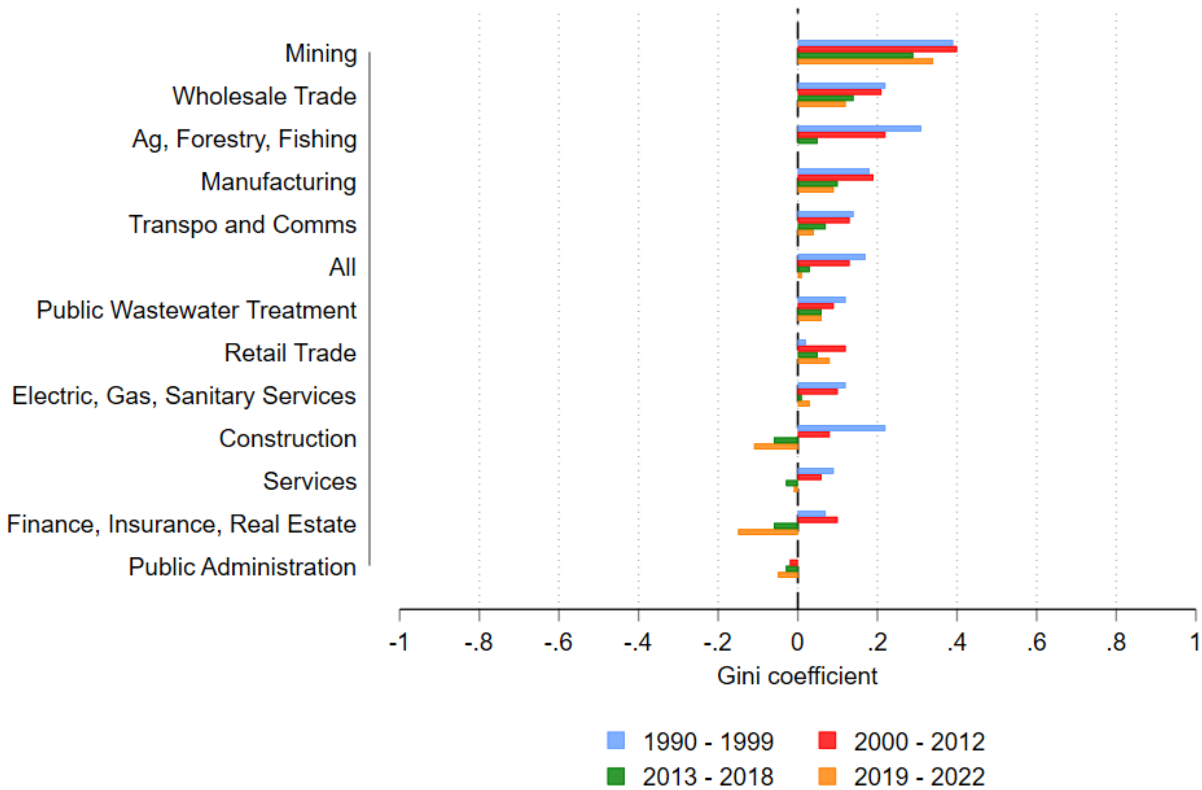
Notes: This table presents the names of the major groups, identified by the two-digit SIC code, within each industrial division. We split industrial Division E (Transportation, Communications, Electric, Gas, and Sanitary Services) into two: (1) E1: Transportation and Communications and (2) E2: Electric, Gas, and Sanitary Services. For additional information on these industries, please refer to https://www.osha.gov/data/sic-manual.

Appendix Figure 1. Gini Coefficients for Distribution of Outfalls by Industry and by Race for the Top 40% of Census Block Groups in Terms of Rural Population Share.



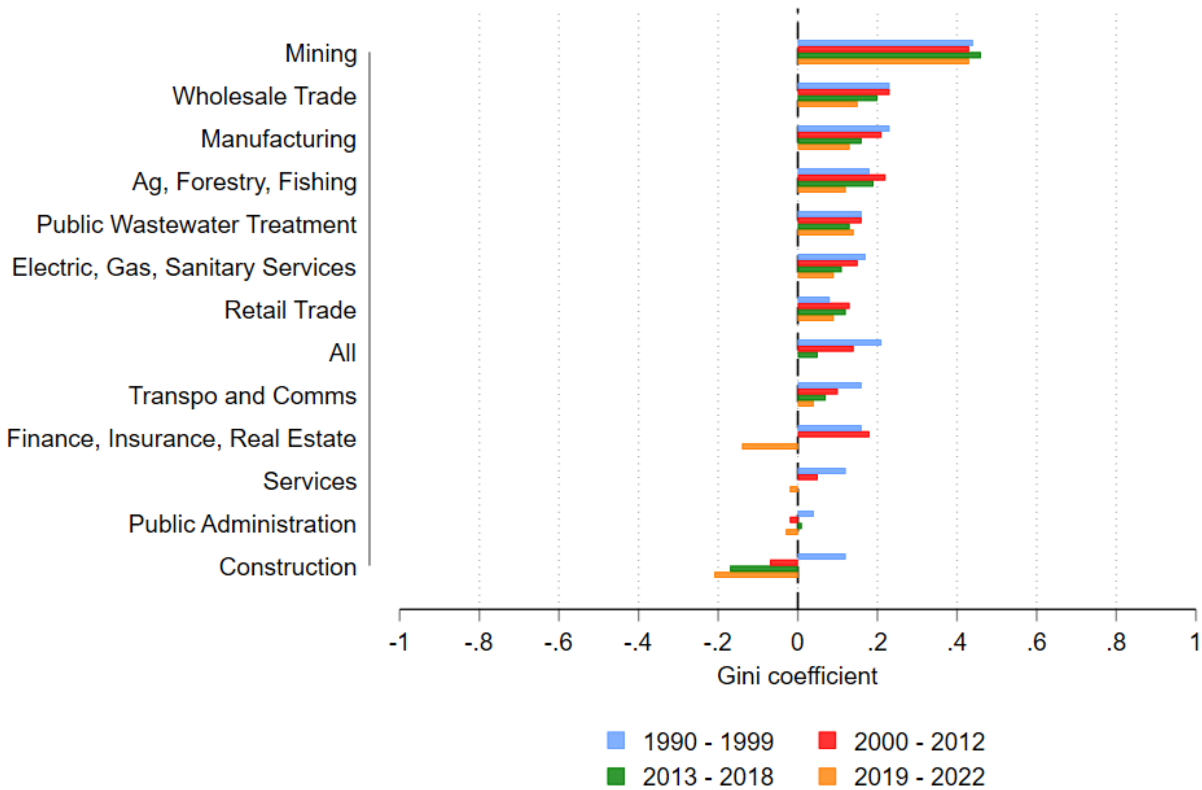
Notes: This graph presents Gini coefficients by industry for the share of the population that is non-White. The category "All" depicts Gini coefficients across all facilities, regardless of the availability of the industrial code. Blue bars are for the 1990 - 1999 period, red bars are for the 2000 - 2012 period, green bars are for the 2013 - 2018 period, and orange bars are for the 2019 - 2022 period. We restrict the sample to the top 40% of census block groups in terms of rural population share.

Appendix Figure 2. Gini Coefficients for Distribution of Outfalls by Industry and by Poverty for the Top 40% of Census Block Groups in Terms of Rural Population Share.



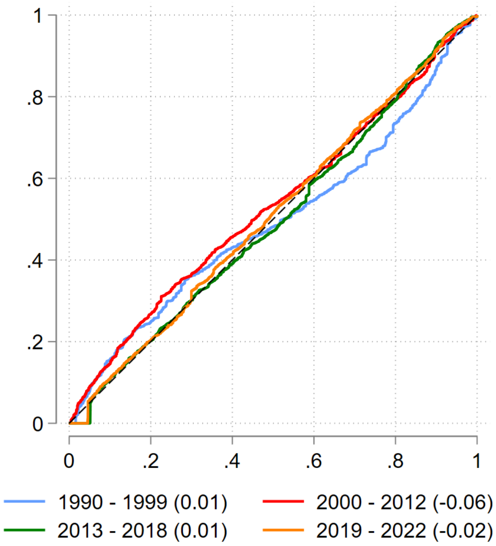
Notes: This graph presents Gini coefficients by industry for the share of the population that is below the poverty line. The category "All" depicts Gini coefficients across all facilities, regardless of the availability of the industrial code. Blue bars are for the 1990 - 1999 period, red bars are for the 2000 - 2012 period, green bars are for the 2013 - 2018 period, and orange bars are for the 2019 - 2022 period. We restrict the sample to the top 40% of census block groups in terms of rural population share.

Appendix Figure 3. Gini Coefficients for Distribution of Outfalls by Industry and by Education for the Top 40% of Census Block Groups in Terms of Rural Population Share.

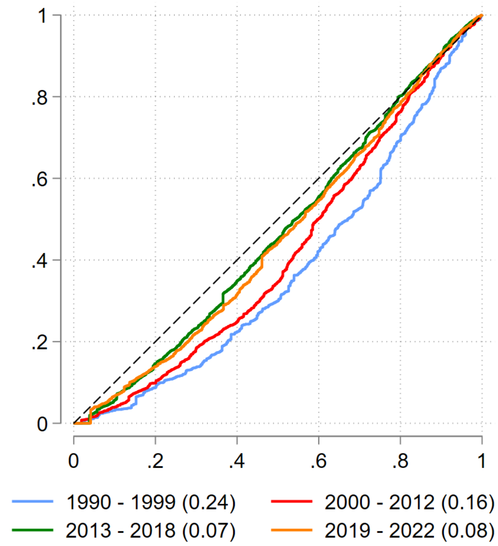


Notes: This graph presents Gini coefficients by industry for the share of the population without a college degree or higher. The category "All" depicts Gini coefficients across all facilities, regardless of the availability of the industrial code. Blue bars are for the 1990 - 1999 period, red bars are for the 2000 - 2012 period, green bars are for the 2013 - 2018 period, and orange bars are for the 2019 - 2022 period. We restrict the sample to the top 40% of census block groups in terms of rural population share.

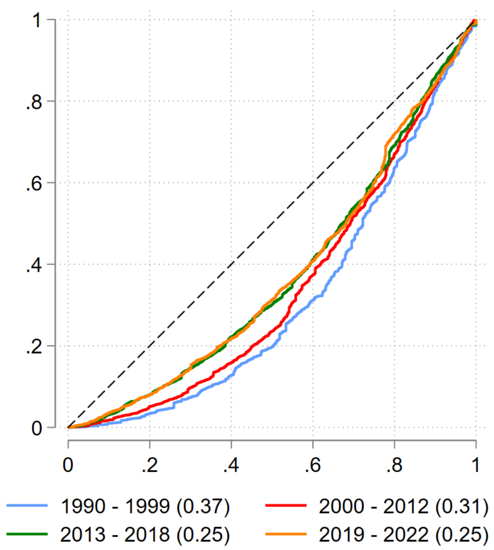
Appendix Figure 4. Pseudo-Lorenz Curves for the Count of All Outfalls by Race, Poverty, Education for Facilities in Chemicals and Allied Products.



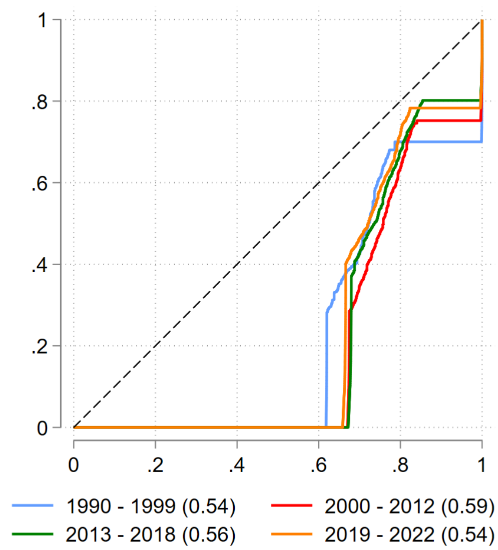
(a) Share Non-White



(b) Share Poverty



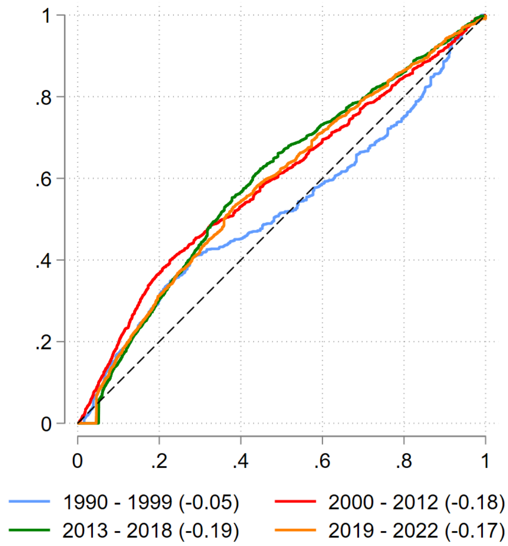
(c) Share Non-College



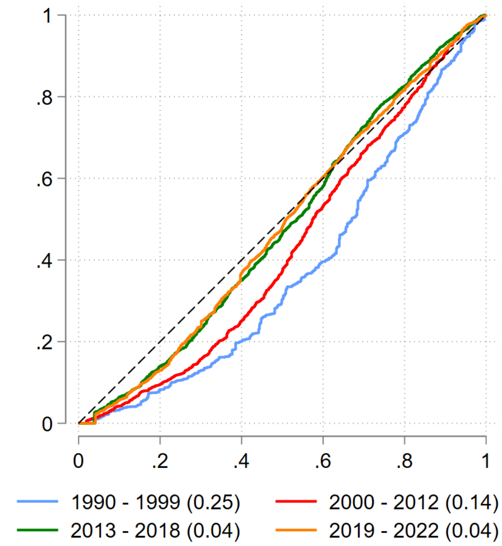
(d) Rural Population Share

Notes: This figure displays pseudo-Lorenz curves for the count of all outfalls from Chemicals and Allied Products (SIC 28) by the share of the population non-White (App. Figure 4a), share of the population below the poverty line (App. Figure 4b), the share of the population without a college education or higher (App. Figure 4c), and the share of the population living in rural areas (App. Figure 4d). The 45-degree line represents equal distribution. Each figure shows how these counts change over time from the 1990-1999 period (blue line) to the 2019-2022 period (orange line). Gini coefficients for each period are in parentheses in the legend.

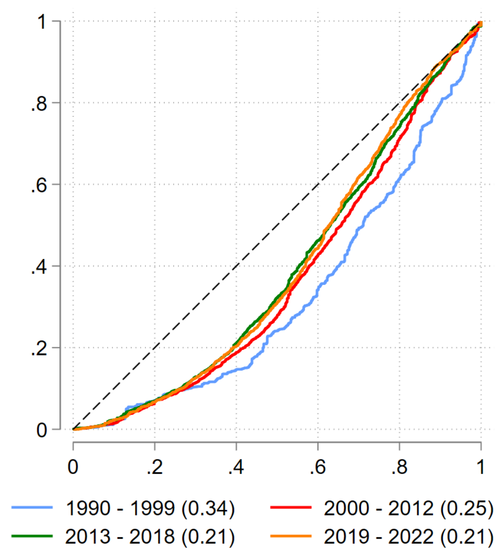
Appendix Figure 5. Pseudo-Lorenz Curves for the Count of All Outfalls by Race, Poverty, Education for Facilities in Petroleum Refining and Related Industries.



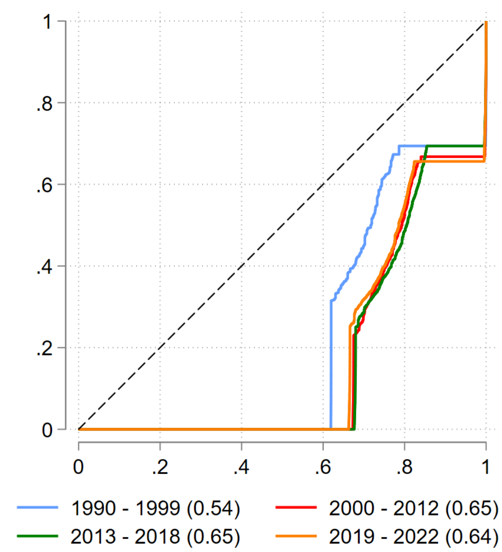
(a) Share Non-White



(b) Share Poverty



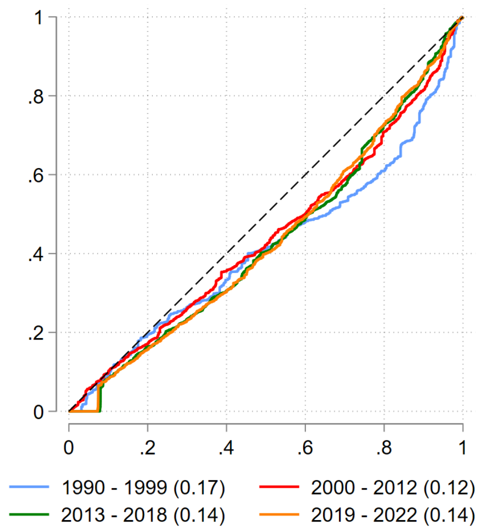
(c) Share Non-College



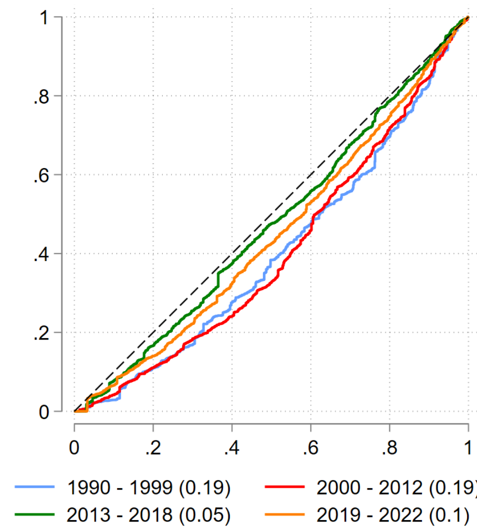
(d) Rural Population Share

Notes: This figure displays pseudo-Lorenz curves for the count of all outfalls from Petroleum Refining and Related Industries (SIC 29) by the share of the population non-White (App. Figure 5a), share of the population below the poverty line (App. Figure 5b), the share of the population without a college education or higher (App. Figure 5c), and the share of the population living in rural areas (App. Figure 5d). The 45-degree line represents equal distribution. Each figure shows how these counts change over time from the 1990-1999 period (blue line) to the 2019-2022 period (orange line). Gini coefficients for each period are in parentheses in the legend.

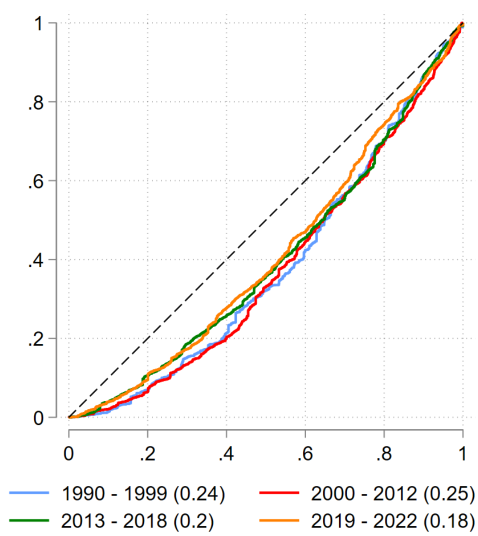
Appendix Figure 6. Pseudo-Lorenz Curves for the Count of All Outfalls by Race, Poverty, Education for Facilities in Chemicals and Allied Products for the Top 40% of Census Block Groups in Terms of Rural Population Share.



(a) Share Non-White



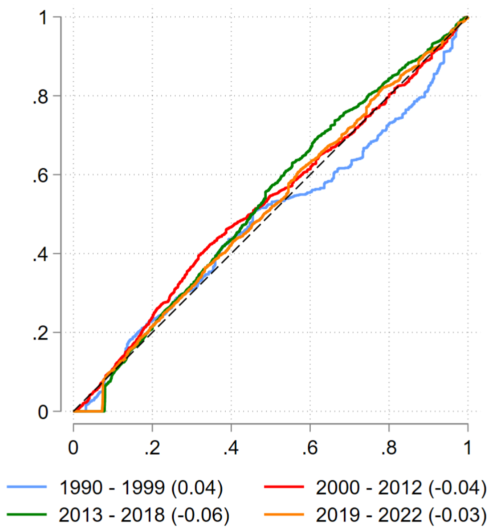
(b) Share Poverty



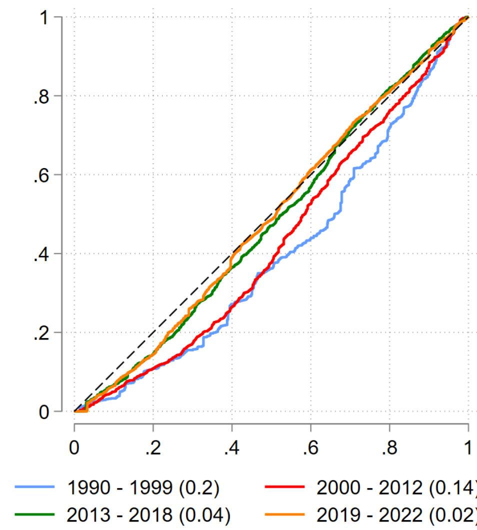
(c) Share Non-College

Notes: This figure displays pseudo-Lorenz curves for the count of outfalls from Chemicals and Allied Products (SIC 28) by the share of the population non-White (App. Figure 6a), share of the population below the poverty line (App. Figure 6b), and the share of the population without a college education or higher (App. Figure 6c). The 45-degree line represents equal distribution. Each figure shows how these counts change over time from the 1990-1999 period (blue line) to the 2019-2022 period (orange line). Gini coefficients for each period are in parentheses in the legend. We restrict the sample to the top 40% of census block groups in terms of rural population share.

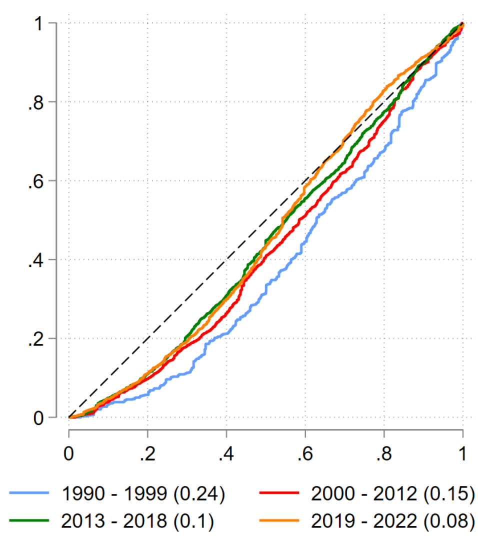
Appendix Figure 7. Pseudo-Lorenz Curves for the Count of All Outfalls by Race, Poverty, Education for Facilities in Petroleum Refining and Related Industries for the Top 40% of Census Block Groups in Terms of Rural Population Share.



(a) Share Non-White



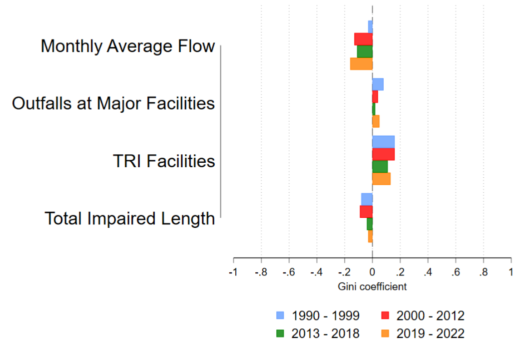
(b) Share Poverty



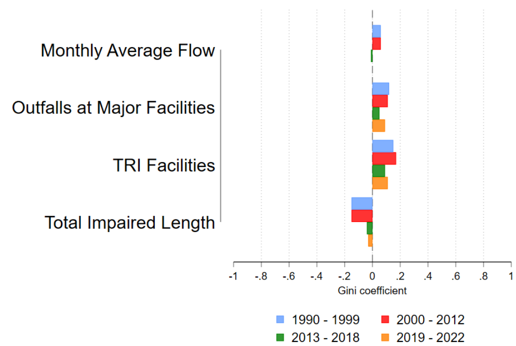
(c) Share Non-College

Notes: This figure displays pseudo-Lorenz curves for the count of outfalls from Petroleum Refining and Related Industries (SIC 29) by the share of the population non-White (App. Figure 7a), share of the population below the poverty line (App. Figure 7b), and the share of the population without a college education or higher (App. Figure 7c). The 45-degree line represents equal distribution. Each figure shows how these counts change over time from the 1990-1999 period (blue line) to the 2019-2022 period (orange line). Gini coefficients for each period are in parentheses in the legend. We restrict the sample to the top 40% of census block groups in terms of rural population share.

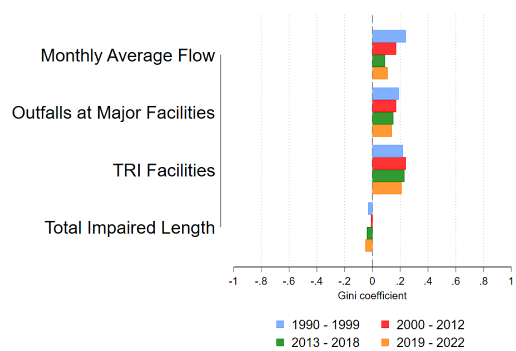
Appendix Figure 8. Gini Coefficients for Other Measures of Pollution for the Top 40% of Census Block Groups in Terms of Rural Population Share.



(a) Share Non-White



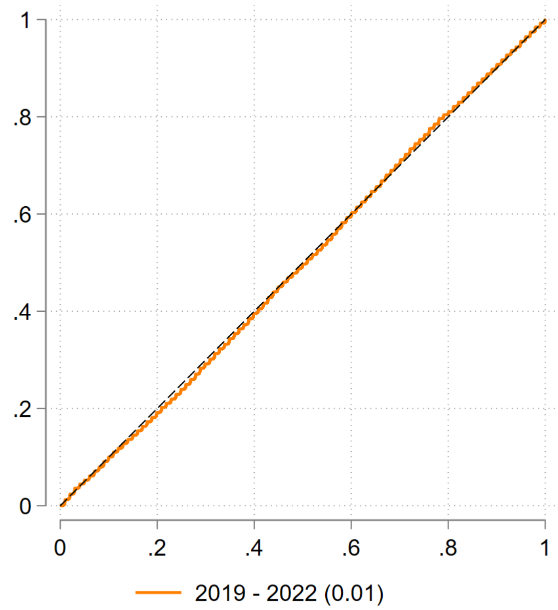
(b) Share Poverty



(c) Share Non-College

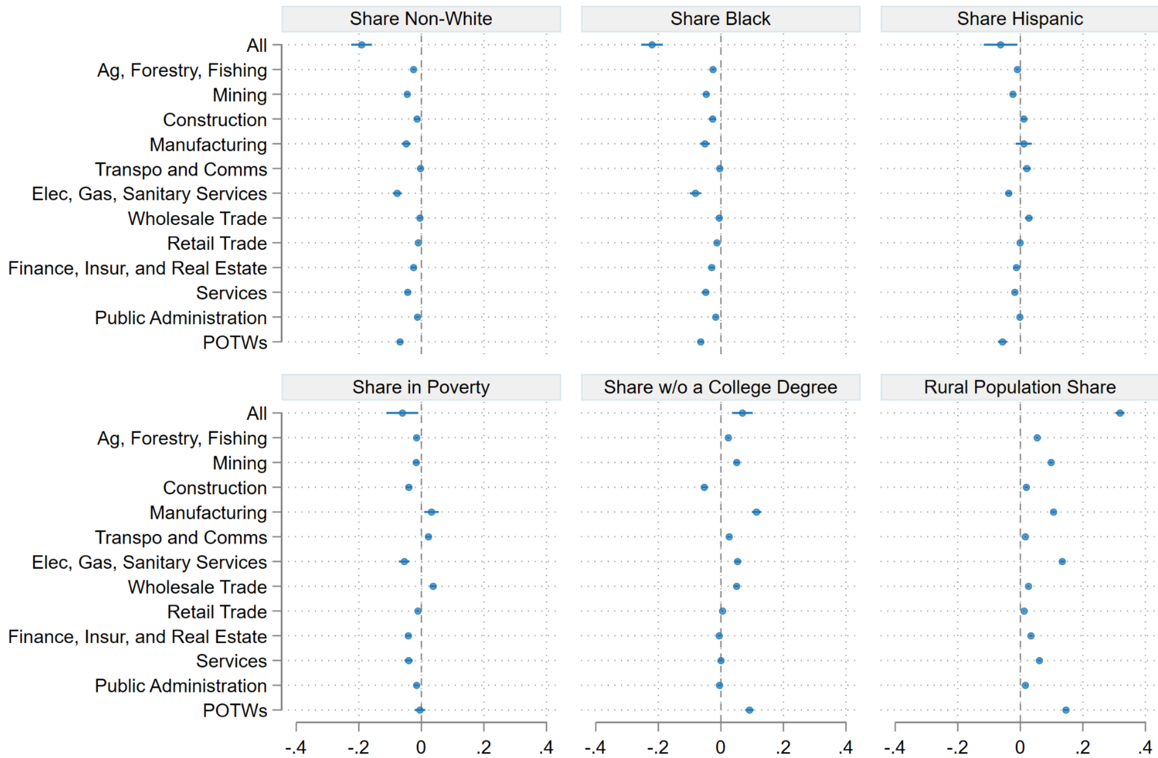
Notes: This figure displays bar graphs depicting Gini coefficients using the following measures: (1) monthly average flow, (2) outfalls at facilities deemed "Major" by USEPA, (3) the number of facilities on the TRI with a water pollution discharge permit, and (4) the total impaired waterway length. For each measure of pollution, we examine the distribution across the share of the population non-White (App. Figure 8a), the share of the population below the poverty line (App. Figure 8b), and the share of the population without a college education or higher (App. Figure 8c). We restrict the sample to the top 40% of census block groups in terms of rural population share.

Appendix Figure 9. Pseudo-Lorenz Curves and Wastewater Discharge.



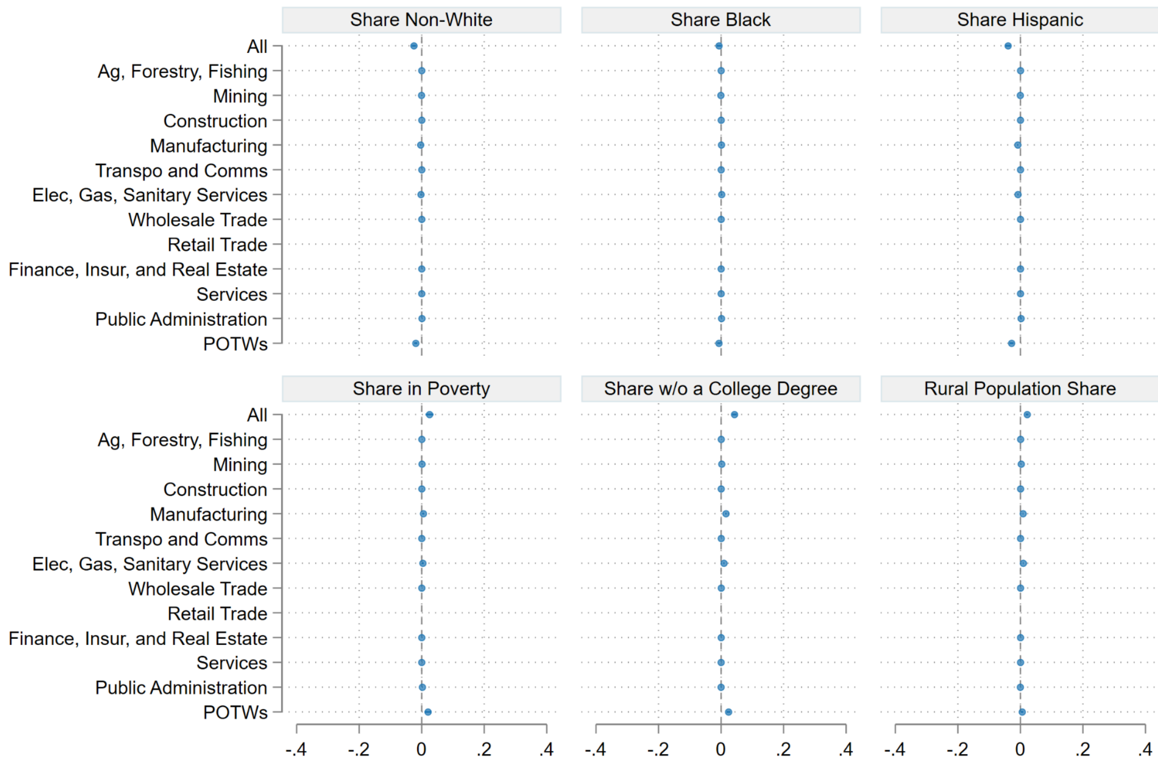
Notes: This figure displays a pseudo-Lorenz curve for the count of all outfalls by the share of wastewater discharge for the 2019 - 2022 period (orange line). The 45-degree line represents equal distribution. The Gini coefficient is in parentheses in the legend.

Appendix Figure 10. Cross-Sectional Results Between Presence of Outfalls and Demographics and State Fixed Effects.



Notes: This figure displays results from cross-sectional regressions of an indicator for the presence of outfalls in a census block group versus a given measure of demographics and state fixed effects. The outcome variable for the category "All" is an indicator for the presence of an outfall in a census block group regardless of industrial classification. This category includes facilities with and without SIC information. For the remaining categories, we use industry-specific indicators for the presence of an outfall in a census block group as the outcome variable. Coefficient estimates are shown in blue dots and 95% confidence intervals are shown by the corresponding lines. Standard errors are clustered at the county level. Results are grouped by demographic variable and by industrial classification. The sample includes all four time periods.

Appendix Figure 11. Cross-Sectional Results Between Presence of Outfalls from Major Dischargers and Demographics.



Notes: This figure displays results from cross-sectional regressions of an indicator of an outfall at major water pollution dischargers in a census block group versus a given measure of demographics. The outcome variable for the category "All" is an indicator for the presence of an outfall at a major facility in a census block group regardless of industrial classification. This category includes major facilities with and without SIC information. For the remaining categories, we use industry-specific indicators for the presence of an outfall at a major facility in a census block group as the outcome variable. Coefficient estimates are shown in blue dots and 95% confidence intervals are shown by the corresponding lines. Standard errors are clustered at the county level. Results are grouped by demographic variable and by industrial classification. The sample includes all four time periods. There are no major dischargers in the retail trade industrial division.

Appendix Figure 12. Cross-Sectional Results Between Mean Monthly Average Flow and Demographics.

Notes: This figure displays results from cross-sectional regressions of the mean of monthly average flow at outfalls versus a given measure of demographics. The outcome variable for the category "All" is the mean of monthly average flow from outfalls in a census block group regardless of industrial classification. This category includes major facilities with and without SIC information. For the remaining categories, we use the mean of monthly average flow from industry-specific outfalls in a census block group as the outcome variable. Coefficient estimates are shown in blue dots and 95% confidence intervals are shown by the corresponding lines. Standard errors are clustered at the county level. Results are grouped by demographic variable and by industrial classification. The sample includes all four time periods.

Appendix Table 3. Number of Outfalls by Industrial Classifications and Corresponding Demographics (1990 - 1999).

Industry	No. of Outfalls	No. of CBGs	Non-White	Non-College	Poverty	Rural Population
Agriculture, Forestry, Fishing	1,754	776	8%	87%	15%	86%
Construction	743	460	13%	84%	16%	55%
Electric, Gas, Sanitary Services	15,247	7,662	11%	85%	14%	66%
Finance, Insurance, Real Estate	3,465	2,253	6%	86%	13%	79%
Manufacturing	22,793	8,253	14%	88%	16%	53%
Mining	12,448	2,923	8%	89%	17%	80%
Public Administration	2,528	1,247	13%	84%	14%	59%
Retail Trade	1,022	839	10%	85%	12%	57%
Wholesale Trade	3,817	2,092	16%	88%	17%	47%
Services	5,132	3,514	8%	85%	14%	78%
Transportation and Communications	3,779	1,873	18%	87%	16%	42%
POTWs	17,666	11,102	11%	86%	15%	62%
All Industries - w/ SIC code	90,394	30,352	12%	86%	15%	61%
All Industries	92,761	32,957	12%	86%	14%	59%
All CBGs (CONUS Average)	92,761	216,330	18%	80%	13%	28%

Notes: This table provides summary statistics of the total number of outfalls and the distribution of outfalls by industrial classification for the 1990 - 1999 period. The table also shows corresponding census block group demographic information for all census block groups and by industrial classification. "All Industries - with industrial code" summarizes these statistics for outfalls that correspond to facilities with at least one industrial code. "All Industries" summarizes these statistics for all outfalls, regardless of the availability of the industrial code. "All CBGs" summarizes these statistics for all census block groups in the conterminous US (CONUS) for comparison purposes.

Appendix Table 4. Number of Outfalls by Industrial Classifications and Corresponding Demographics (2000 - 2012).

Industry	No. of Outfalls	No. of CBGs	Non-White	Non-College	Poverty	Rural Population
Agriculture, Forestry, Fishing	9,210	3,665	11%	85%	13%	84%
Construction	47,263	12,844	13%	77%	13%	42%
Electric, Gas, Sanitary Services	38,125	16,274	13%	81%	12%	60%
Finance, Insurance, Real Estate	7,785	4,092	11%	80%	11%	61%
Manufacturing	55,881	20,005	17%	83%	14%	49%
Mining	42,304	7,451	12%	85%	14%	75%
Public Administration	10,158	4,889	14%	78%	12%	42%
Retail Trade	4,183	3,245	16%	82%	13%	44%
Wholesale Trade	11,257	6,034	18%	84%	15%	46%
Services	13,079	8,069	13%	78%	12%	58%
Transportation and Communications	13,784	6,714	18%	81%	14%	38%
POTWs	26,857	16,419	12%	82%	13%	60%
All Industries - w/ SIC code	279,886	59,633	15%	80%	13%	48%
All Industries	338,288	72,952	16%	79%	12%	43%
All CBGs (CONUS Average)	338,288	216,330	22%	75%	13%	23%

Notes: This table provides summary statistics of the total number of outfalls and the distribution of outfalls by industrial classification for the 2000 - 2012 period. The table also shows corresponding census block group demographic information for all census block groups and by industrial classification. "All Industries - with industrial code" summarizes these statistics for outfalls that correspond to facilities with at least one industrial code. "All Industries" summarizes these statistics for all outfalls, regardless of the availability of the industrial code. "All CBGs" summarizes these statistics for all census block groups in the conterminous US (CONUS) for comparison purposes.

Appendix Table 5. Number of Outfalls by Industrial Classifications and Corresponding Demographics (2013 - 2018).

Industry	No. of Outfalls	No. of CBGs	Non-White	Non-College	Poverty	Rural Population
Agriculture, Forestry, Fishing	10,806	4,166	12%	82%	14%	82%
Construction	89,534	21,192	18%	74%	16%	35%
Electric, Gas, Sanitary Services	47,015	17,966	16%	78%	15%	55%
Finance, Insurance, Real Estate	10,340	4,983	15%	75%	13%	50%
Manufacturing	87,741	26,982	21%	80%	17%	41%
Mining	76,632	10,098	14%	81%	15%	70%
Public Administration	12,185	5,934	18%	75%	15%	35%
Retail Trade	4,845	3,427	20%	79%	17%	41%
Wholesale Trade	19,291	9,340	23%	82%	19%	38%
Services	14,400	8,595	16%	75%	14%	51%
Transportation and Communications	26,483	11,285	24%	78%	18%	30%
POTWs	27,512	16,777	14%	80%	16%	58%
All Industries - w/ SIC code	426,784	72,015	19%	77%	16%	41%
All Industries	631,160	96,629	20%	75%	16%	35%
All CBGs (CONUS Average)	631,160	216,330	25%	73%	16%	21%

Notes: This table provides summary statistics of the total number of outfalls and the distribution of outfalls by industrial classification for the 2013-2018 period. The table also shows corresponding census block group demographic information for all census block groups and by industrial classification. “All Industries - with industrial code” summarizes these statistics for outfalls that correspond to facilities with at least one industrial code. “All Industries” summarizes these statistics for all outfalls, regardless of the availability of the industrial code. “All CBGs” summarizes these statistics for all census block groups in the conterminous US (CONUS) for comparison purposes.