Policy Spillovers to Biodiversity: The Impact of Clean Water Act Investments on Birds

*Anthony Ponce

Last Updated: November 10, 2026

Abstract

North America's birds have declined by 29% since 1970, yet few studies causally identify whether environmental policy slows such losses. This paper asks how federal investment in public infrastructure affects biodiversity, using the wastewater treatment plant construction projects commissioned through the Clean Water Act's construction grants program. Leveraging the plants' staggered commissioning from 1967 to 1999 and the Audubon Christmas Bird Count in an event-study difference-in-differences design, I find that grant-funded commissioning raised the number of water-linked bird species recorded downstream of grant-receiving plants by about 5%, roughly 1 more species per survey. The number of water-linked birds rose about 14%, roughly 830 more birds per survey, after filtering extreme count values. Compared with a counterfactual without the grants, this amounts to about 3% more water-linked species from 1967 to 2024, about 870 more species records across all count sites each winter. However, survey sites receiving more of a plant's treated effluent tied to grant-funded commissioning, relative to the river's flow, showed no greater gains than sites receiving less. These estimates suggest that federal environmental investment reaches biodiversity, which the Act's benefit-cost assessments have so far left out.

Keywords: valuation of environmental effects, environmental policy, ecosystem services, biodiversity conservation, infrastructure and public investment

*Cornell University Dyson School of Applied Economics & Management. Email: ap795@cornell.edu. I thank my advisors, Catherine L. Kling, Nicholas J. Sanders, and Sheila M. Olmstead, for their guidance and support throughout this paper. Many thanks also to Wendong Zhang, Todd Gerarden, Rihannon Jerch, Ivan Rudik, Ariel O. Balboa, Amanda D. Rosewald, Matthew Gibson, Sophie Gilbert, Eric Edwards, Prakash Mishra, Patrick Baylis, and Eric Lewis for thoughtful discussions and comments throughout various stages of the paper. I am grateful for the feedback of participants and the communities at AAEA, NAREA, PERC's Graduate Fellowship, Camp Resources, the CU Environmental & Resource Economics Workshop, and the Cornell SC Johnson SEERE Seminar Series. I an deeply grateful to all of these organizations and their community. All errors are my own.

1 Introduction

Since 1970, the breeding birds of the United States and Canada have fallen by nearly three billion, 29% of their 1970 abundance (Rosenberg et al., 2019). Over the same half century, federal grants funded new and improved municipal wastewater treatment plants across American cities, projects costing approximately $650 billion over their operational lifetimes in 2014 dollars (Keiser and Shapiro, 2019a). Most of the benefits of cleaner surface waters accrue through recreational use and ecosystem health (Olmstead, 2010), yet a half-century review of US pollution policy mostly excludes biodiversity, as it does fisheries, forests, and groundwater (Shapiro, 2022). Valuing so large a public investment for the ecosystem services it provides begins with measuring what it accomplished for the wildlife that depends directly on clean water.

What are the impacts of federal investments in public infrastructure on ecological changes in biodiversity? The construction grants under Title II of the Clean Water Act (CWA), and the bird communities downstream of the treatment plants they funded, provide the setting in which this paper studies that question. I begin by estimating the level effect of a plant's grant-funded commissioning on the birds downstream: how many birds (abundance) and how many species (richness) the surveys record, and how diverse the community is. Then, I test whether that effect scales with each plant's contribution to cleaning, its treated effluent relative to the river's flow where it enters, and set that contribution against how much cleaner the water got.

To estimate the effect, I compare the birds recorded in the survey circles of the National Audubon Society's Christmas Bird Count (CBC) that lie downstream of a treatment plant, before and after the plant's grant-funded commissioning, with circles whose plants were commissioned later or not at all. The comparison is staggered over 29 commissioning cohorts from 1967 to 1999, and the river network links each circle to the plants upstream. I estimate the level effect in an event study with circle and year fixed effects, alongside controls for survey effort, winter weather, and baseline housing and cropland. Its pre-treatment coefficients test whether treated and comparison circles followed parallel trends before commissioning.

The design addresses two identification challenges. First, municipalities might have built treatment plants where bird populations were already changing. However, a circle's pre-program bird counts do not predict when its first plant was commissioned. Second, because volunteers conduct the counts, birder effort might respond to a plant's completion. However, commissioning alters neither which circles are surveyed nor the effort put into a survey in the year of commissioning.

To evaluate whether the effect scales with how much treated effluent a circle receives,

I measure each plant's treated effluent relative to the river's flow where it enters and sum it across the commissioned plants upstream of a circle. The sum is the circle's effluent treated flow share, which I call the dose, as the literature on continuous treatments calls such a measure (Callaway et al., 2025). The main estimation sample holds 85,934 survey observations across 2,975 circles, of which 2,372 lie downstream of a grant plant commissioned during the sample.

In the main specification, commissioning a grant-funded treatment plant upstream raises species richness for all birds by about 2%, roughly one additional species per survey on the treated circles' average survey of 65 species and on their median survey of 59.[1] The estimate keeps its sign across nine specifications. Abundance on the full counts and the three diversity statistics (Shannon entropy, the effective number of species, which is Hill-Shannon diversity of order 1, the exponential of the Shannon entropy, and the effective number of common species), which I read as robustness checks, change in none of the six species groups at the 5% level. The rise concentrates among the water-linked species, the 248 of the survey's 935 species that forage on or in the water or live in wetlands and take a food whose supply water quality governs. I classify them from the foraging traits of EltonTraits (Wilman et al., 2014), the habitats of the AVONET trait database (Tobias et al., 2022) and of the Partners in Flight Avian Conservation Assessment Database (Partners in Flight, 2024), and the foraging guilds of De Graaf et al. (1985). Among these species, commissioning raises richness by about 5%, roughly one additional species on their average survey of 18 and one more every two surveys on their median survey of 13. Their abundance rises by about 14% after filtering each year's extreme counts, approximately 830 additional water-linked birds on the average survey and 85 on the median one, while the estimate on the full counts is imprecise. Within the water-linked species, richness rises about 8% for the bottom-sediment feeders, about 7% for the scavenging gulls, and about 6% each for waterbirds, waterfowl, and the submerged-plant grazers, and the grazers' abundance rises in both forms, by about 42% on the full counts and 28% with the extreme counts dropped. About half of the water-linked species' rise in richness continues a linear pre-commissioning trend.

The effect does not rise with the dose. Circles that received more of a plant's treated effluent relative to the river's flow recorded no more species than circles that received less, in a comparison of the circles above and below the median dose, whose pre-trend tests do not reject for richness and the diversity statistics. Each additional commissioned plant up-

[1] The average survey divides the total count of the treated circles' surveys in the nine years before commissioning by the number of those surveys. The median survey is the one in the middle when those surveys are ranked by their count: half recorded fewer and half more. A few surveys record very large counts, which puts the average survey far above the median one. Thus, one percentage is many more birds on the average survey than on the median one.

stream, by contrast, raises richness by about 1% for waterbirds and waterfowl and 3% for the submerged-plant grazers, and not at all for landbirds, 20 of whose 350 species are water-linked. Responses thus follow the physical presence of commissioned plants rather than their share of the river's flow. Downstream water quality also improved after commissioning: dissolved oxygen saturation rose by 0.15 within-circle standard deviations, about 2 percentage points of saturation, and total phosphorus fell by 0.15, about 0.1 milligrams per liter. However, the ways of separating the water's part of the effect from the plants' other channels disagree, and the data cannot yet say how much of the effect operates through water quality.

Over the full program horizon, from the first commissioning in 1967 through 2024, the estimates imply that the surveys recorded 3% more water-linked species than under the counterfactual scenario in which no plant upstream of a circle is commissioned with a grant (95% interval 1 to 5). That difference is about 870 more entries of these species on the surveys' species lists in an average year (320 to 1,400). Scaled to continental populations following Rosenberg et al. (2019), with the extreme counts dropped, the rise in water-linked birds totals 65.4 million (5.4 to 118.7 million). This continental figure describes the populations at the date of those estimates, not a sum over years, and it is an upper bound because it treats the surveyed circles as standing for the continent.

The Title II construction grants are well suited to evaluating the ecological effects of public infrastructure for three reasons. First, the program was large: it constituted "the largest nonmilitary public works program since the Interstate Highway System" (Ramseur, 2023, p. 2). Second, project timing is precisely recorded: the grant record holds 8,778 plants built with construction grants, nearly all with a dated completion, and their commissioning was staggered over three decades, which lets a circle below a plant commissioned in one year be compared with circles whose plants were commissioned later. Third, bird monitoring spans the entire program era: CBC circles keep fixed boundaries surveyed winter after winter from 1900 to the present, capturing decades of pre-treatment history.

The direction of bird community responses to wastewater treatment is theoretically ambiguous. On the one hand, removing nutrients can restore submerged vegetation and the waterbirds that feed on it: phosphorus removal at a plant on a Dutch lake was followed by the return of stoneworts and of the waterbirds that graze them (Noordhuis et al., 2002). On the other hand, untreated sewage outfalls enriched the invertebrates that some diving ducks wintered on, and a primary treatment plant that began taking Edinburgh's sewage cut those ducks' numbers (Campbell, 1984). The net direction of the response is therefore an empirical question.

Bird populations serve as an ecological policy outcome in their own right. Society values wildlife directly, through recreational birdwatching demand and non-use existence values,

as well as indirectly, through ecosystem services. In formal ecosystem service accounting, species populations constitute the "final ecosystem service" behind birding and existence value, whereas chemical water quality parameters such as dissolved oxygen serve as intermediate ecological processes (Boyd and Banzhaf, 2007). Methodological guidelines prioritize indicators located further along the causal chain from policy inputs to ecological benefits (Olander et al., 2018), making bird counts a more direct measure of a policy's ecological effect than readings of the water alone.

This paper makes three contributions. First, it advances environmental policy valuation by providing causal estimates of unmeasured ecosystem services and biodiversity spillovers within the ongoing research agenda evaluating the Clean Water Act. The Clean Water Act's Title II construction grants provide a setting for environmental economics: the 1972 law directly overlaps with the half-century decline in North American bird populations, a net loss of nearly three billion birds since 1970 (Rosenberg et al., 2019), and supplies multi-decade capital investment variation, staggered across receiving river networks. While an active literature continues to evaluate the Clean Water Act's full spectrum of benefits, retrospective studies have focused primarily on housing market appreciation (Keiser and Shapiro, 2019a), municipal sewerage capital accumulation (Flynn and Smith, 2022), downstream infant birth weights (Flynn and Marcus, 2025), municipal user fees (Jerch, 2026), or wetland flood mitigation (Taylor and Druckenmiller, 2022).[2] While these evaluations uncover economic co-benefits, they abstract from direct ecological impacts and biodiversity changes, leaving non-market ecosystem services unquantified (Keiser and Shapiro, 2019b; Keiser et al., 2019; Shapiro, 2022). Grounded in ecosystem service valuation frameworks (Daily et al., 2009; Boyd and Banzhaf, 2007; Polasky et al., 2005; Olander et al., 2018) and in the federal guidelines of Circular A-4 and Circular A-94 (Office of Management and Budget, 2023a,b), which ask the analyst to present non-monetized physical ecosystem effects alongside market metrics, this paper uses the Clean Water Act's plausibly exogenous infrastructure rollout to show that federal water pollution policy generates spillovers to biodiversity in natural physical units.

Second, it adds a public infrastructure investment to the small literature using volunteer survey networks to evaluate environmental policy impacts on biodiversity (Baylis et al., 2026). Closely related quasi-experimental studies evaluate air pollution regulation (Liang et al., 2020), shale gas wells and wind turbines (Katovich, 2024), and forest infrastructure in India (Madhok, 2025). Here the treatment is a federal grant that built public infrastructure, its timing is staggered over 29 commissioning cohorts, and the design tests

[2]Of each grant dollar, 45 cents went to municipal sewerage capital (Flynn and Smith, 2022). Infant birth weight increased by 8.21 grams in counties downstream of plants that had to upgrade (Flynn and Marcus, 2025). Secondary treatment compliance tripled residential sewer fees (Jerch, 2026). Each hectare of wetland lost raised annual flood insurance claims by $1,840 to $1,900 (Taylor and Druckenmiller, 2022).

observer participation and effort as potential selection outcomes before reading the counts. Volunteers choose where and how hard to survey, and analyses of the CBC are limited by that choice (Link et al., 2006; Meehan et al., 2019). A change in surveying that followed commissioning would therefore read as a change in birds, and commissioning changed neither.

Third, it measures treatment continuously, by how much treated effluent a circle receives relative to the flow of its river, and asks whether ecological responses scale with that measure and whether the effect operates through water quality or through other channels. Following the flow share that Fan (2025) builds for dams, I divide each commissioned plant's design flow by the river's mean annual flow where it enters and sum the ratios across the plants upstream of a circle. Circles below the same plant then differ in how much of their water it treated, and a circle upstream of one plant can still receive effluent from another. With the literature on continuous treatments (Callaway et al., 2025), I ask whether the effect grows with a plant's contribution to cleaning and how much of it operates through water quality. Richness does not rise with the dose: responses follow the number of commissioned plants upstream rather than their share of the river, and the part of the effect that operates through the water cannot yet be separated from other channels. The measure thus separates an effect of a plant's presence from one that grows with its effluent's share of the river.

The paper proceeds as follows. Section 2 reviews how studies measure a policy's effect on biodiversity and describes how the construction grants worked. Section 3 describes the data and the sample. Section 4 presents the method and results for the extensive margin, whether a grant-funded plant was commissioned upstream of a circle, and Section 5 does the same for the intensive margin, how much the effect scales with the dose and the water. Section 6 sets the effects against the program's benefits and costs and states the study's limitations and the questions that follow. Section 7 concludes.

2 Background

2.1 Ecosystem Services and Biodiversity Valuation

How do changes in the birds downstream of municipal wastewater infrastructure enter economic welfare? The policy valuation literature measures environmental impacts in two ways: through recreational use value, the demand for birdwatching and wildlife viewing, and through the ecosystem services that wildlife populations supply.

The first approach estimates recreational use values for observable species presence and abundance. In a site-choice model of eBird members' birding trips in Washington and

Oregon, Kolstoe and Cameron (2017) estimate that birders are willing to pay $3.38 per trip in June, in 2012 dollars, for an additional expected species at a site. In Alberta, Jayalath et al. (2023) estimate that viewing an additional species is valued at $0.68 per trip in 2016 Canadian dollars, with a species at risk valued at $5.61 against $0.41 for a species not at risk. Using the 2017 recognition of the Cassia crossbill as a full species as a natural experiment, Wardle (2025) estimates that the newly recognized rare species raised the annual consumer surplus of eBird visitors to the Idaho ranger district that holds it by $73,857, or $277 per visitor, with the value concentrated in "a small but intensely interested group of birders." At 665 Indiana sites during the fall migration of 2021, counts of rails, cranes, and waterfowl carried no premium over other species (Melstrom and Reeling, 2026). Beyond birds, Lewis et al. (2022) estimate that Pacific Northwest households valued a permanent increase of 1,190 returning coho salmon at $63 million in present value.

The second approach estimates the value of the ecosystem services that wildlife populations supply. Waterbirds act "as predators, herbivores and vectors of seeds, invertebrates and nutrients" (Green and Elmberg, 2014). Appendix Table B12 lists the services that review reports for the feeding guilds of five species the survey records. After veterinary diclofenac drove India's vultures to functional extinction, all-cause human mortality rose 4.7% in districts highly suitable for vultures relative to districts poorly suited to them (Frank and Sudarshan, 2024). Similarly, a 10% decline in avian biodiversity, measured from the variation in migratory birds' return that winter habitat in South America produces, reduces US crop revenues by 1.1% through the pest control these birds supply (Noack et al., 2025).

When defining biodiversity metrics, empirical studies often report abundance, species richness, diversity indices, or the frequency of species observation on the same datasets without consistently motivating their choice of metric (Baylis et al., 2026).[3] As Dasgupta (2021, p. 53) emphasizes, an ecosystem's productivity and resilience depend less on the taxonomy of its species than on the functions they perform, and "a mere headcount of species can mislead." These frameworks motivate the paper's three measures: total abundance (N), species richness (S), and effective species diversity ($\exp(H)$). The valuation literature prices the abundance of named species, which supports abundance as a metric, and the presence of an additional species at a site, which supports species richness. The ecosystem services birds provide depend on the functions species perform and on their distribution, which supports the effective number of species, read as a robustness check on the counts. Section 3 defines

[3]Early biodiversity economic theory defined diversity strictly by species extinction risk rather than population size (Solow et al., 1993). Ecological theory defines community diversity through the relative abundances of species, with richness, the Simpson index, and the Shannon index belonging to one one-parameter family (Polasky et al., 2005).

these statistics.

A review of causal inference in conservation finds that few studies rigorously identify the drivers of biodiversity decline, the policies that mitigate it, or its costs (Baylis et al., 2026). Of 186 studies in a global meta-analysis comparing biodiversity under a conservation action against a counterfactual, 44.74% relied on simple before-after or treatment-control comparisons.[4] Recent quasi-experimental studies measure avian responses directly, on the same statistics this paper uses. In England, 100 more hectares of set-aside raised bird abundance by 1 to 2% and richness by 0.3 species, left the Shannon index unchanged, and lowered phylogenetic diversity (Palmer et al., 2026). Designation as a national key rural tourism village in China lowered a county's predicted bird species richness by 0.61 species (Hu et al., 2026). In India, ten square kilometers of infrastructure encroachment into forests reduced the species diversity an eBird user observed by 4%, a reduction more than halved where local institutions gave tribal communities informed consent (Madhok, 2025).[5] (Baylis et al., 2026) identifies Liang et al. (2020) as the only rigorous study that clearly documents the impact of a pollution control regulation on biodiversity. Under the NOx Budget Trading Program, land bird abundance rose, and the estimates for waterfowl, shorebirds, and waterbirds were small and statistically insignificant. Among the consequences of economic activity, Liang et al. (2025) identify noise and water pollution as channels with few or no causal studies. A causal estimate of a water-quality channel to birds exists for harm: construction of a coal-fired power plant lowered dissolved oxygen by 5.1% at nearby monitoring stations and cut the stopover duration of tracked migratory birds within 50 kilometers by 54.4% (Chen et al., 2026). Reviewing the federal benefit-cost record of water-quality policy, Keiser and Shapiro (2019b) note that "existing estimates abstract from ecological consequences such as loss of biodiversity."

The first half of this paper therefore estimates the level effect of grant-funded commissioning on these measures of the birds downstream of a treatment plant (Section 4), the program's extensive margin: whether a grant-funded plant was commissioned upstream of a circle. The second half asks how much, the intensive margin: whether the effect scales with the dose and operates through water quality (Section 5). The two terms follow Fan (2025), who distinguishes "the extensive margin decision to build a dam" from the intensive margin, on which policy can reduce the extent of the dam's effects.

[4] Of the 26 quasi-experimental designs among the 186 studies, 84.62% used land use as a proxy rather than a direct measure of biodiversity (Baylis et al., 2026).

[5] Exogenous environmental shocks move similar metrics: larger farms east of the former inner German border lowered species richness (Noack et al., 2022), onshore wind turbines lowered abundance and richness, with waterfowl the least affected group (Meng et al., 2025), and wildfire smoke lowered richness, the Shannon index, and phylogenetic diversity (Meier and Strobl, 2026).

2.2 Title II Construction Grants

The Clean Water Act of 1972 (33 U.S.C. § 1251 et seq.) declared its objective "to restore and maintain the chemical, physical, and biological integrity of the Nation's waters." Its interim goal was "water quality which provides for the protection and propagation of fish, shellfish, and wildlife and provides for recreation in and on the water" by July 1, 1983 (U.S. Code, 2018). To achieve these objectives, the Act declared as its fourth policy "that Federal financial assistance be provided to construct publicly owned waste treatment works" (33 U.S.C. § 1251(a)). While the statutory goals named wildlife, the rules that allocated the grants gave their primary purpose as assisting municipalities in meeting the Act's enforceable requirements (40 CFR 35.901; Office of the Federal Register, 1999). State priority lists ranked candidate projects on the severity of the pollution problem, the population affected, and the preservation of high-quality waters (40 CFR 35.915(a)), and the facilities plan chose the alternative with "the lowest present worth or equivalent annual value unless nonmonetary costs are overriding" (40 CFR Part 35, Subpart E, Appendix A, item 6a).[6] The 1972 amendments required every publicly owned treatment plant to achieve secondary treatment by July 1, 1977 (33 U.S.C. § 1311(b)(1)(B)).[7] Title II funded 75% of eligible construction costs (33 U.S.C. § 1282(a)(1)), and federal grants disbursed nearly $52 billion between fiscal years 1973 and 1990, "the largest nonmilitary public works program since the Interstate Highway System" (Ramseur, 2023, p. 2). Appendix Figure B1 shows the grants by period of commissioning, split in its three panels by the category of need they paid for, the plant's level of treatment, and its technology.

Congress appropriated funds against five-year authorizations, and the Environmental Protection Agency (EPA) allotted each appropriation among the States by a statutory formula of needs and population (33 U.S.C. § 1285; Ramseur, 2016, Table 1, p. 5).[8] Each State ranked candidate projects on an annual priority list under a priority system EPA approved, whose mandatory criteria were the severity of the pollution problem, the existing population affected, and the need to preserve high-quality waters, and which could not weigh development needs, geographic region, or population growth (40

[6] Where the regulations wanted habitat weighed, they wrote it in: the Clean Lakes program's review criteria included "the estimated improvement in fish and wildlife habitat" (40 CFR 35.1640-1(a)(3)). Within the grants, habitat entered at the margin of the constructed wetland: "over 30 such systems" were financed through fiscal 1990 (U.S. General Accounting Office, 1994, chapter 2), against 600 innovative and 2,100 alternative projects funded in fiscal 1979 to 1987 (U.S. General Accounting Office, 1994, chapter 1).

[7] Secondary treatment means 30 milligrams per liter of five-day biochemical oxygen demand and of suspended solids as a thirty-day average, 85% removal, and a pH between 6.0 and 9.0 (40 CFR 133.102; Office of the Federal Register, 2026).

[8] The formula's weights changed at every reauthorization. From 1983 to 1986 they gave backlog needs 50%, population 37.5%, and total needs 12.5%, with an 80% hold-harmless (Ramseur, 2016, pp. 5, 12).

CFR 35.915(a); U.S. Environmental Protection Agency, 1980, II-8 to II-9).[9] Compliance lagged the statute. By July 1, 1977, about 40% of major municipal dischargers had secondary capability, against about 80% of major industrial dischargers using best practicable technology (U.S. General Accounting Office, 1980, p. 28), and more than 10,000 municipalities missed the deadline (U.S. Environmental Protection Agency, 1979, p. 5; U.S. General Accounting Office, 1980, p. 31). Before 1984, the federal failure to provide funds "was an acceptable excuse for noncompliance" (Stiefel, 1994), and a plant whose federal assistance came too late could extend its permit deadline to a date keyed to the availability of funds, no later than July 1, 1983, and later July 1, 1988 (33 U.S.C. § 1311(i)(1)). The 1984 National Municipal Policy then required every plant to comply by July 1, 1988 "whether or not they receive Federal funds" (U.S. Environmental Protection Agency, 1984, 49 FR 3832).[10] The Comptroller General reported to Congress in 1980, under the title "Costly Wastewater Treatment Plants Fail To Perform As Expected," that "many are not treating wastewater at the efficiency levels they were designed to achieve" (U.S. General Accounting Office, 1980, cover and p. i). Plant commissioning dates thus spread across the years the program paid out, determined by the grant queue and the enforcement schedule, and the criteria that ranked projects named the population affected and the uses of the water.

Title II construction grants ended with fiscal year 1990, and the 1987 amendments replaced them with capitalization grants to State revolving funds, which lend at or below market rates against a State match of 20% (33 U.S.C. §§ 1381, 1382(b)(2), 1383(d)(1); U.S. Environmental Protection Agency, 1988; Ramseur, 2025, p. 4).[11] The 1981 amendments had already reduced the federal share to 55% for grants awarded after September 30, 1984, and restricted eligible expenditures (33 U.S.C. § 1282(a)(1); Ramseur, 2016, pp. 12–13). This paper studies the construction grants rather than the loans that followed. A grant directly funded the plant and was never repaid (33 U.S.C. § 1284(b)(1)(A)), which makes a plant's commissioning a direct federal infrastructure investment. The Grants Infor-

[9] A State's list carried a one-year fundable portion, capped at the funds available less reserves, and a five-year planning portion (40 CFR 35.915(b), (c)). From 1984 the criteria became the impairment of classified water uses by existing municipal discharges and the extent of use restoration or public health improvement a project would bring (40 CFR 35.2015(b)(2)).

[10] By the 1988 deadline 86% of cities complied, and the rest were placed on administrative or court-ordered schedules (Copeland, 2016, p. 5, citing EPA). In the General Accounting Office's 1980 random sample of 242 plants in ten States, 87% had violated their permits during a twelve-month period, and enforcement "varied from none to minimal" (U.S. General Accounting Office, 1980, pp. i–ii, 28).

[11] Loans were repaid from within a year of completion and amortized within 20 years under the 1990 rule (40 CFR 35.3120(a)). Each fund's first loans were reserved by law for the plants on the State's National Municipal Policy list, and leveraging the capitalization grant as security for State bonds was permitted from the start (U.S. Environmental Protection Agency, 1988, pp. 4, 6–7).

mation and Control System (GICS), EPA's grant record, dates each project's steps and its completion.

Once commissioned, a treatment plant modifies the receiving water through several mechanisms, each of which can influence bird populations downstream. Secondary treatment reduces oxygen-demanding organic matter and suspended solids, while advanced treatment removes nutrients. Appendix Table B13 describes the technologies, their lifetime costs, and what the literature reports for birds at each. For example, after Tampa Bay's municipal plants implemented advanced treatment in 1980, the total nitrogen load from that source fell by 90% (Greening and Janicki, 2006). Keiser and Shapiro (2019a) construct an index of fishable water from biochemical oxygen demand, dissolved oxygen, fecal coliforms, and total suspended solids. Appendix Table B1 details three of these constituents and the pathways by which each reaches birds downstream.[12] The first mechanism is oxygen. As oxygen levels in heavily enriched water recover, vegetation and invertebrates recolonize (Pringle and Burton, 2017): on the Lea estuary, a sewerage scheme shifted the benthos from dominance by first-order opportunists in 1995 to nearly 100% pollution-sensitive taxa by 1997-98 (Borja et al., 2000). The second mechanism is clarity. Submerged vegetation requires 4 to 29% of incident light, and its survival depends on water clarity (Dennison et al., 1993). The upper Potomac River, nearly devoid of submerged vegetation until 1983, regained a thirteen-species community that year as improvements in sewage treatment, together with unusual weather, raised water clarity (Dennison et al., 1993). In Chesapeake Bay, invertebrates replaced submerged vegetation in the waterfowl's diet (Perry and Deller, 1996). Across three Danish lakes, water clarity, more consistently than nitrogen or phosphorus, drove the changes in the breeding waterbird community, whose abundance rose up to fivefold in the two lakes restored to clear water (Fox et al., 2020). The third mechanism is enrichment. Sewage can directly feed birds such as gulls, but "mostly it is the algal, plant, invertebrate and fish communities sustained by the outfalls that provide food for a variety of birds" (Pringle and Burton, 2017). Nutrient removal can restore submerged vegetation and the waterbirds that graze it. On Lake Veluwemeer, almost empty of waterbirds in the 1970s and 1980s, phosphorus removal at the local treatment plant and flushing with phosphorus-poor water were followed by the return of stoneworts and of as many as 65,000 waterbirds in November 1998, the grazers tracking their food (Noordhuis et al., 2002).

Beyond the chemistry of the water, treatment plants create physical structures and discharge warm effluent. Below a Swiss municipal plant, the stream averaged 8.3 degrees Celsius

[12]The fourth, fecal coliforms, indexes pathogens, whose evidence on birds is correlative: avian pathogens, *Clostridium botulinum* among them, were more prevalent in the more effluent-affected of three Spanish wetlands, and the authors find "not enough evidence" that such waters attract birds and favor outbreaks (Anza et al., 2014).

over a winter against 4.8 degrees above it, and in January and February insectivorous birds were eight times as dense as upstream, on six times the emergence of chironomid midges (Lepori, 2022). British works with percolating filters served several passerines as feeding sites, and works with percolating filters and tertiary treatment by surface irrigation "are likely to support the most varied bird communities" (Fuller and Glue, 1980).

Conversely, treatment can reduce the food that untreated sewage had supplied. After a primary treatment plant began processing Edinburgh's sewage in 1978, the scaup and goldeneye wintering between Leith and Levenhall fell from peaks of 3,640 and 1,638 in the winter before treatment to 675 and 608 two winters later (Campbell, 1984). The gulls using the tidal River Tyne fell by 37% after the untreated sewage discharged into it was cut by 97% (Raven and Coulson, 2001). On one lagoon of Melbourne's Western Treatment Plant, waterfowl numbers declined after it began receiving treated effluent instead of raw sewage and increased after partially treated sewage was restored, a pattern not observed on the plant's other lagoons (Loyn et al., 2023). Reviewing this evidence, Pringle and Burton (2017) judge that treating sewage can be expected to affect waterbirds adversely in some situations, that the effects are site- and species-specific, and that the associations are "correlative, or otherwise postulated," most apparent within a site and, at the scale of the whole site, hard to separate from other factors.

These pathways operate through the receiving water. Thus, each plant's contribution to water quality is its treated effluent relative to the river's flow where it enters, and the water at a surveyed circle reflects the sum of these ratios over the commissioned plants upstream, which Section 4.1 constructs. The timing of commissioning defines the event estimated in the first half, and the mixed evidence on direction motivates the second, which sets the birds' response against the measured change in water quality (Section 5).

3 Data

To estimate the ecological effects of the construction grants, I assemble a panel that joins the counts of a historically rich volunteer bird survey, the Title II grant records, EPA's periodic censuses of treatment plants, high-resolution river hydrography, and ambient water quality monitoring records. This section describes the data sources, spatial matching protocols, species classifications, and sample restrictions used in the analysis. The extensive margin (Section 4) uses the counts, the survey effort, the commissioning dates, the network link, and the covariates. The intensive margin (Section 5) adds the design flow, the mean annual flow, the dose, water quality, and the plants' effluent and technology.

3.1 The Christmas Bird Count

I measure birds with the National Audubon Society's Christmas Bird Count (CBC) (National Audubon Society, 2025b). The CBC has two advantages for evaluating the Title II grants. Its record runs for seven decades before the 1972 amendments, which allows balanced windows of surveys before and after a plant's commissioning, and its circles are fixed sampling frames, which keep birder effort from moving across space in response to a plant's cleaning. Other monitoring platforms offer neither. The North American Breeding Bird Survey began in 1966, much after CBC had been well-established and six years before the amendments (U.S. Geological Survey, 2026). eBird's submissions, though geographically much richer and resolution finer, begin with its launch in November 2002, twelve years after the last construction grants of fiscal year 1990 (Sullivan et al., 2009).

The CBC began on Christmas Day of 1900, and in each winter since, volunteers guided by a compiler have counted the birds inside a circle 15 miles (24 kilometers) in diameter on one calendar day of the count period, December 14 to January 5 (National Audubon Society, 2024).[13] I call one circle's count in one winter a survey. Audubon calls it a "count." I index each survey by its year, the December in which the count period opens, which is Audubon's "count year" (National Audubon Society, 2025a). A circle "cannot be moved once established as a count circle" and keeps the same center every year (National Audubon Society, 2024).[14] I therefore take the circle as the spatial unit. The circle is fixed in place and surveyed winter after winter. Thus, the same place is observed before and after a plant upstream of it is commissioned. A circle covers about 450 square kilometers, three quarters of the area of the city of Chicago, nearly eight times the land area of Manhattan, 3.7 times that of Boston, and half that of Dallas (U.S. Census Bureau, 2020).[15] Among the thirteen empirical studies I reviewed of how a policy, an economic shock, or new infrastructure affects birds, one uses the CBC. Katovich (2024) finds that shale wells within 5 kilometers of a circle cut the birds counted there by 15% and that wind turbines have no measurable effect.

The surveys record counts by species, from which I construct two main outcomes and three diversity statistics for each survey. Let N_{jct} denote the count of species j in circle

[13] A bird seen in the three days before or after the count day but not on it is recorded as present without a number, since "Numbers of count week birds are NOT recorded" (National Audubon Society, 2024), and such records enter none of my statistics. Where a circle lies is its compiler's choice: "where you draw it is up to you" (Contreras and Hoyer, 2022).

[14] Katovich (2024) prefers the CBC to eBird for "its stable and transparent methodology, which allows adjustment for counting effort and avoids endogenous changes in birders' choice of where to observe and report" and for its numbers by species, since eBird users are liable to "under-report common species and over-report remarkable species."

[15] A circle 15 miles in diameter covers 177 square miles. The Census Bureau's land areas for 2020 are 22.66 square miles for New York County, 48.34 for Boston, and 339.58 for Dallas.

c and year t. Abundance, $N_{ct} = \sum_j N_{jct}$, is the number of birds counted in a group of species, and richness, S_{ct}, is the number of its species with a positive count. As robustness checks, I compute three diversity statistics. The Shannon entropy, $H_{ct} = -\sum_j p_{jct} \ln p_{jct}$, where $p_{jct} = N_{jct}/N_{ct}$ is the share of the counted birds that belong to species j, has units of nats, and higher entropy means greater uncertainty about the species of a bird drawn at random. The effective number of species, $\exp(H_{ct})$, is the Hill number of order one (Hill, 1973), the number of equally abundant species that would give the observed entropy. The effective number of common species, $1/\sum_j p_{jct}^2$, is the inverse Simpson index, the Hill number of order two, which weights the dominant species more heavily.

The species names in the raw data "follow the American Ornithological Union (AOU) nomenclature as it existed in the early 2010s" (National Audubon Society, 2025a). I therefore match every name to the 2023 eBird taxonomy. The match gives each species one name across 125 years of changing nomenclature, and it links the surveys to the BirdLife ranges and the trait databases. A subspecies or a form counts toward its species, while a bird identified only to a pair of species, a hybrid, or a domestic type counts toward none. The group of all species pools 749 species as the CBC names them.[16] A species missing from a survey is entered as a zero, or zero-filled, when its BirdLife resident, non-breeding, or passage range lies within 100 kilometers of the circle's center and it was reported in the circle's Bird Conservation Region (BirdLife International and NatureServe, 2024), and only in a year with recorded effort. A year without a survey therefore stays missing. Figure 1 draws the rule in New York for a resident, the Black-capped Chickadee, and a winter visitor, the Yellow-rumped Warbler, with invented statuses.

Because a survey records only the birds its observers detect, the data's provider "strongly suggest[s] that analysts minimally include predictors of the observation process, namely, the effort expended by participants" (National Audubon Society, 2025a). I use the three measures of effort that each survey reports: party-hours, field observers, and miles traveled. Party-hours are the hours all parties spend in the field from dawn to dusk, excluding nocturnal birding and feeder watching (National Audubon Society, 2024). Field observers are counted apart from nocturnal birders and feeder watchers (National Audubon Society, 2024), and miles traveled sum the distance over every mode.[17] Hours are recorded on at least 96% of the surveys from 1953 onward and on few before. Appendix Figure B2 shows the mean

[16] On the 2023 eBird taxonomy these are 744 species, since five pairs of recorded names, Thayer's Gull and Iceland Gull among them, map to one eBird species. I keep the CBC's names, and a survey recording both names of a pair counts two species, as 524 surveys do.

[17] The miles field of 2021 records 0.00019 miles per party-hour, a fall by a factor of about 65,000 that is gone in 2022. I treat it as a recording error and set that year's miles to missing. Thus, 2021 leaves every fit that conditions on effort. Where a circle was surveyed on more than one date in a year, I sum its birds and its effort across the dates.

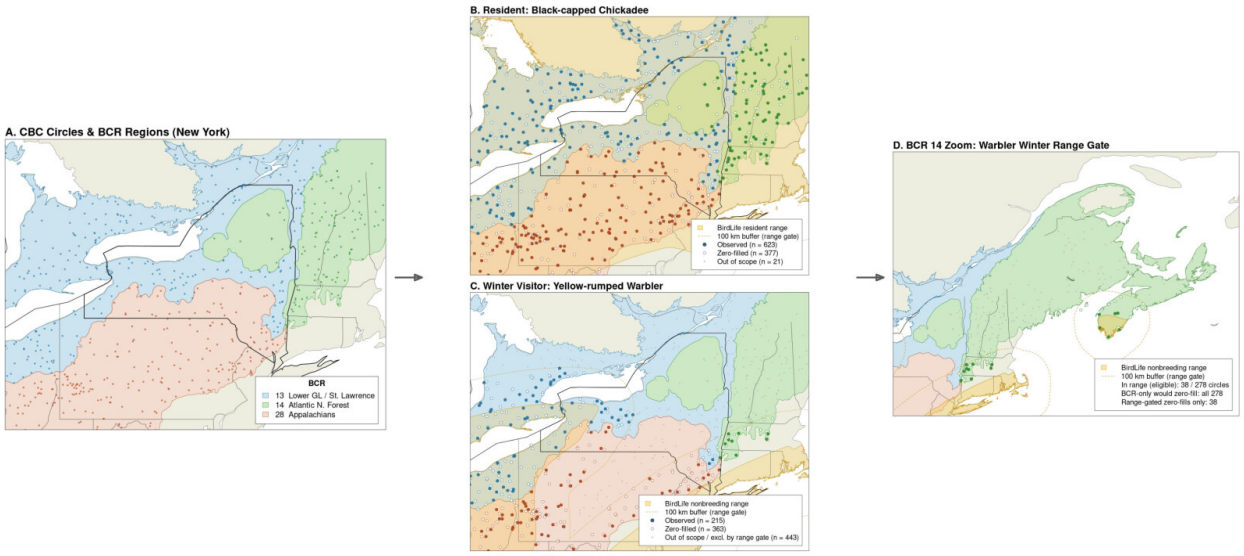


Figure 1: A schematic of the zero-filling in New York's three Bird Conservation Regions. Panel A: the circles and the regions. Panels B and C: a resident, the Black-capped Chickadee, and a winter visitor, the Yellow-rumped Warbler, each with its BirdLife range and the 100 kilometer buffer. Panel D: the warbler's range and buffer within one region. Which circles observed each species, and which were zero-filled, are invented for illustration.

and the median of each measure per circle in each year.

Figure 2 shows the distributions of the two main outcomes over the surveys of the panel defined below, drawn separately for the circles that are ever treated and for the never-treated circles in the comparison, the partition the estimates use. The surveys of the ever-treated circles record a median of 8,759 birds and 63 species, against 3,629 birds and 44 species at the never-treated circles. The estimates are read in units on the treated circles' surveys in the nine years before commissioning, where a survey records 65 species on average and 59 at the median across all birds, and 18 on average and 13 at the median among the water-linked species defined in Section 3.2.

3.2 Species Groups and Feeding Guilds

These statistics are computed for all birds and for categories of species drawn from the ornithological literature, each a mapping of the CBC's species to the classes of a published classification or review. The first are the four groups of a continental assessment of North American birds, whose data assign each species to waterbirds, waterfowl, shorebirds, or landbirds (Rosenberg et al., 2019). Among the species the CBC reports, 77 are waterbirds, 41 waterfowl, 44 shorebirds, and 351 landbirds. Of the four, landbirds hold the fewest water-

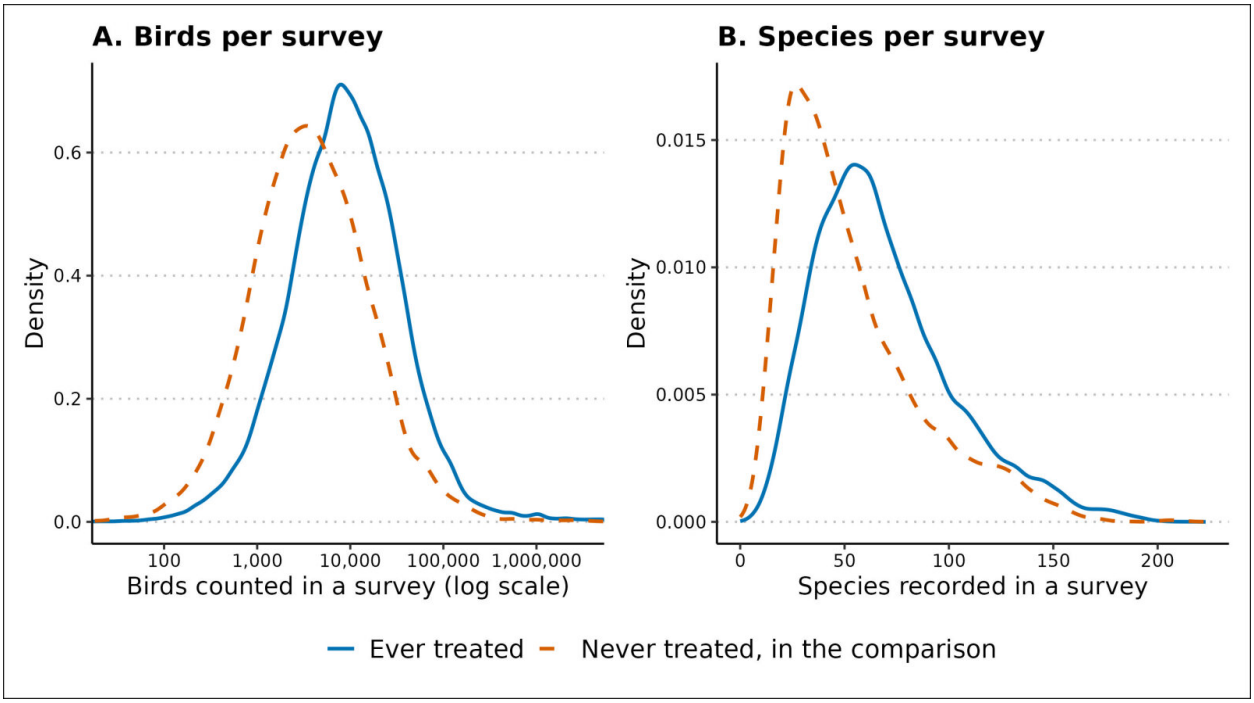


Figure 2: The birds counted (panel A, on a logarithmic axis) and the species recorded (panel B) per survey, over the 75,485 surveys of the 2,794 ever-treated circles and the 28,215 surveys of the 1,386 never-treated circles in the comparison, 1919 to 2024.

linked species, 20 of 350.[18] The second are eight feeding guilds, each a set of species that take one food governed by water quality in the same way. Each guild is drawn from the species that meet one of two screens for water linkage. A species meets the first if it forages on or in the water, its below-surface and around-surface strata in EltonTraits (Wilman et al., 2014) summing to at least 20% of its foraging (281 species). It meets the second if it carries a wetland, riverine, coastal, or marine habitat in the AVONET trait database (Tobias et al., 2022) or a wetland, coastal, or ocean habitat outside the breeding season in the Partners in Flight database (Partners in Flight, 2024), and forages off the surface, its aerial, ground, and understory strata summing to at least 30% (47 species). I assign the guilds from the foraging guilds of De Graaf et al. (1985) and the literature on how birds respond to water quality. The guilds are the visual fish-eaters (68 species), the marsh invertebrate-eaters (94), the submerged-plant grazers (33), the bottom-sediment feeders (21), the crustacean-eaters (11), the scavenging gulls (17), the scavenging crow (1), and the aerial insect-eaters (3).

[18] Of the landbirds, 92 species carry a water linkage on the four trait flags the guild assignment uses, and 22 forage on or in water or eat fish, among them the Osprey and the Belted Kingfisher. The definition depends on where a species forages and which habitats it uses, not on where its prey grew. A landbird that takes insects emerging from the water is therefore not counted as water-linked. Below a Swiss sewage treatment plant in January and February, insectivorous birds were 8 times denser than upstream, on 6 times the emergence of chironomid midges (Lepori, 2022).

Their foods are fish caught by sight, the invertebrates of marshes and shores, submerged water plants, the invertebrates of the bottom sediment, crustaceans, the food the gulls and the crow scavenge, and insects caught in the air. Pooled, the 248 species of the eight guilds form the water-linked species.

3.3 Treatment Plants

The plants come from the grant record EPA kept in its Grants Information and Control System (GICS), into which "information concerning the award of each construction grant is entered" (U.S. Environmental Protection Agency, 1980, IV-14). GICS codes each grant by the phase of work it funds. Step 1 grants fund facility plans and studies, Step 2 grants fund construction plans and specifications, and Step 3 grants fund the building of a treatment plant. Step 4 grants fund design and construction together, Step 7 grants fund design and building under one award, and Step 9 grants fund the design and construction of facilities for training operators (U.S. Environmental Protection Agency, 2017, pp. 6–7). Because physical completion, rather than planning or design, determines when a plant enters service, I count a plant as commissioned once it holds a completed, funded construction grant of Step 3, 4, 7, or 9. I date its commissioning by the earliest such completion. The completion date records "the date when all but minor components of a project have been built, all equipment is operational and the project is capable of functioning as designed" (40 CFR 35.2005(b)(9); Office of the Federal Register, 1999).[19] A plant's earliest award precedes its earliest completion by a median of three years, with an interquartile range of two to five. I therefore date treatment at commissioning rather than at award. Appendix Table B14 gives the years from award to commissioning, the design flows, and the grants by period of commissioning. The grant record holds 8,778 such construction-grant plants, 8,769 of them with a completion date. Figure 3 shows the years from a plant's first construction grant to its commissioning and the years in which the circles' first upstream plants were commissioned.

3.4 The River Network

I link each plant to the circles below it on NHDPlus Version 2.1, the national network of river reaches, which carries a modeled mean annual flow on every reach (McKay et al., 2012). The river is followed 25 miles downstream of the reach each plant discharges into, and 50 miles in a robustness check. A circle is linked to the plant when that stretch of river crosses it, and

[19] In the grant record, the completion date's status code marks the final inspection on 97% of the plants that set a commissioning year, and on the other 244 the date behaves as a schedule. A plant's earliest year of initiation of operation equals its earliest completion year on 56% of plants and lies within a year of it on 81%.

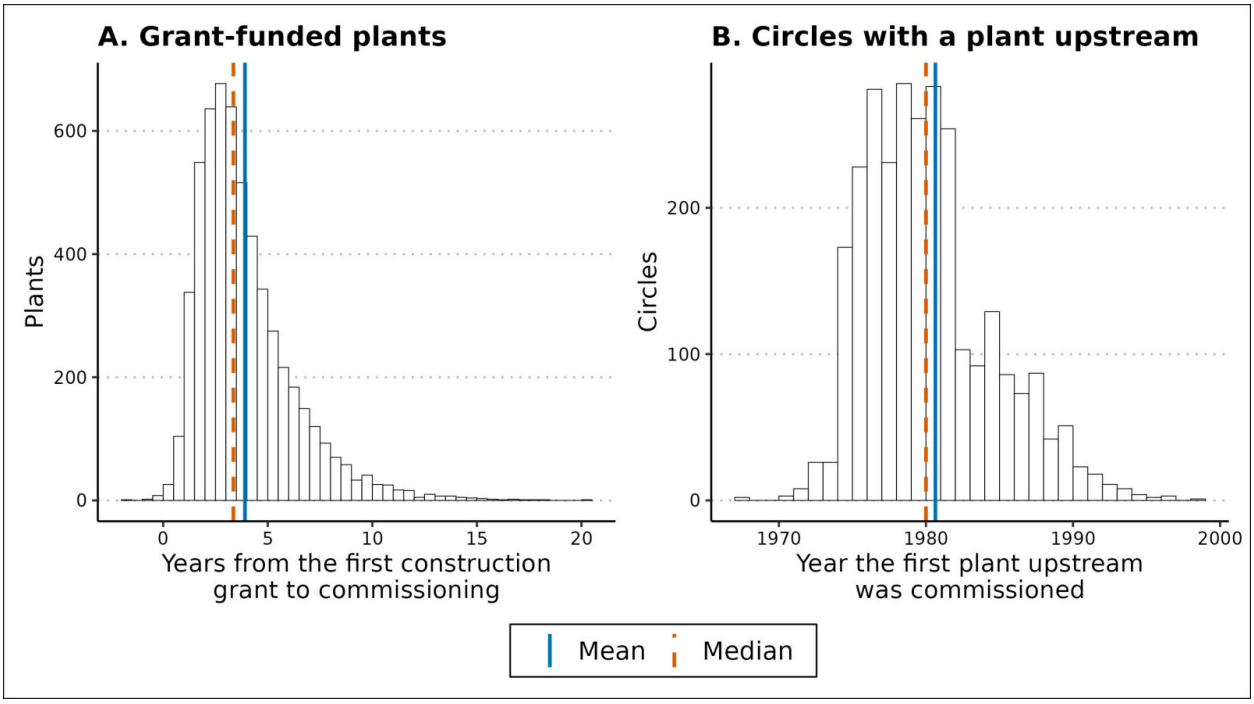


Figure 3: The grant record. Panel A: the 5,641 grant-funded plants within 25 river miles upstream of a surveyed circle, by the years from their first construction grant to their commissioning, with a median of 3.4. Panel B: the 2,794 ever-treated circles, by the year their first upstream plant was commissioned, with a median of 1980. Solid lines mark the mean and dashed lines the median.

the circle's outlet is the most downstream reach of the stretch inside the circle. The band is my choice, long enough to reach the circles that receive a plant's effluent and short enough to keep out effects that operate at the scale of a whole watershed.[20] Of 68,864 geocoded locations of treatment plants, grant-funded or not, 36,402 are linked to at least one circle, in 113,751 plant-circle pairs over 4,378 circles, and 15,214 of these pairs join a grant plant's design flow to a mean annual flow at the outlet. Two quality controls follow. A pair is dropped when its outlet carries less than 0.001 million gallons a day (MGD) of modeled flow per square kilometer of drainage, about 1.4 millimeters of runoff a year, drier than any real basin. The ratio of design flow to mean annual flow can exceed one where a plant's effluent outweighs a small headwater stream. I therefore cap it at its 99th percentile, 4.44 at 25 river miles. Leaving out the plants with no surface discharge and those that send their effluent to another plant removes 4.4% of these ratios summed over grant plants, and leaves 221 circles with no contributing grant plant. Of the 2,794 circles in my sample with a commissioned grant plant within 25 river miles upstream, 72.0% lie below two or more, with a median of three.

3.5 The Dose

The intensive margin needs two further quantities, the design flow of a plant and the mean annual flow of the river at the circle's outlet. Each plant's design flow comes from EPA's Clean Watersheds Needs Survey, the periodic census of treatment plants from which the States' priority lists were built (33 U.S.C. § 1375(b); U.S. Code, 2018; U.S. Environmental Protection Agency, 1985, pp. 1, 5). Its fifteen waves, from 1973 to 2022, record each plant's design flow, the hydraulic capacity it is permitted to treat in MGD, and whether it discharges its effluent to surface water. I read the design flow from the wave nearest the plant's completion year, the later wave on a tie. The 1973, 1976, and 1982 waves entered an unreported design flow as zero. I therefore read a zero in those three waves as missing. Plants with no surface discharge, which apply their effluent to land or reuse it, and plants that send their effluent to another plant still count as commissioned but contribute no design flow.[21] Among the commissioned grant plants upstream of the circles in my sample, the median design flow is 0.5 MGD, with an interquartile range of 0.13 to 2. This is 0.022

[20] Fan (2025) draws the main sample from cells within 25 kilometers of a river reach downstream of a dam, a distance from the river rather than along it.

[21] Design flow is the capacity a plant is permitted to treat, not the effluent it discharges in a given winter. The ideal measure would be each plant's effluent, in volume and in quality, before and after its commissioning. The Needs Survey reports effluent quality only in some waves, with biochemical oxygen demand on both sides of the commissioning year for 44% of commissioned plants, and to my knowledge no record reports the volume each plant discharged over those years.

cubic meters a second, about 0.6% of the Chicago River's 3.9 cubic meters a second that Fan (2025) gives for reference. The mean annual flow is read at the outlet, where the stretch of river below the plant leaves the circle.

Each plant's design flow, as a fraction of the mean annual flow at the outlet where its effluent enters a circle, is that plant's flow share into the circle. The flow share is fixed by the geometry of the river network and carries no date. A plant's flow share is thus its treated effluent relative to the river's flow where it enters. A circle's effluent treated flow share in a year is the sum of the flow shares of its upstream plants commissioned by that year, zero before the first of them, and rising as each further plant comes online. This sum is the dose, as the literature on continuous treatments calls such a measure (Callaway et al., 2025). A dose of one means that the design flows of a circle's commissioned upstream plants sum to the river's mean annual flow at its outlet. For a circle below a single plant the dose equals that plant's flow share, and for a circle below several plants, as most treated circles are, it does not. I therefore keep the two words apart: flow share for a plant's contribution and dose for a circle's total. The construction follows Fan (2025), whose flow share is the share of the water reaching a cell that passed through a dam upstream, one just below the dam and falling at each confluence with an undammed tributary. Fan (2025) counts a dam only when no other dam lies between it and the cell, since water passing the lower dam already passed the upper one. The dose has no such exclusion, as each plant discharges effluent from its own service area rather than water that passed through another plant. Figure 4 draws the construction with invented round numbers as an example. A plant of design flow 20 on a river of mean annual flow 100 gives the circle just below it a dose of 0.20. Once a tributary of mean annual flow 100 joins, the next circle receives 20 of 200, a dose of 0.10. Below the join with a second river of mean annual flow 200 and its own plant of design flow 60, the circle at mean annual flow 400 has a dose of $(20 + 60)/400 = 0.20$. Doses thus dilute at joins and sum across plants.

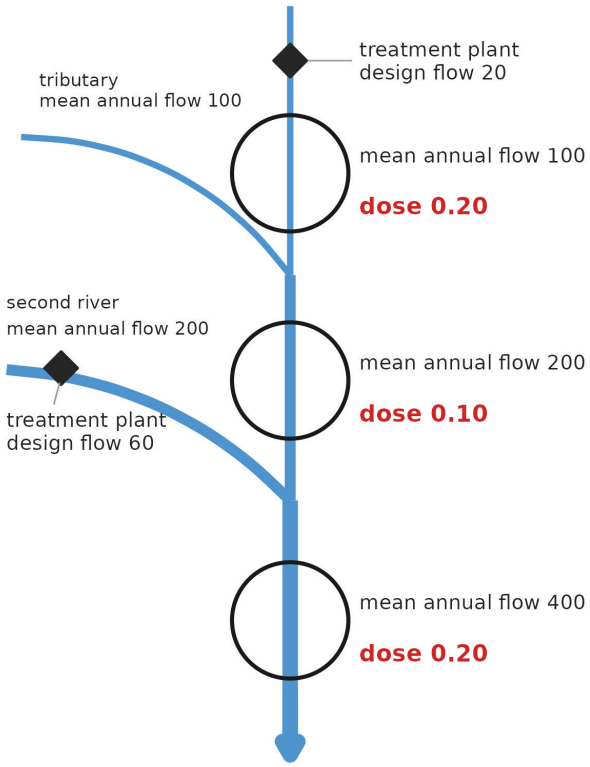


Figure 4: A schematic of the dose, drawn with invented round numbers. Design flows and mean annual flows are in the same invented units, and a dose is their ratio, summed over the commissioned plants upstream of a circle.

The dose rises in steps as later plants come online, and I use it in two forms. The dose at first commissioning is a circle's dose in the year of its first commissioning, held fixed thereafter, and the dose in force is its dose in each year, which rises again with each later plant. Beside the dose, the number of commissioned upstream plants within 25 river miles, hereafter the plant count, is the dose with every flow share set to one. The binary indicator and the plant count include every commissioned grant plant upstream. The dose includes only those of them that have a measured design flow and send their effluent to surface water. A plant that applies its effluent to land, reuses it, or sends it to another plant is commissioned and counted, but it adds nothing to the dose.

Figure 5 plots the dose of each treated circle twice, in the year of its first commissioning and in 2024. Before the sample restrictions of Section 3.7, 3,649 circles have a commissioned grant plant within 25 river miles upstream. Of these, 3,390 carry a positive dose in 2024, and 259 carry a dose of zero, 221 because no plant upstream discharges to surface water and 38 because none has a measured design flow. Over all 3,649 treated circles, zeros included, the

median dose in 2024 is 0.011: the design flows of a circle's commissioned plants sum to 1.1% of the river's mean annual flow at its outlet. The 90th percentile is 0.71 and the 99th 6.64, and the positive doses span eight orders of magnitude. In Figure 5, the median dose among the circles with a positive dose is 0.0030 at first commissioning and 0.015 in 2024. A dose can exceed one where effluent outweighs a small stream, and 8.7% of treated circles carry such a dose in 2024. The cap of Section 3.4 limits each plant's flow share to 4.44, its 99th percentile, and does not limit the dose. Thus, a circle below several plants can carry a dose above 4.44, up to the maximum of 15.0. The flow share itself depends on where the plant sits. Over the plant-circle pairs at 25 river miles, its median is 0.0037 for a plant on one of the circle's own reaches and 0.0007 for a plant outside the circle. The first commissionings came within a short window: among the circles treated by 2024, the circle-years before a circle's first commissioned plant number 31,186 in the 1970s, 6,661 in the 1980s, 220 in the 1990s, and none after. Appendix Figure B7 plots the number of circles receiving a positive dose in each year.

Figure 6 draws three configurations of a plant's downstream and upstream circles with invented round numbers, each a case the dose must handle. In panel A a second plant upstream doses both circles: the upstream circle receives 10 of 100, a dose of 0.10, and the downstream circle sums the two plants, $(10 + 40)/200 = 0.25$. In panel B the upstream circle lies above the only plant and has a dose of zero, while the downstream circle's is $30/200 = 0.15$. In panel C a distant plant doses only the upstream circle, $30/100 = 0.30$, because its water leaves the 25 river miles before the downstream circle, whose own plant gives it $40/200 = 0.20$. A dose thus sums over a circle's plants, is zero above a plant, and ends where the plant's water leaves the 25 river miles.

3.6 Water Quality Portal Monitoring and Additional Covariates

Ambient water quality records are drawn from the Water Quality Portal (WQP), which compiles the monitoring observations of the U.S. Geological Survey and of EPA's modern water-quality database.[22] I track eight analytes in three classes. The regulated constituents are five-day biochemical oxygen demand, total suspended solids, fecal coliform, and dissolved oxygen saturation. Total phosphorus is partly regulated, and the nitrogen constituents are ammonia, total nitrogen, and nitrate plus nitrite. Lead is the one falsification analyte: a response of lead to commissioning would point to a shared trend rather than to treatment.

Water quality stations within a circle's 24-kilometer disk are mapped to the circle, and

[22]The portal does not serve EPA's legacy holdings. Thus, only 0.66% of EPA's records and 6.46% of the Geological Survey's in my pull predate 1972, which thins the baselines of the earliest cohorts only. Keiser and Shapiro (2019a) obtain the legacy records separately.

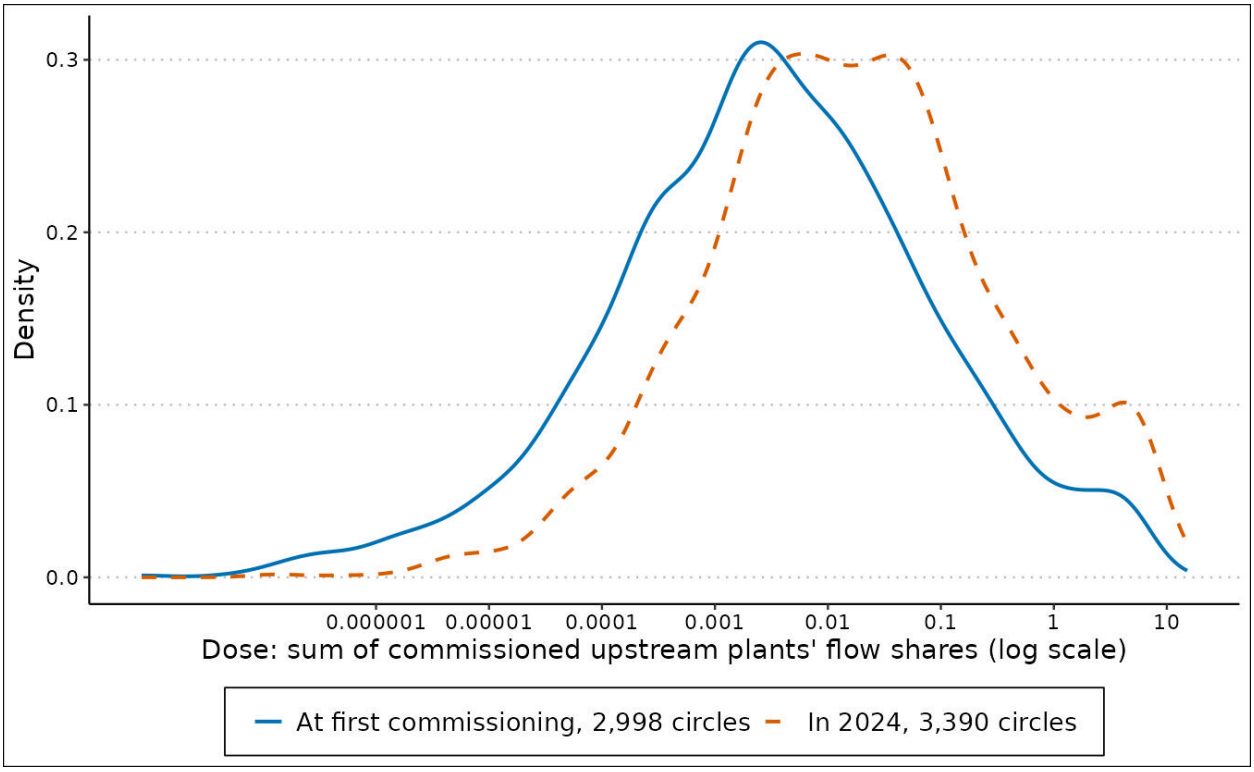


Figure 5: The dose of the treated circles at 25 river miles, on a logarithmic axis, in the year of each circle's first commissioning (2,998 circles with a positive dose) and in 2024 (3,390 circles). A dose of one means that the design flows of a circle's commissioned upstream plants sum to the river's mean annual flow at its outlet. The 392 circles with no positive dose at their first commissioning are counted in neither curve.

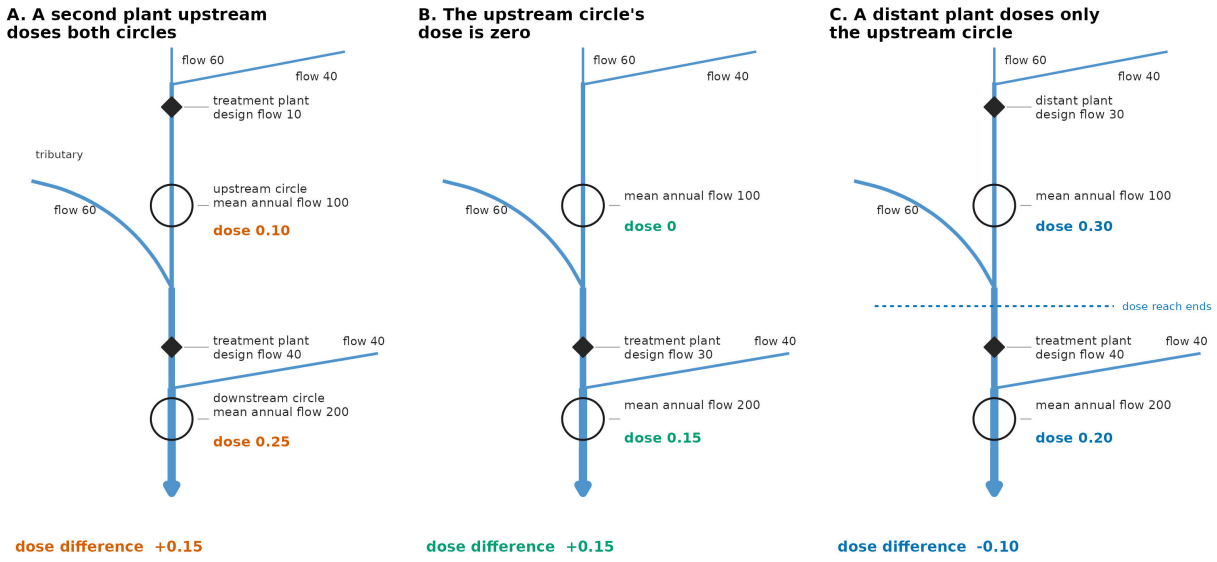


Figure 6: Three configurations of a plant's downstream and upstream circles, drawn with invented round numbers. Diamonds are treatment plants and open circles are survey circles. Black labels are given, the design flows of plants and the mean annual flows of rivers, and labels in a panel's color are derived, the doses of its two circles and their difference, downstream minus upstream. In panel C the dashed line marks where the distant plant's 25 river miles end, so its effluent counts in the upstream circle's dose and not in the downstream circle's.

annual analyte observations are the mean reading during the growing season, May to September, preceding the winter bird survey. Readings are capped at the 99th percentile across circle-years, and fecal coliform is transformed logarithmically, given its extreme skewness. To report standardized effects in physical concentration units, I compute within-circle baseline standard deviations before commissioning: 4.5 milligrams per liter (mg/L) for biochemical oxygen demand and 83.8 mg/L for total suspended solids. Following Keiser and Shapiro (2019a), I measure a circle's baseline water quality over every station in its eight-digit hydrologic unit, which raises the share of circles with a baseline from 65% to 82%, and a circle's readings after commissioning are an outcome of commissioning and never enter a bird regression as a control.

The Needs Survey also characterizes the plants themselves. For the comparisons of Section 5, I assign each treated circle the flow-share-weighted mean effluent concentration of its commissioned plants before their grants and split the circles at the median within each commissioning cohort bin into a cleaner and a dirtier half. The split covers 2,392 of the 2,567 treated circles with a positive dose, with a median effluent oxygen demand of 15.5 mg/L in the cleaner half against 39 in the dirtier. The split rests on the plants' discharge rather than on the circle's own readings, which would build reversion to the mean into every later comparison. A plant's effluent cut is the fall in its effluent concentration from the latest wave at or before commissioning to the first wave after it, and 2,863 of the 3,390 circles with a positive dose carry a flow-share-weighted mean cut. The survey's 1984 to 1992 waves place each plant in a technology class: of the grant plants upstream of a circle, 60.8% are conventional, 29.8% use ponds, lagoons, or wetlands, 7.4% ponds with land treatment, and 2.0% land treatment alone.

The main specification controls for three further dimensions of time-varying and baseline environment. Survey measurement controls are the logarithms of party-hours, field observers, and miles traveled. Winter climate controls are the monthly mean air temperature at 2 meters and total precipitation of the ERA5 reanalysis, averaged across December and January over the circle (Hersbach et al., 2023). Baseline attribute-by-year trends interact four pre-grant attributes with year fixed effects. Housing density is census tract housing units per square mile of habitable land at nine census years from 1940 to 2019 (Markley et al., 2022), and cropland share is the county share of harvested cropland at sixteen agricultural censuses from 1940 to 2012 (Haines et al., 2014). The other two are the summed flow shares and number of the upstream works the grants did not fund, and the grant plants' treatment level before their grants. Each baseline is read at the last census before the circle's first commissioning, or before 1972 for a circle never treated, as Fan (2025) uses values for 2000 "so as to capture pre-dam baselines." The works the grants did not fund are dated by their first

crossing to secondary treatment between two Needs Survey waves and form 4,942 plant-circle pairs on 2,159 works at 25 river miles.

3.7 Sample Restrictions, Identification, and Descriptives

I construct the primary estimation panel from the 6,042 mapped CBC circles. Restricting to the 5,135 circles located within the 30 Bird Conservation Regions covered by the zero-filling, and requiring at least one survey with recorded participant effort, yields 4,224 circles and 105,057 survey observations spanning 1919 to 2024. At 25 river miles, 2,794 of these circles are ever treated and 1,430 never treated, and the 44 never-treated circles that hold a grant plant on one of their own reaches leave the comparison, since the plant's water misses their outlet.

Three further requirements define the estimation sample. The three effort measures remove 2,472 surveys, 2,223 of them in 2021. ERA5 climate data from 1940 onward remove 84 circles, and baseline land-use data measured on at least half of the circle remove 724, 634 of them in Canada or Mexico. The 397 circles then left with a single survey also leave, because a circle effect fits one survey exactly. The final estimation sample is 85,934 survey observations across 2,975 circles (2,372 ever treated and 603 never treated) spanning 1940 to 2024. Within it, the sample differs across species groups and statistics: a group's panel holds only the surveys in which one of its species was recorded or zero-filled, the Poisson fits drop every circle whose count of the group is zero in every surveyed year, and the diversity statistics are undefined where a survey records no bird of the group. Each table gives the surveys behind each column.

Figure 7 illustrates the spatial distribution and survey coverage of the treated and never-treated circles across the contiguous United States. Identification is separated across the two margins. The extensive margin (Section 4) is identified by the staggered dates of first commissioning of a grant-funded plant upstream of a circle, relative to the not-yet-treated and never-treated circles. The intensive margin (Section 5) is identified by continuous variation in the dose, how much of a river the commissioned plants' effluent makes up, on this panel with the dose attached. Of the 2,794 ever-treated circles, 2,567 carry a positive dose, the other 227 having their plants screened out or their design flows unmeasured, and the water-quality regressions run on the circle-years with a reading.

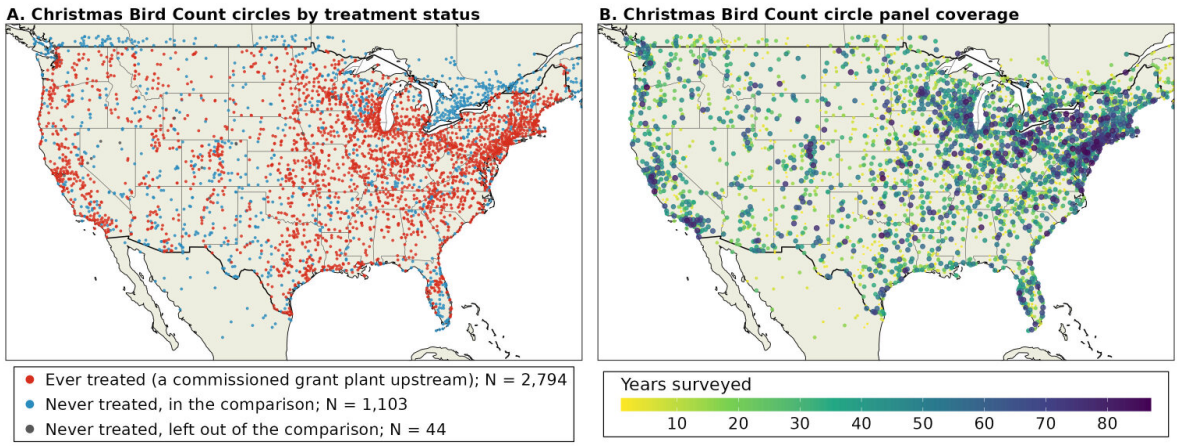


Figure 7: The 4,224 circles of the panel at 25 river miles. Panel A: circles ever treated, never treated and in the comparison, and never treated but left out of the comparison. Panel B: the number of years each circle was surveyed, 1919 to 2024. The map shows circle centers within the contiguous United States, and 283 never-treated circles lie outside it.

4 Downstream Effects of Grant-Funded Treatment Plants on Birds

This section estimates the extensive margin of the construction grants: whether the commissioning of a grant-funded treatment plant upstream of a circle changed the birds counted downstream of it. Section 4.1 sets out the event study and the tests of its assumptions, and Section 4.2 reports the estimates and then the tests.

4.1 Empirical Method

4.1.1 Estimating the Effect of Grant-Funded Commissioning

To identify the level effect of grant-funded commissioning, I compare the circles downstream of newly commissioned plants with the circles that have no commissioned plant upstream. Treated and untreated circles differ before any grant is made, for two reasons. First, governments typically regulate a plant's effluent to remain within dilution limits, and its outfall therefore discharges into a river large enough to dilute it (Ehalt Macedo et al., 2022). Second, grants went first to the projects ranked highest under the rules of Section 2, by the severity of the pollution and the population affected. The circles downstream of commissioned plants therefore lie on larger rivers and on dirtier water. In the estimation sample, the standardized difference between the treated and the never-treated circles, the difference in means over the pooled standard deviation, is 0.69 for the largest river's mean annual

flow and 0.44 for its drainage area (Appendix Table B7). Before commissioning, the treated circles' watersheds also carried more fecal coliform, a standardized difference of 0.40 and a median of 1,050 against 443 most probable number per 100 milliliters, and more oxygen demand, 0.26 and a median of 2.8 against 2.2 milligrams per liter. Their dissolved oxygen was lower, a standardized difference of 0.21, at 87% against 92% saturation, and their suspended solids the same (Appendix Table B8). A comparison across circles would credit the grant with this siting and ranking, and a comparison of each circle before and after commissioning would credit it with every change common to those years. Circle effects absorb the levels that siting and ranking set, year effects absorb the common shocks, and the event study compares the two sets of circles over event time. Moreover, in the sample of Section 3.7, circles with no grant-funded plant upstream differ systematically with their distance from the nearest treated circle (Section 4.2.2). A plant's downstream and upstream circles share its region and ranking, and only the downstream circle receives its effluent. Thus, Section 5 compares each plant's downstream circle with its own upstream circle.[23]

I construct the treatment from the plants upstream of a circle c on the river network. A plant commissions in the year its earliest grant-funded construction project is completed, and G_c, the year of the circle's first commissioning, is the first year with a commissioned plant upstream. Event time is $k = t - G_c$. This half uses only whether and when a circle's first plant was commissioned. The dose, constructed in Section 3.5 from the same plants, enters in Section 5.

Abundance, richness, and the three diversity statistics of Section 3.1 enter on their own scales. Abundance N_{ct} and richness S_{ct} enter through their conditional mean, the exponential index of a fixed-effects Poisson model, and never as a logged count (Chen and Roth, 2024). The logarithm of the effective number of species is the entropy H_{ct}. However, parallel trends in the two are different assumptions, a common change in effective species against a common proportional change, and I report both. The level effect is therefore proportional for abundance and richness, in species for the effective number of species, and in nats for the entropy. The tables report the three diversity statistics standardized within each group, in standard deviations, and convert the main estimates back to these units.

I estimate the level effect under two assumptions, each stated on the statistic's own scale. The first is parallel trends. For abundance and richness, it is the index version of parallel trends of Wooldridge (2023): without commissioning, the expected count in treated and untreated circles would have grown by a common factor between any two years, which is

[23] Fan (2025) sets aside three "naive" comparisons for dams, among them places upstream against places downstream of a dam. The failure of balance and parallel trends the paper reports concerns a comparison with undammed rivers, and the upstream and downstream comparison is listed without a test.

parallel trends in the logarithm of the expected count. For the effective number of species and the entropy, it is parallel trends in levels, as Callaway and Sant'Anna (2021) state it: without commissioning, the mean of the statistic would have changed by the same amount in treated and untreated circles from one year to the next. Under this assumption, I compare each treated circle with every circle still untreated in the same year, never treated or not yet treated (Callaway and Sant'Anna, 2021; Callaway et al., 2025), a design that Katovich (2024) applies to shale wells on the CBC's circles against never-treated circles, with circle and year effects and the number of observers as effort.[24] The never-treated circles are those with no grant-funded plant upstream. A survey records the birds its observers detect, and detection rises with effort. Thus, survey effort in hours, observers, and miles traveled enters as a measurement covariate, and winter temperature and precipitation enter in levels. Time-varying covariates "should not be influenced by the policy intervention" (Wooldridge, 2023). Thus, baseline housing density and cropland share enter as attribute-by-year trends rather than as current values,[25] and so do the plants upstream that the grants did not fund, the sum of their flow shares and their number in the year before the circle's first commissioning, and the grant plants' treatment level before their grants.

The second is no anticipation and strict exogeneity of timing: the birds do not anticipate commissioning, and no plant upstream of a circle, first or later, is commissioned in response to the birds counted there (Wooldridge, 1997; Callaway et al., 2025). Under this assumption, a circle's whole history of commissioning can be held fixed, as the fixed-effects estimator of equation (1) requires (Wooldridge, 1999). The timing of a locality's grant was set by its project's place on the State's priority list and by the appropriation cycle, and none of the criteria used to rank projects concerned birds.[26] I measure the level effect against the counterfactual scenario in which the circle's plants receive no grant. That scenario still contains the treatment plants localities financed themselves. Thus, the estimate does not measure the effect of treatment plants regardless of who paid for them.

For the survey of circle c in year t, I estimate the event study of commissioning on

[24] A circle first treated after the year being compared is untreated over the whole comparison window, and its change over the window is an untreated change (Callaway and Sant'Anna, 2021). It contaminates the comparison only if its birds respond to its own coming treatment, and a circle never treated has no coming treatment to respond to.

[25] Keiser and Shapiro (2019a) estimate that a grant added about $7 million to home values within 25 miles of the river it cleaned, against a project cost near $31 million. A current value of housing density could therefore carry part of the effect.

[26] A project ranked low on its State's priority list could wait "10 years or more" for its grant (Congressional Budget Office, 1985, p. 13). States could also add criteria of their own, among them readiness for construction and "a locality's influence at the state level" (Congressional Budget Office, 1985, p. 8).

abundance,

$$\mathrm{E}\Big(N_{ct} \mid \cdot\Big) = \phi_c \exp\left(\lambda_t + \sum_{k \neq -1} \beta_k \mathbf{1}(t - G_c = k) + \mathcal{C}'_{ct}\delta + \mathcal{T}_{ct}\right), \tag{1}$$

where ϕ_c and λ_t are circle and year effects, the mean is conditional on the circle effect and on the regressors the display shows, and the indicator of event time k equals one when year t falls k years from G_c and is zero for a circle with no plant upstream. The vector $\mathcal{C}_{ct}$ holds the logarithms of the three effort measures, party-hours, field observers, and miles traveled.[27] It also holds winter temperature and winter precipitation. The vector $\mathcal{T}_{ct}$ interacts the circle's baselines, its housing density, its cropland share, the plants upstream that the grants did not fund, and the grant plants' prior treatment level, with indicators for every year. In estimation, year -1 is the omitted event time. The main specification then reports every coefficient, year -1 included, relative to the average of the pre-treatment coefficients over years -9 to -1, which are constrained to average to zero, and the additional specifications that I run keep year -1 as theirs.

The window follows Miller (2023). Rather than "blindly choosing the period immediately before the event for the normalization," that paper recommends "a reasonable pre-event window" with the coefficients "constrained to average to zero within the pre-period window." The window shifts "the whole pattern of coefficients up or down" while "retaining the same shape," and because the noise of a single reference year no longer enters every coefficient, the standard errors "can be noticeably smaller." The pre-trend is then judged on "the overall trend in coefficients" rather than year by year. Fixed-effects Poisson quasi-likelihood conditions out the circle effect without assuming a distribution for abundance or restricting its serial dependence (Wooldridge, 1999, 2023). The specification assumes a multiplicative circle effect, the conditional mean assumption of Wooldridge (1999), given the circle's covariates in every year, and a probability of detecting a bird that is unrelated to abundance given the circle's effect and covariates. That is, β_k is the proportional difference in expected abundance between circles k years from their first commissioning and the circles not yet or never treated in that year, net of the circle and year effects and the covariates, measured at the main specification from the average of the nine years before the commissioning. Richness enters the same index, and the three diversity statistics enter a linear model with the same right-hand side, a circle effect α_c, and an error term.

Note that the circle and year effects absorb a plant's presence before its grant-funded construction and the shocks common to every circle. The coefficients thus identify the effect of commissioning through the water a plant sends downstream and through whatever else

[27] Miles traveled are not recorded for 2021. Thus, every specification with effort drops that year.

arrives with commissioning in that year. The event study does not show whether the effect grows with the dose, or how much of it operates through the water quality the plant changed. Section 5 takes up both questions, after the tests of the event study's own assumptions. Appendix A states this result formally and proves it.

4.1.2 Specifications and the Tests of Assumptions

Since commissioning is staggered across years, one static coefficient cannot reveal a divergence between treated and untreated circles before commissioning. Under no anticipation and parallel trends every pre-treatment coefficient of equation (1), β_k for $k < -1$, is zero, and their joint test is the pre-trend test for each statistic (Callaway and Sant'Anna, 2021). I report each estimate under a sequence of specifications: two-way fixed effects alone, the covariates added in turn, the main specification on the full sample with the attribute trends added and the pre-treatment coefficients recentered to average zero (Miller, 2023), and that specification on the treated circles observed at every event time, first with and then without the extreme counts. Standard errors are clustered by circle. When the sample is not drawn by cluster, Abadie et al. (2023) cluster at the level of treatment assignment. Here a commissioning takes effect at the plant and reaches every circle within 25 river miles downstream of it, and the circle is the unit the panel follows over time. A rejected pre-trend test, however, does not identify the assumption that failed. Thus, I group the assumptions into four categories, each with its own test.

First, parallel trends is tested by the pre-treatment coefficients of equation (1), estimated on the full sample and on the balanced sample, whose cohorts are the same at every event time. Second, under no anticipation and strict exogeneity of timing, a circle's residuals before its first commissioning are unrelated to when its plants are commissioned, a restriction the pre-treatment coefficients alone cannot test. Third, for the observation process, an event study of effort on commissioning measures a trend in effort over event time, and a test of whether a survey's residual predicts the next year's effort detects observers responding to the birds. A circle's exit regressed on its earlier residuals, and the gap between the balanced and full samples, detect circles dropped after poor years, and an event study of whether a circle is surveyed shows any change in which circles are surveyed. Fourth, a difference-in-differences against untreated circles subtracts any spillover onto them. Following Butts (2026), I test no interference, the part of the stable unit treatment value assumption that rules out one circle's treatment moving another circle's birds, with the gradient of the outcome in distance from the nearest treated circle among circles with no plant upstream. Figure 8 draws the channel this test targets: birds that relocate between a treated river and a separate stream with no water link to it. When a test rejects, I report the rejection for its category of assumption

and do not change the specification to remove it.

The three diversity statistics also allow a test that detects a change in the community without its direction. All three are unchanged when species exchange their abundances. Thus, a change in any of them detects that commissioning altered the community but cannot identify the species that rose. For example, if commissioning replaced one species with another of equal abundance, every Hill number, the family that includes the effective number of species, would stay where it was. I test the null hypothesis that commissioning leaves each statistic unchanged.[28]

The results open on the estimates. The checks of the observation process, the pre-trend tests, and the falsification tests follow them in Section 4.2.2, the observation process first, since every later test uses the surveys that were taken, with effort in the index.

4.2 Results

4.2.1 The Effect of Commissioning

Commissioning a grant-funded treatment plant upstream raised the number of species recorded on the surveys downstream and left the diversity statistics unchanged. Table 1 reports the estimates for the six groups the paper follows: all birds, the water-linked species pooled, waterbirds, waterfowl, shorebirds, and landbirds.

The estimates are read the same way on every surface of the paper. A coefficient b on the log scale is a change of $100(e^b - 1)$ percent in the expected count. That percent times the count of a survey at the treated circles in the nine years before commissioning gives the effect in species or birds, on the average survey and on the median survey. Because the model applies one percent to every survey, the effect on the average survey is also the average effect over those surveys. The two surveys differ most for birds: with the extreme counts dropped, the average survey recorded 6,021 water-linked birds and the median survey 620. A group's effect in units depends on how many species or birds its surveys record. Thus, percents and units rank the groups differently, and a range across groups gives the focal group's percent with its units, then the other groups' percents. For example, among the water-linked groups whose richness rose, the pooled water-linked species have the smallest percent, about 5%, and the largest rise in species, about one on the average survey. The scavenging gulls' 6.5% is about one more species every 12 surveys. Birds add across surveys. Thus, a total is birds counted, given for a year because the same birds are counted every winter. Species do

[28] On the entropy, parallel trends can hold for the statistic itself, fitted by least squares, or for the logarithm of the mean effective number of species, fitted by Poisson quasi-likelihood. The two differ by a Jensen gap that moves too little over the panel to matter for a headline estimate but is large relative to the change this test detects. Thus, I choose one form in advance and report it.

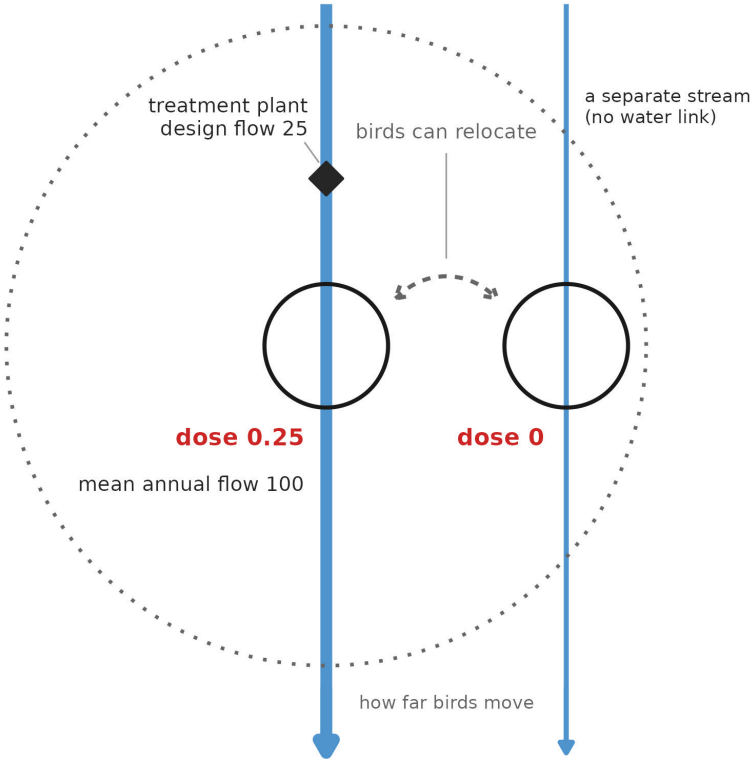


Figure 8: A schematic of birds relocating between a treated river and a separate stream with no water link to it, drawn with invented round numbers. The circle on the treated river lies below a plant of design flow 25 on a river of mean annual flow 100, a dose of 0.25, and the circle on the separate stream has a dose of zero. The dotted ring marks how far the birds move.

not add: a species recorded on two surveys is two entries on surveys' species lists and one species, and the number of distinct species that rose is not identified. A total scaled to the continent's populations treats the surveyed circles as standing for the continent, and it is an upper bound.

For all birds, richness rises 0.015 log points (SE 0.008), about 2%. In the nine years before commissioning the average survey at the treated circles recorded 65 species and the median survey 59. Thus, the rise is about one more species on either, about a seventh of the change a circle's richness shows from one year to the next. Among the water-linked species, richness rises 0.045 log points (SE 0.015), about 5%: about one more species on the average survey, which recorded 18 water-linked species, and one more every two surveys on the median survey, which recorded 13. Among waterbirds and waterfowl it rises 0.054 log points (SE 0.019) each, about 6%. For waterbirds that is one more species every three surveys on the average survey (about 6 species) and every six on the median survey (3). For waterfowl it is one more every two surveys on the average survey (about 8) and every three on the median survey (7). For shorebirds and landbirds the rises lie within sampling error, 0.014 (SE 0.020) and 0.007 (SE 0.007).

Abundance depends on whether the largest counts are kept. For all birds it falls 0.148 log points (SE 0.188) on the full counts and rises 0.114 (SE 0.049) once I drop each year's records above the 97.5th percentile of that year's distinct count values, with its pre-trend test then at the edge of rejection (p-value 0.050). The water-linked species move from −0.181 (SE 0.256) to 0.129 (SE 0.061), and the bottom-sediment feeders from −0.499 (SE 0.099) to 0.203 (SE 0.139) (Appendix Table B2). I therefore report abundance in both forms, the two rows of Figure 9, and I rest no claim on its sign. Abundance on the full counts and the three diversity statistics change in none of the six groups at the 5% level.[29]

With the extreme counts dropped, the water-linked species' abundance rises 0.129 log points (SE 0.061), about 14%: about 830 more birds on the average survey, which recorded 6,021 with the extreme counts dropped, and about 85 on the median survey, which recorded 620. On the full counts it falls 0.181 log points (SE 0.256), about 17%: about 6,800 fewer birds on the average survey of the full counts (41,119) and about 100 fewer on the median survey (632), an estimate whose interval spans zero. Of the 935 species the CBC records, 290 belong to a species group whose richness or abundance rose after commissioning, and with the extreme counts dropped none belongs to a group whose richness or abundance fell, against 21 on the full counts. The 290 are an upper bound on the species that rose, since

[29]For all birds' effective number of species, the average of −0.008 (SE 0.028) lies between the averages from a fit with a separate coefficient for every commissioning cohort and year, 0.029 weighted by treated surveys and −0.028 weighted equally across cohorts, 1.3 and 0.7 standard errors from them.

Table 1: Event-study estimates of the effect of grant-funded commissioning on bird abundance and richness downstream, by species group

	(1) All birds	(2) Water-linked species	(3) Waterbirds	(4) Waterfowl	(5) Shorebirds	(6) Landbirds
Panel A. Abundance and richness, Poisson quasi-likelihood, in log points						
Abundance	−0.15	−0.18	0.19	0.18	−0.10	−0.0059
	(0.19)	(0.26)	(0.12)	(0.17)	(0.12)	(0.24)
Abundance, extreme counts dropped	0.11**	0.13**	0.11	0.15	−0.11	0.080
	(0.049)	(0.061)	(0.092)	(0.11)	(0.12)	(0.056)
Richness	0.015**	0.045***	0.054***	0.054***	0.014	0.0072
	(0.0076)	(0.015)	(0.019)	(0.019)	(0.020)	(0.0068)
Surveys	85,934	85,534	82,090	85,135	70,677	85,934
Panel B. Pre-trend test p-values						
Abundance	0.07	0.06	0.48	0.17	0.20	0.13
Abundance, extreme counts dropped	0.05	0.20	0.34	0.24	0.08	0.03
Richness	0.36	0.27	0.36	0.49	0.45	0.56
Panel C. The effects in percent and in species per survey						
Richness, percent	2	5	6	6	1	1
Abundance, extreme counts dropped, percent	12	14	11	16	−10	8
Species per survey before commissioning, mean	65	18	6	8	3	
median	59	13	3	7	1	
Surveys per additional species, at the mean	1	1	3	2	26	
at the median	1	2	6	3	72	
Circle and year effects	Yes	Yes	Yes	Yes	Yes	Yes
Effort controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
Attribute-by-year trends	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each estimate is the average of the event-study coefficients of equation (1) over years 0 to 14 after commissioning, less their average over years −9 to −1, by Poisson quasi-likelihood in log points. Every regression has circle and year effects and the controls of Section 4.1: the logarithms of party-hours, field observers, and miles traveled; winter temperature and precipitation; and the circle's baseline housing density, cropland share, plants the grants did not fund, and the grant plants' prior treatment level, each interacted with year indicators. Standard errors clustered by circle are in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels. Extreme counts are the records above the 97.5th percentile of their year's distinct count values. Panel B gives the p-value of the Wald test that every pre-treatment coefficient, the pooled endpoint included, is zero in the regression that omits year −1. Panel C converts each estimate b to $100(e^b - 1)$ percent and to the surveys per additional species at the mean and median richness of the treated circles over the nine years before commissioning, not computed for landbirds. For all birds, 2,975 circles, 2,372 of them treated, remain in the estimation. The effective number of species, the number of common species, and the entropy are in Appendix Table B9.

more of a group's species recorded need not mean that each of them rose.

Figure 9 plots the event study for all birds. Richness's post-treatment coefficients rise from about the third year after commissioning and lie at about 0.02 to 0.03 log points from the eighth year on. Abundance's are positive in the first years after commissioning and negative later, each with a wide interval. The interior coefficients rest on 728 to 1,124 treated circles each (Appendix Table B3). Richness's pre-treatment coefficients rise by about 0.001 log points a year (SE 0.0009) toward the year of commissioning. Their pre-trend test does not reject, but it does not bear on parallel trends, because next year's effort follows this year's richness (Section 4.2.2). From year 0 to the endpoint richness rises 0.027 log points (SE 0.010), and the slope of the pre-treatment coefficients carried over those fifteen years accounts for 0.016 of that rise. Net of it, the rise is 0.012 (SE 0.013).[30] The richness estimate thus rests partly on the circles' own pre-trend, and the part of its rise net of that pre-trend, under half, cannot be distinguished from zero.

Appendix Table B2 shows that, among the water-linked species, the rise in richness concentrates in three feeding guilds: the bottom-sediment feeders, by 0.078 log points (SE 0.023), the submerged-plant grazers, by 0.060 (SE 0.022), and the scavenging gulls, by 0.063 (SE 0.026). The rises for the visual fish-eaters and the marsh invertebrate-eaters lie within sampling error. For shorebirds, richness rises 0.014 log points (SE 0.020), and for landbirds 0.007 (SE 0.007). The pre-trend test on richness does not reject in any of these seven categories. However, it bears on parallel trends only for the shorebirds, the scavenging gulls, and the marsh invertebrate-eaters, where next year's effort does not follow this year's richness. The grazers are also the one category whose abundance rises at the 5% level both on the full counts and with the extreme counts dropped, by 0.351 (SE 0.144) and 0.247 (SE 0.124), the direction in which grazing waterbirds returned with their food to Lake Veluwemeer (Noordhuis et al., 2002). The rises in abundance for the crustacean-eaters and the aerial insect-eaters come with pre-trend tests that reject (p-values below 0.001 and 0.002), and I report the rejections.

Compared with the counterfactual scenario, in which no plant upstream of a circle is commissioned with a grant, the surveys recorded about 3% more of the water-linked species from the first commissioning in 1967 to 2024 (95% interval 1 to 5), by that group's post-period mean. The difference between the observed and counterfactual entries on the surveys'

[30] Carrying the slope forward assumes that a failure of parallel trends would continue in a straight line, one model of such a failure among several. It measures how much of the rise a continuing pre-trend would produce and does not test the assumption.

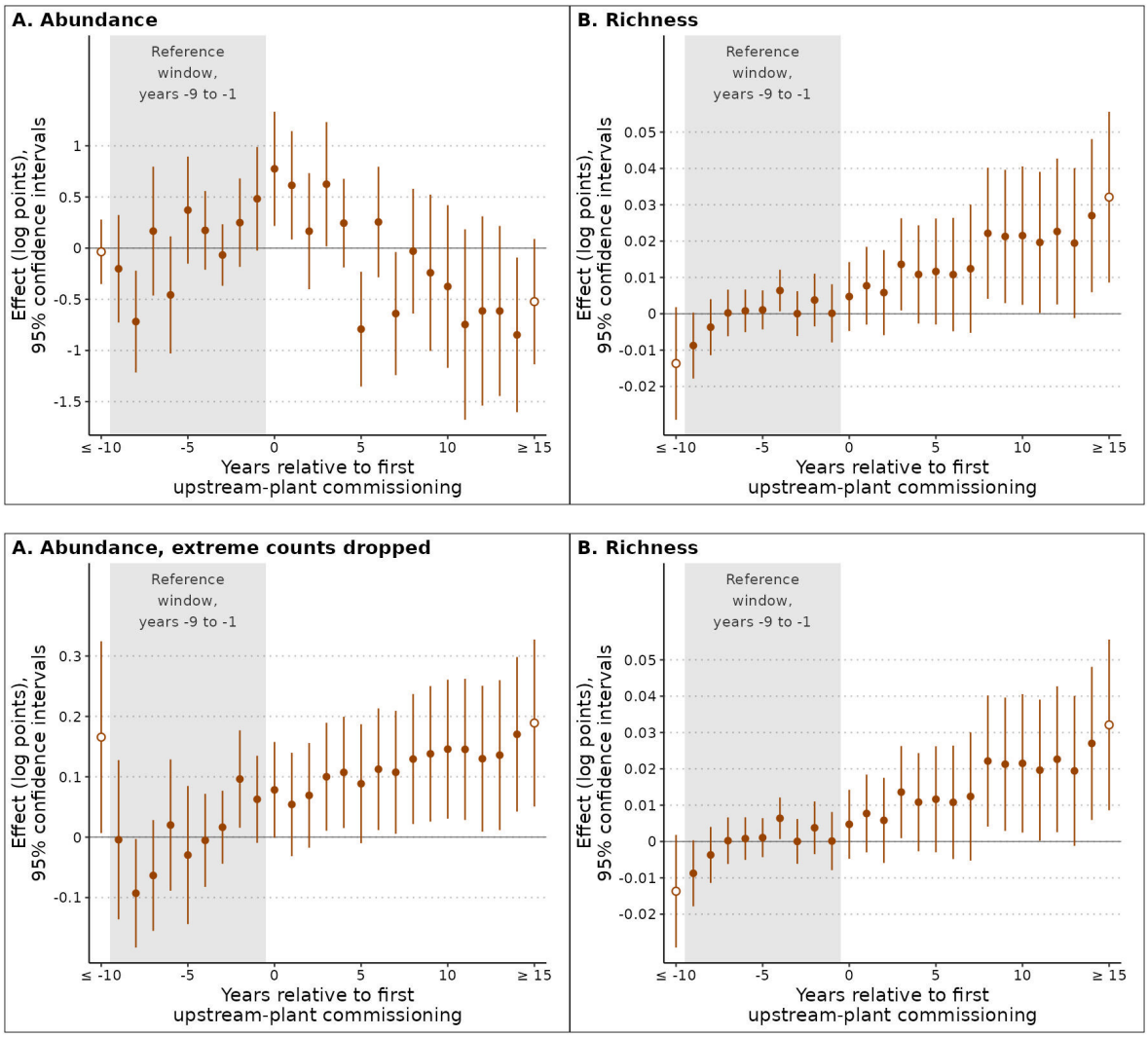


Figure 9: Abundance and richness of all birds by year relative to the first commissioning of a grant-funded plant upstream, on the full counts (upper row) and with the extreme counts dropped (lower row), Christmas Bird Count circles, 1940 to 2024.

Notes: Event study of equation (1) at the main specification, on 85,934 surveys at 2,975 circles, by Poisson quasi-likelihood with circle and year effects and the controls of Section 4.1. In each row, panel A is abundance and panel B richness, both in log points. The upper row uses the full counts. The lower row drops each year's records above the 97.5th percentile of that year's distinct count values, the extreme counts. Richness is drawn in both rows, and the drop moves its coefficients only in the third decimal. Each coefficient is measured against the average of years −9 to −1, the shaded window (Section 4.1). Thus, year −1 carries its own interval. The first and last points pool every year 10 or more before and 15 or more after the first commissioning. They are estimated differently from the interior coefficients and are drawn hollow. Whiskers are 95% confidence intervals from standard errors clustered by circle. The pre-trend test, a Wald test that every pre-treatment coefficient is zero in the regression that omits year −1, has a p-value of 0.07 for abundance on the full counts, 0.05 with the extreme counts dropped, and 0.36 and 0.38 for richness in the upper and lower rows. The window averages around which the coefficients are normalized are −0.48 (SE 0.26) for abundance on the full counts, −0.063 (SE 0.037) with the extreme counts dropped, and −0.0001 (SE 0.0041) and −0.0002 (SE 0.0041) for richness in the upper and lower rows. Appendix Table B3 counts the treated circles at each event time.

species lists is about 870 in an average year (320 to 1,400).[31] This estimate is an upper bound if richness would have kept rising as it did before commissioning. The water-linked species' pre-treatment coefficients for richness slope upward, and net of that pre-trend carried forward fifteen years, their richness rises 0.050 log points (SE 0.026) from year 0 to the endpoint, about half the 0.092 (SE 0.017) without the adjustment. Scaled to the continent as Liang et al. (2020) scale their estimate, with the population estimates of Rosenberg et al. (2019), the water-linked birds that rose with commissioning come to 65.4 million with the extreme counts dropped (95% interval 5.4 to 118.7 million), out of a population of 653.5 million birds of 149 species. On the full counts the figure is −111.5 million (−549.8 to 154.2 million).[32]

The three diversity statistics show no change in any category. In Appendix Table B9 the largest of the 18 estimates in absolute value is 0.052 standard deviations (SE 0.030), for waterbirds' entropy, and none of them differs from zero at the 5% level, nor does any of the diversity estimates for the other categories in Appendix Table B2. The test of Section 4.1 therefore does not reject: the data detect no change in how a survey's birds are shared among species. For all birds, the result is qualified by the rejection of the diversity statistics' pre-trend test (Section 4.2.2). Since the statistics are unchanged when species exchange their abundances, the absence of a change cannot show that no species rose or fell.

Appendix Figure B6 and Figures S1 to S5 of the online supplement show that richness keeps its sign across the nine specifications for all birds, for waterbirds, waterfowl, and shorebirds, and for the water-linked species pooled.[33] Abundance does not. For waterbirds and waterfowl (Supplement Figures S1 and S2), richness lies between 0.025 and 0.059 log points in every specification and is smallest on the balanced sample, where its intervals cross zero. For shorebirds (Supplement Figure S3) it is positive throughout and distinguishable from zero in no specification. For the water-linked species pooled (Supplement Figure S5),

[31] In the counterfactual scenario, each treated survey from the year of its circle's first commissioning on carries its richness times e^{-b}, with b the post-period mean, held at its value for years 0 to 14 after year 14, and every other survey is unchanged. The percent is the difference over the total across all surveys less the difference. A species counts once at each circle that records it in a year. Thus, the 870 are entries of a species on a survey's list, not distinct species.

[32] The share of the group's counted birds that fell in treated circles over the years 2008 to 2017, the years nearest the date the population estimates stand for, times $1 - e^{-b}$ gives the share attributable to commissioning, 10.0% with the extreme counts dropped, which multiplies the population. The calculation treats the surveyed circles as standing for the continent and a rise in birds counted as a rise in the population. The figure is an upper bound if the population outside the surveyed circles rose by less than the surveyed circles did.

[33] The nine specifications are two-way fixed effects alone, with effort, with effort and climate, the main specification, the main specification with the extreme counts dropped, the main specification without the treated circles whose commissioned plants upstream have no surface discharge, the main specification with the never-treated circles that hold a plant on their own reaches kept, the balanced sample, and the balanced sample with the extreme counts dropped.

richness lies between 0.036 and 0.049 log points in the seven specifications on the full sample, each distinguishable from zero, and falls to 0.024 (SE 0.018) on the balanced sample, the one place its interval crosses zero. For landbirds (Supplement Figure S4), richness lies between −0.005 and 0.010 log points and is distinguishable from zero in no specification. For all birds, abundance is −0.30 to −0.37 log points without the housing and cropland trends, where its pre-trend test rejects, −0.148 with them, and positive once the extreme counts are dropped. Dropping the treated circles whose commissioned plants upstream have no surface discharge, or keeping the never-treated circles that Section 3.7 sets aside, moves the main specification's estimate by at most 0.6 and 0.8 of its standard error over the 67 estimates of Appendix Table B4, with medians of 0.18 and 0.21.

4.2.2 The Observation Process and the Falsification Tests

I begin with how the surveys were taken, and first with which circles were surveyed.[34] Appendix Table B5 shows that commissioning does not move circles into or out of the survey: on all 443,080 circle-years, every circle of the panel in every year, surveyed or not, the probability that a circle is surveyed is flat over the nine years before commissioning (p-value 0.22) and changes by −0.8 percentage points (SE 0.8) over the fifteen after.[35] Nor do circles leave the survey after poor years: a circle's residuals from the main fits one to three years before a survey do not predict that the survey is its last (joint p-values 0.09 to 0.61 across the five statistics, on 576 last surveys among 67,232).[36] Appendix Table B6 shows that restricting the treated circles to the 366 surveyed in every year from ten before commissioning to fifteen after leaves the event-study coefficients of richness, the effective numbers of species and of common species, and the entropy indistinguishable from the full sample's (joint p-values 0.37, 0.22, 0.44, and 0.07). The abundance coefficients differ (p-value below 0.001), with a post-period mean 0.21 log points higher (SE 0.18). I report that difference for abundance. One circle's roost holds 42% of all the birds counted, and the difference is not read as a different response among the circles surveyed throughout until a fit without that circle separates its weight from selection.

Among these circles, survey effort, the other part of the observation process, does not step

[34] In 27 of the 29 commissioning cohorts, 1967 to 1999, at least one treated circle was surveyed in the year before its first commissioning. Each of the other two cohorts holds a single circle.

[35] Waterbirds, waterfowl, and shorebirds are present in more treated surveys after commissioning, by 2.7, 1.2, and 2.2 percentage points (SE 0.6 to 0.7). Their diversity statistics, defined only where the group is present, are therefore read conditional on presence, and the change in presence belongs to their richness and abundance.

[36] Where a circle does stop, it stops after its effort fell: a fall of one log point in party-hours or in field observers between consecutive years raises the probability that the second is the circle's last survey by 0.27 and 0.23 percentage points (SE 0.14 and 0.12; joint p-value 0.017), on a base rate of 0.9%.

at commissioning either. Appendix Figure B3 shows that in the year of the first commissioning party-hours change by 0.008 log points (p-value 0.51), field observers by −0.014 (p-value 0.24), and sample coverage, the estimated share of a survey's birds that belong to the species it detected (Chao and Jost, 2012), by less than one ten-thousandth (p-value 0.99). Nor does the way the survey is run change: the nocturnal, feeder, boat, and foot shares of effort, the number of parties, and the number of modes of travel show no step in that year (p-values 0.11 to 0.87). Over longer horizons, however, party-hours and field observers decline in the treated circles. Effort also follows the birds: a survey one Poisson standard deviation above its fitted richness raises the next year's party-hours by 2.9% (SE 0.4) and its miles by 3.8% (SE 0.5), and the test that effort is strictly exogenous rejects for richness (z 5.2) and for no other statistic (z 1.5 and below). I report this rejection for richness and keep effort in the main specification as a control,[37] reading the late coefficients against its pre-trend.[38]

Second, the pre-treatment coefficients: Appendix Table B2 shows that for all birds the pre-trend test on richness does not reject (p-value 0.36), nor does the test on abundance (p-value 0.07).[39] Because next year's effort follows this year's richness, effort is not strictly exogenous, and the residuals behind the richness test need not have mean zero even when the trends are parallel. I therefore report the result of the test without reading it as evidence until an estimator that allows effort to respond to the previous year's count has been run. The pre-trend tests bear on parallel trends only in the categories where neither check of the observation process rejects: whether effort in the next year follows richness in this year, for the richness test, and whether the balanced sample's event-study coefficients equal the full sample's, for the abundance test. The fourteen categories are all birds, the four groups of the continental assessment, the eight guilds, and the water-linked species pooled. In six of them, the shorebirds, the aerial insect-eaters, the crustacean-eaters, the marsh invertebrate-eaters, the scavenging crow, and the scavenging gulls, next year's effort does not follow this year's richness, and the richness test bears on parallel trends. In the other eight it does not bear on them, and I report the test only (p-values 0.10 to 0.97 across all fourteen). The abundance

[37] Omitting effort from the main specification raises the richness estimate from 0.0154 to 0.0242 log points and moves the abundance estimate from −0.148 to −0.047. The gap on richness, 0.0088, is reproduced to within 0.0001 by the effort elasticities times effort's own event-study coefficients. The elasticities are 0.56 for party-hours on abundance (SE 0.12) and 0.10 on richness (SE 0.004), 0.09 and 0.05 for field observers, and 0.09 and 0.03 for miles traveled.

[38] With the three effort measures, party-hours, field observers, and miles traveled, entered in logarithms, as natural splines with four degrees of freedom, or in deciles, the post-period mean ranges from 0.0154 to 0.0159 log points for richness and from −0.163 to −0.145 for abundance, and from −0.0083 to −0.0062, −0.0160 to −0.0138, and −0.0308 to −0.0282 standard deviations for the effective numbers of species and of common species and the entropy.

[39] The test is a Wald test, clustered by circle, that every pre-treatment coefficient, the pooled endpoint included, is zero in the regression that omits year −1. The event-study figures print the same p-value.

test does not bear on parallel trends for all birds, shorebirds, landbirds, the submerged-plant grazers, and the water-linked species, where the balanced sample's event-study coefficients differ from the full sample's.

The tests of the diversity statistics bear on parallel trends in every group. For all birds they reject (p-values 0.008, 0.03, and 0.0006 for the effective numbers of species and of common species and the entropy). Appendix Figure B5 shows the rejection carried by a single year, six before commissioning, in an otherwise flat window. Miller (2023) recommends assessing the window as an overall trend rather than year by year (Section 4.1). I report the rejection as the test returns it. The diversity statistics of waterbirds, waterfowl, shorebirds, and landbirds do not reject (p-values 0.06 to 0.63). On the balanced sample, richness does not reject in any category (p-values 0.15 to 0.79). Abundance rejects there for all birds, shorebirds, landbirds, and the water-linked species (p-values 0.049 and below).

Third, the assignment: pre-program birds do not predict when a circle's first plant was commissioned: given the circle's largest river and its pre-program effort, the 1951 to 1972 level and trend of richness and of abundance leave the first commissioning year unpredicted (p-values 0.23 to 0.83). A pre-program richness level one standard deviation lower raises the probability that a circle is ever downstream of a grant plant by 6.2 percentage points (SE 2.4) and the probability that its best priority-list position is the highest by 5.2 (SE 2.4). The pre-program richness trend does not predict assignment (-1.5 percentage points per standard deviation, SE 1.4). The circle effects absorb that level. Thus, the selection the test finds is on a circle's standing and not on its trend.

Fourth, the gradient: Appendix Figure B4 shows that among circles with no plant upstream, those near a treated circle do not differ from those more than 200 kilometers away: for abundance, richness, and the three diversity statistics, the estimates for the bins within 25 kilometers are jointly indistinguishable from zero (p-values 0.68, 0.31, 0.26, 0.52, and 0.24). The comparison rests on spillovers that reach no circle beyond some distance of a treated one (Butts, 2026). Thus, it tests spillovers near the treated circles, and a draw of birds spread across a whole flyway, which would reach the reference circles too, would not show in it. Richness is lower at 50 to 100 kilometers, by 0.069 log points (p-value 0.006), and at 100 to 200 kilometers, by 0.037 (p-value 0.022). A spillover from treated circles would be largest nearest them. Thus, a dip at 50 to 200 kilometers with none within 25 more likely reflects the regions in which circles without a plant lie than birds moving between circles.

Commissioning a grant-funded plant raised the number of species recorded on the surveys of the circles downstream of it, by about 2% for all birds and about 5% for the water-linked species (Table 1). The literature points to three channels. The first is the water, through the prey base and the vegetation of Section 2: less nitrogen in treatment-plant effluent was

correlated with more submerged vegetation on the Potomac (Ruhl and Rybicki, 2010), and in the other direction diving ducks off Edinburgh had fed on invertebrates enriched by untreated sewage and on grain in the effluent (Campbell, 1978). The second is the plant as a place: British works with percolating filters served birds as feeding sites (Fuller and Glue, 1980), some treatment wetlands are built with wildlife habitat and public use as benefits beside treatment (Knight, 1997), and along the Nile the plants with the most advanced treatment held fewer species than the others (Noby et al., 2025). The third is warmth: below a Swiss plant the stream averaged 8.3 degrees Celsius over a winter against 4.8 above it, and the author attributes the greater insect emergence that drew wintering birds in part to that heat (Lepori, 2022). Warmth comes from the discharge itself rather than from the level of treatment. Thus, a new plant would warm the stream from its commissioning on, while a plant upgraded with a grant already discharged warm effluent before the upgrade and kept doing so after it. The comparison of treated and untreated circles does not identify whether the rise grows with how much of a plant's treated effluent reaches a circle, or which of the three channels the rise runs through, the two questions Section 4.1 leaves open. Section 5 takes up the first question on the intensive margin. It asks whether the rise grows with the dose, the share of the river's flow that a plant's treated effluent makes up at its design capacity. It compares each plant's downstream circle with its own upstream circle, which shares what the plant's siting brings to both and receives none of its effluent (Section 5.1.1).

5 The Role of the Dose and of Water Quality

5.1 Empirical Method

The event study of Section 4.2 leaves two questions open: whether the change in the birds grows with a plant's dose, and how much of it operates through the water quality the plant changed.

5.1.1 The Dose-Response Forms

With the dose in hand, the static form replaces the event indicators of equation (1) with a function of the dose at first commissioning, the dose with which a circle is first treated, in the year of its first commissioning (Callaway et al., 2025):

$$\mathrm{E}\big(N_{ct} \mid \cdot\big) = \phi_c \exp\left(\lambda_t + \sum_{v=1}^{V} \beta_v b_v\big(D_c^{(0)}\big)\mathrm{Post}_{ct} + \mathcal{C}'_{ct}\delta + \mathcal{T}_{ct}\right), \tag{2}$$

where $D_c^{(0)}$ is circle c's dose at first commissioning, $\text{Post}_{ct} = \mathbf{1}(t \geq G_c)$ is the commissioning indicator, equal to one from the year of the circle's first commissioning on, and $b_1, \ldots, b_V$ form a basis of functions of the dose. That is, $\sum_v \beta_v b_v(d)$ is the level effect of dose d, on the scale of the index, among the circles first treated at dose d. It is a level, the $\text{ATT}(d \mid d)$ of Callaway et al. (2025), not the slope of the dose-response curve.[40] For the effective number of species and the entropy, I stack each plant's downstream circle with its own upstream circle:

$$Y_{c,pt} = \alpha_c + \psi_{pt} + \lambda_t^{\text{dir}} \text{Down}_c + \sum_{v=1}^{V} \beta_v b_v\Big(D_{c,pt}\Big) + \mathcal{C}'_{c,pt}\delta + \mathcal{T}_{c,pt} + \varepsilon_{c,pt}, \tag{3}$$

where p indexes a plant and its two circles, ψ_{pt} is an effect common to the circle pair in year t, λ_t^{dir} is a year effect that every pair's downstream circle carries and its upstream circle does not, with Down_c the indicator of the downstream circle, and each circle carries its own dose $D_{c,pt}$. That is, the coefficients compare the two circles on either side of a plant's outfall in the same year, net of how that difference moved in the other circle pairs over the same years, and the upstream circle receives none of that plant's effluent. Equations (2) and (3) define the level effect at each dose that the design identifies: the circle panel is the one form that returns a level at a dose for a count outcome, and the stacked form returns it for the two diversity statistics. The estimates this section reports come from the two dynamic forms that follow, equations (4) and (5), and from static regressions on the dose in force (Tables 3 and 4), which carry one coefficient per unit of dose rather than a level at each dose.

Figure 10 places the three pairs of Figure 6 on one curved response. The pairs of panels A and B both differ by 0.15 in dose, and those equal steps change the outcome by 5.25 and 2.25, because the same step moves the outcome more where it lands higher on the curve. A pair therefore measures the response over the interval between its two doses, and only where its upstream circle sits at zero, as in panel B, does it give the level of the curve itself.

Beyond the static forms, two dynamic forms test the dose's timing. The first is the Poisson analog of the distributed lag in dose changes of Fan (2025),

$$\text{E}\Big(N_{ct} \mid \cdot\Big) = \phi_c \exp\left(\lambda_t + \sum_{\tau \neq -1} \varphi_\tau \Delta D_{c,t-\tau} + \mathcal{C}'_{ct}\delta + \mathcal{T}_{ct}\right), \tag{4}$$

where $\Delta D_{c,t-\tau}$ is the change in circle c's dose in force in year $t - \tau$, so that the terms with $\tau < 0$ are dose steps still to come and those with $\tau \geq 0$ steps already taken, with $\tau = -1$ omitted and the steps more than ten years ahead or fifteen behind pooled at the ends. That

[40] Read as a weighted integral of the level effects, the two-way fixed-effects coefficient on a continuous dose has weights that integrate to zero, negative below the mean dose and positive above it. The coefficient is therefore not an average treatment effect (Callaway et al., 2025). Under parallel trends alone, its decomposition into slopes also carries a selection term.

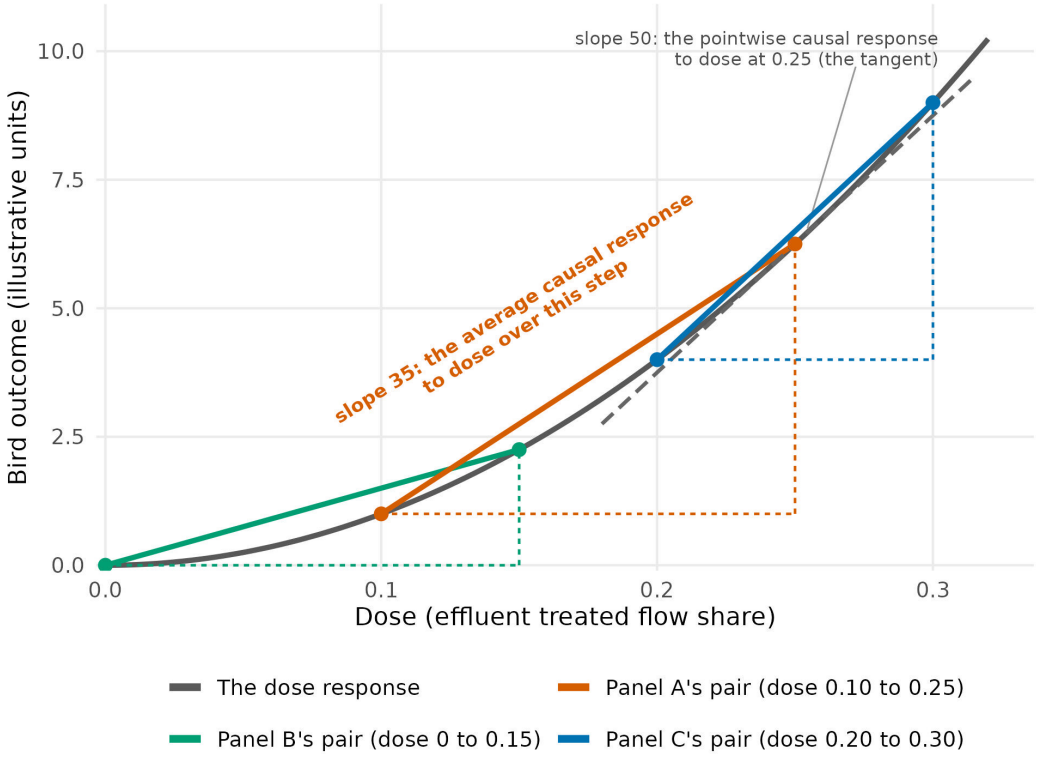


Figure 10: A dose-response curve, drawn with invented round numbers as 100 times the dose squared, with the three pairs of Figure 6 drawn over their own steps in dose, each in its panel's color. The solid colored line joins the two points of each circle pair, and its slope, 35 in panel A, is the average response over the pair's step in dose. The dashed line is the tangent at a dose of 0.25, whose slope of 50 is the response at that dose.

is, φ_τ is the proportional change in expected abundance or richness τ years after a dose step of one, relative to the year before the step, and a circle that steps several times carries the sum of its steps' contributions. In each year, the comparison is between a circle whose dose rises and the circles whose dose does not rise that year, the never-treated circles among them, with the circle effects absorbing each circle's level. The pre-trend test, the joint test that the pre-treatment coefficients are zero, checks that the dose's whole history is strictly exogenous, the steps after the first commissioning included, which is where a later plant sited where birds were already recovering would show. For the diversity statistics the same form is estimated in first differences, as Fan (2025) estimates it, so that each coefficient is the year-to-year increment of the response. The second form keeps the treated circles only and splits them within each cohort bin at the median of their dose at first commissioning (Fan, 2025):

$$\mathrm{E}\Big(N_{ct} \mid \cdot\Big) = \phi_c \exp\left(\lambda_{b(c)t} + \sum_{k \neq -1} \beta_k^{\mathrm{H}} \mathrm{High}_c \mathbf{1}(t - G_c = k) + \mathcal{C}_{ct}'\delta + \mathcal{T}_{ct}\right), \tag{5}$$

where $b(c)$ is circle c's cohort bin, consecutive cohorts pooled until both halves of the split hold six circles, $\lambda_{b(c)t}$ is a bin-by-year effect, and High_c is the indicator of a dose at first commissioning above the bin's median. That is, in each year the circles in the higher half of a cohort bin are compared with the circles in the lower half of the same bin, each half measured from its own nine years before commissioning, and no untreated circle enters. The coefficient β_k^{H} says whether the circles that received more of the effluent rose more than the circles that received less, and nothing about whether either rose, and its pre-treatment coefficients test parallel trends between the halves. I estimate both forms, and the static forms above, with the plant count in place of the dose. A response that scales with the plant count and not with the dose points at the plants as habitat or at where they were sited, rather than at the dilution of their effluent.

The slope of the dose-response curve is not identified under parallel trends alone: a comparison of two doses carries a selection term that only strong parallel trends, a common change at every dose across dose groups, removes (Callaway et al., 2025). The level at each dose is identified under parallel trends, which the pre-treatment coefficients of these forms test. No pre-period can test strong parallel trends, since before commissioning only untreated outcomes are observed (Callaway et al., 2025).

5.1.2 The Water-Quality Channel

Within each half of the split by the plants' effluent before their grants (Section 3.6), the cleaner or the dirtier, indexed by ς, the dose can affect the birds in two ways, which add on

the scale of the index: through the water quality it changes, and directly, through anything else a commissioning brings to a circle. Write a_ς for the treatment effect on water quality in the half, the change in an analyte's reading per unit of dose, β_M^{wq} for the birds' response to water quality, the change in the index per unit of the reading, and β_D^{dir} for the direct effect. The reduced form of a bird statistic on the dose, the regression of the statistic on the dose alone with no water-quality variable beside it,[41] then has slope

$$\tau_\varsigma = \underbrace{\beta_D^{\text{dir}}}_{\text{the direct effect}} + \underbrace{\beta_M^{\text{wq}} a_\varsigma}_{\text{the effect through the water-quality channel}}, \tag{6}$$

That is, a unit of dose moves the index by its direct effect plus the birds' response to water quality times the treatment effect on water quality. The data identify τ_ς and a_ς, but not the split between the two terms: any pair of effects with the same sum produces the same data once the circle and year effects absorb each circle's level of the reading. Because the dose also affects the birds directly, the ratio $\tau_\varsigma / a_\varsigma$ does not identify the birds' response to water quality (Conley et al., 2012). A restriction on signs turns this into a bound. If neither effect offsets the other, the effect through the water-quality channel lies between zero and τ_ς, and the birds' response to water quality between zero and the ratio $\tau_\varsigma / a_\varsigma$. That ratio is what a two-stage least squares regression with the dose as its instrument would estimate. It equals the birds' response only where the direct effect is zero, and I therefore report it as the far end of the bound and never as an estimate. On the proportional scale of abundance and richness the two effects multiply rather than add. Appendix A states the decomposition and the bound formally (Proposition A.2).

I define the effect through the water-quality channel broadly, as every part of the dose's effect that moves with the measured cleaning. Its leading interpretation is trophic, through the prey and the plants that cleaner water supports. Its rival is the plant as habitat, which the technology classes test.

I estimate the treatment effect on water quality, each analyte's response to commissioning, from the readings of Section 3.6 in three steps. First, I put each reading on its within-circle standard score: the reading less the mean of the circle's readings, over their standard deviation. A coefficient is then in standard deviations of a circle's own readings, and a circle whose readings vary widely counts no more than one whose readings vary little. Second, I regress the score on commissioning by least squares with circle and year effects and the controls of equation (1), in two forms. The event-study form is that of Section 4.1, with each coefficient measured against the average of the nine years before commissioning.

[41] The term follows van Kippersluis and Rietveld (2018), whose equation (2.4) regresses the outcome on the instrument alone, with a coefficient equal to the instrument's direct effect plus its effect through the endogenous variable.

The static form regresses the score on the commissioning indicator Post_{ct} of equation (2), and its coefficient is the average change in the score over the years after commissioning. Third, I regress the score on the dose in force, separately in each half of the split, which is equation (7) below. An effect in standard deviations converts to the analyte's physical unit by the pooled within-circle standard deviation of the readings before commissioning over the growing season (Section 3.6): for dissolved oxygen that standard deviation is 13.3 percentage points of saturation, for phosphorus 0.66 milligrams per liter, and for fecal coliform 1.5 on the logarithmic scale. The regressions run on the circle-years with a reading, fewer than the surveys, and a circle's readings after commissioning are an outcome of commissioning and never enter a bird regression as a control.

The split is identified where the readings respond differently to the same dose, which the comparison of the two halves uses. Its first step is the regression of the standard score on the dose in force in each half, the treatment effect on water quality there:

$$M_{ct} = \alpha_c + \lambda_t + \sum_{\varsigma} \beta^M_{\varsigma} \mathbf{1}\big(\varsigma(c) = \varsigma\big) D_{ct} + \big(\mathcal{C}^w_{ct}\big)' \delta^w + \mathcal{T}_{ct} + \varepsilon^M_{ct}, \tag{7}$$

where M_{ct} is circle c's standard score in year t, $\varsigma(c)$ its half, D_{ct} its dose in force, $\mathcal{C}^w_{ct}$ the controls of equation (1), and β^M_{ς} estimates a_{ς}. Predicted cleaning is the fitted slope of a circle's half times its dose,

$$\hat{P}_{ct} \equiv \hat{\beta}^M_{\varsigma(c)} D_{ct}, \tag{8}$$

That is, $\hat{P}_{ct}$ is the change in the reading that circle c's dose would produce at its half's estimated response.[42] Entered beside the dose in the static form and in the two dynamic forms, predicted cleaning takes the coefficient that is the birds' response to water quality, and the dose takes the coefficient that is the direct effect. The split holds provided neither differs between the halves, so that the effluent before the grants affects the birds only through the treatment effect on water quality (Conley et al., 2012). Before any of this, a Wald test checks whether the treatment effect on water quality differs between the halves at all, since where it does not the halves identify nothing.

Three further comparisons rest on other assumptions, and their agreement tests each. The capacity comparison uses the commissionings that only added capacity, after which the readings downstream should not change: their effect on the birds estimates the direct effect, under the assumption that a commissioning of capacity alone carries the same direct effect per unit of dose as any other, the "homogenous direct effects" of van Kippersluis and Rietveld

[42] Predicted cleaning is the fitted value of equation (7) at the circle's dose. A second version uses the effluent cut instead of the slope of the half: it is the coefficient of a water-quality regression on the dose times the cut, multiplied by the circle's dose and its mean cut. The two versions predict the same change from different variation, and their agreement checks both.

(2018). The cut comparison identifies the split from the variation in the plants' effluent cut at a given dose, under the assumption that the cut enters the dose's effect only through water quality. The location comparison uses the circles whose dose comes only from plants upstream and outside them, under the assumption that a plant's channels other than water quality reach only the birds of the circle that contains it. Because the comparison of the two halves, the capacity comparison, and the cut comparison cannot tell the effect through water quality from a direct effect that grows in proportion to the cleaning itself, each of the three estimates an upper bound on the effect through the water-quality channel. Only the location comparison escapes this, and only for a direct channel that stays inside the circle.

The circle effects absorb what a plant's presence meant for a circle before its grant, and the year effects the shocks common to every circle. The direct effect is what remains of the dose's effect apart from water quality, and the tests of the plant as a place at the end of this section are its falsification tests.

5.1.3 Addressing Confounders for the Dose

Three things could confound the dose's effect on birds. The first is selection on gains. A flow share is design flow over the river's mean annual flow, and design flow is chosen: a town sizes its plant for its present population and its expected growth,[43] and a growing town is also where the birds' trends, and their response to cleaner water, may differ. Three facts address it. First, most of the flow share's variation is the river's: across plant-circle pairs at 25 river miles, the outlet's mean annual flow accounts for 65.7% of the variance of the flow share's logarithm and the plant's design flow for 41.6%, with a correlation of 0.07 between the two. Second, a plant's circles share its design flow. A comparison within one plant across the circles it reaches therefore holds the chosen term fixed and lets only the river vary. I estimate the response within plants, with the flow share split into its design flow and its mean annual flow, and with a dose that gives every plant one common design flow. Third, I test whether a circle's birds before the program predict the dose it later received, given the size of its largest river.

The second is the technology margin. A grant could have paid for a technology rather than for capacity, and ponds and lagoons need cheap, flat land with a high water table, land often next to wetlands and rich in birds. The choice of technology can therefore mark site characteristics that predict the birds' levels and trends. The commissionings that only added capacity carry no cleaning by construction, and their effect therefore estimates the

[43] Design flow is a plant's capacity at year 20 of its assumed life, and across 68 plants the Congressional Budget Office found reserve capacity of 0 to 750% of the initial flow (Congressional Budget Office, 1985, p. 19).

direct effect alone. I split the bird estimates by technology class, read as heterogeneity by a characteristic chosen with the grant and never as the effect of the technology itself. I also test whether a circle's birds before the program predict a pond plant later built upstream of it.

The third is the plant as a place. Ponds, lagoons, and ice-free water at a plant can draw birds regardless of any change in water quality downstream. At sewage works in Britain, bird numbers rose in hard weather, when a sewage farm's flooded meadows stayed unfrozen and its percolating filters kept a food supply in any weather (Fuller and Glue, 1980). A plant's effluent can also raise a stream's winter temperature from near freezing to 5 to 12 degrees Celsius below the outfall (Adams et al., 2023), and such warmth can advance the emergence of the aquatic insects that wintering insectivores feed on (Lepori, 2022). Observers, too, can be drawn to a new stretch of open water. Four tests separate these channels from water quality. First, plants whose effluent reaches no surface water carry every channel at the plant and none through water quality. Their commissioning should therefore move no bird statistic. Second, pond plants and conventional plants should differ at commissioning only where the plant lies inside the circle, since a pond's habitat reaches only the birds near it. Outside the circle the difference is therefore a placebo. Third, aerated lagoons stay open in a freezing winter, the season of the survey, while stabilization ponds freeze. A habitat channel therefore predicts a difference between them in freezing circles. Fourth, outside the circle the effect through water quality decays along the river, while a pull from a plant's ponds falls with its distance over land. None of the four separates birds drawn to the ponds from observers drawn to them, and survey effort, which does not step at commissioning (Section 4.2.2), is the one test of the second.

5.2 Results

This section asks whether the birds' response grew with a plant's dose and how much of it operated through water quality. Table 2 answers the first question for the dose and for the plant count. Tables 3 and 4 answer the second: how each analyte's reading responded to commissioning, and how the dose's effect splits between the water quality it changes and everything else a commissioning brings.

Table 2 reports the distributed lag in dose changes of equation (4), on the dose and on the plant count, for abundance and richness. Appendix Table B10 reports it for the three diversity statistics. Its pre-trend test checks that the dose's whole history is strictly exogenous. It rejects in 76 of the 84 combinations of species group and statistic whose fit converged, and for all birds on richness and the three diversity statistics (p-values below

0.001), not rejecting only on abundance (p-value 0.19). The responses fourteen years after a dose step, among them 0.011 standard deviations of the effective number of species per unit dose (SE 0.002), therefore combine any effect of the dose with differences that were present before the dose rose (Appendix Table B10 and Figure 11). The median split's pre-trend test does not reject for richness and the diversity statistics (p-values 0.25 to 0.63) and rejects for abundance (p-value 0.020). After commissioning, the circles with a higher dose at first commissioning did not rise more than the circles with a lower one: the higher half's coefficients average below the lower half's on richness and the three diversity statistics, and one year of the sixty is distinguishable from zero (Appendix Figure B8).[44]

By contrast, the plant count's pre-trend tests do not reject in 68 of the 85 combinations, among them every statistic for all birds. Each additional commissioned plant upstream raises richness fourteen years on by 0.0032 log points for all birds (SE 0.0014). The rise is larger among the species tied to water: 0.0096 for the water-linked species pooled (SE 0.0029), 0.013 for waterbirds (SE 0.004), 0.011 for waterfowl (SE 0.004), and 0.026 for the submerged-plant grazers (SE 0.005), against 0.0002 for landbirds (SE 0.0014), 20 of whose 350 species are water-linked (Figure 12).[45] Richness thus follows the number of plants among the species tied to water and not among landbirds. The dose's median split, by contrast, finds no rise with the plants' share of the river. A response that follows the plants and not their share of the river points at the plant itself, or at where the plants were placed, rather than at dilution.

Table 3 shows that, after commissioning, two analytes changed in the direction of cleaner water. On the within-circle standard score, dissolved oxygen saturation rose by 0.15 standard deviations (SE 0.066), about 2 percentage points of saturation, and total phosphorus fell by 0.15 (SE 0.066), about 0.1 milligrams per liter, each converted by the pooled within-circle standard deviation of Section 5.1.2. Oxygen demand, suspended solids, fecal coliform, the three nitrogen analytes, and lead show no change distinguishable from zero. These are the static estimates on the commissioning indicator, averages over the years after commissioning. The event studies of the regulated analytes and of phosphorus do not reject in their pre-trend

[44] Quintiles of the dose among the treated circles, estimated for all birds with circle and year effects alone, add nothing further. The pre-trend test does not reject in any quintile for richness (p-values 0.06 to 0.87). It rejects in four of five quintiles for abundance and in 7 of 15 for the three diversity statistics. The quintiles are cut on the dose in force in 2024 before the cap on each plant's flow share (Section 3.4). Thus, the top quintile runs from 0.11 to 813.5, against 0.023 to 0.108 for the fourth, and describes no typical circle.

[45] The submerged-plant grazers' diversity statistics also rise with the plant count, by 0.048 standard deviations of the effective number of species fourteen years on (SE 0.010), and the waterbirds' do not (−0.003, SE 0.006). For both groups, however, the test of whether later plants were sited where these birds already were rejects (p-values 0.0002 to 0.026). The grazers' estimate may therefore reflect where later plants were built as much as what the plants did.

Table 2: Distributed-lag estimates of the effect of the dose and of the number of commissioned grant plants upstream on bird abundance and richness, by species group

	Abundance		Richness	
	(1)	(2)	(3)	(4)
	Dose step	Plant added	Dose step	Plant added
	Response at event time 14, log points			
Panel A. All birds				
All birds	−0.00023	0.10	0.00057**	0.0032**
	(0.0070)	(0.071)	(0.00026)	(0.0014)
Pre-trend test p-value	0.19	0.21	<0.001	0.12
Surveys	85,934			
Panel B. The water-linked species				
Water-linked species, pooled	−0.0027	−0.042	0.00055	0.0096***
	(0.011)	(0.12)	(0.00035)	(0.0029)
Pre-trend test p-value	<0.001	0.05	<0.001	0.03
Surveys	85,534			
Submerged-plant grazers	0.0071***	0.029	0.0019*	0.026***
	(0.0023)	(0.060)	(0.00096)	(0.0052)
Pre-trend test p-value	0.004	0.07	<0.001	0.13
Surveys	82,765			
Panel C. The general groups				
Waterbirds	0.00039	0.015	0.00038	0.013***
	(0.0019)	(0.030)	(0.00037)	(0.0040)
Pre-trend test p-value	<0.001	0.11	<0.001	0.05
Surveys	82,090			
Waterfowl	−0.0027	0.024	0.00053	0.011***
	(0.0019)	(0.035)	(0.00048)	(0.0036)
Pre-trend test p-value	<0.001	0.09	<0.001	0.28
Surveys	85,135			
Shorebirds	−0.0022	−0.0071	0.00071*	0.0029
	(0.0017)	(0.039)	(0.00042)	(0.0068)
Pre-trend test p-value	<0.001	0.36	<0.001	0.03
Surveys	70,677			
Landbirds	0.0042	0.12	0.00065**	0.00018
	(0.011)	(0.11)	(0.00029)	(0.0014)
Pre-trend test p-value	0.24	0.21	<0.001	0.13
Surveys	85,934			
Circle and year effects	Yes	Yes	Yes	Yes
Effort controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Attribute-by-year trends	Yes	Yes	Yes	Yes

Notes: The distributed lag regression, equation (4), is fit twice: once on the dose, and once on the number of commissioned grant plants upstream, which is the dose with every plant's flow share set to one. Each estimate is the coefficient fourteen years after the dose rises by one, or after one more plant is commissioned, measured against the year before, by Poisson quasi-likelihood in log points, with circle and year effects and the controls of Table 1. Standard errors clustered by circle are in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels. Each pre-trend test is the Wald test that the coefficients on the dose steps that lie ahead are zero. The panels follow the order of the paper's groups: all birds, then the water-linked species pooled and the submerged-plant grazers, then the four groups of the continental assessment. The three diversity statistics are in Appendix Table B10, and the median split's pre-trend tests are reported in the text.

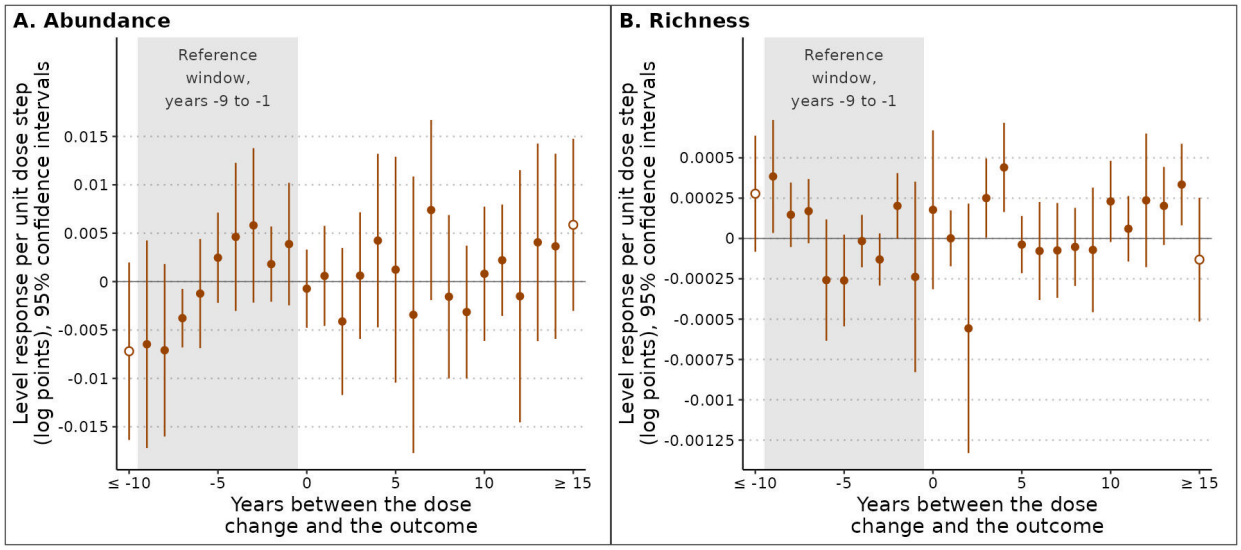


Figure 11: Response of all birds' abundance and richness to a step in the dose, by year relative to the step, Christmas Bird Count circles, 1940 to 2024.

Notes: Event study of equation (4) at the main specification, on 85,934 surveys: panel A, abundance, and panel B, richness, both in log points per unit dose, by Poisson quasi-likelihood with circle and year effects and the controls of Section 4.1. Each coefficient is the response to a step of one in the dose taken that many years earlier, negative years being steps still to come, and is measured against the average of years −9 to −1, the shaded window (Miller, 2023). The first and last points pool the steps 10 or more years ahead and 15 or more behind. They are estimated differently from the interior coefficients and are drawn hollow. Whiskers are 95% confidence intervals from standard errors clustered by circle. The pre-trend test, the joint test that the coefficients on the dose steps still to come are zero, has a p-value of 0.19 for abundance and below 0.001 for richness. The window averages around which the coefficients are normalized are −0.0039 (SE 0.0032) for abundance and 0.00024 (SE 0.00030) for richness.

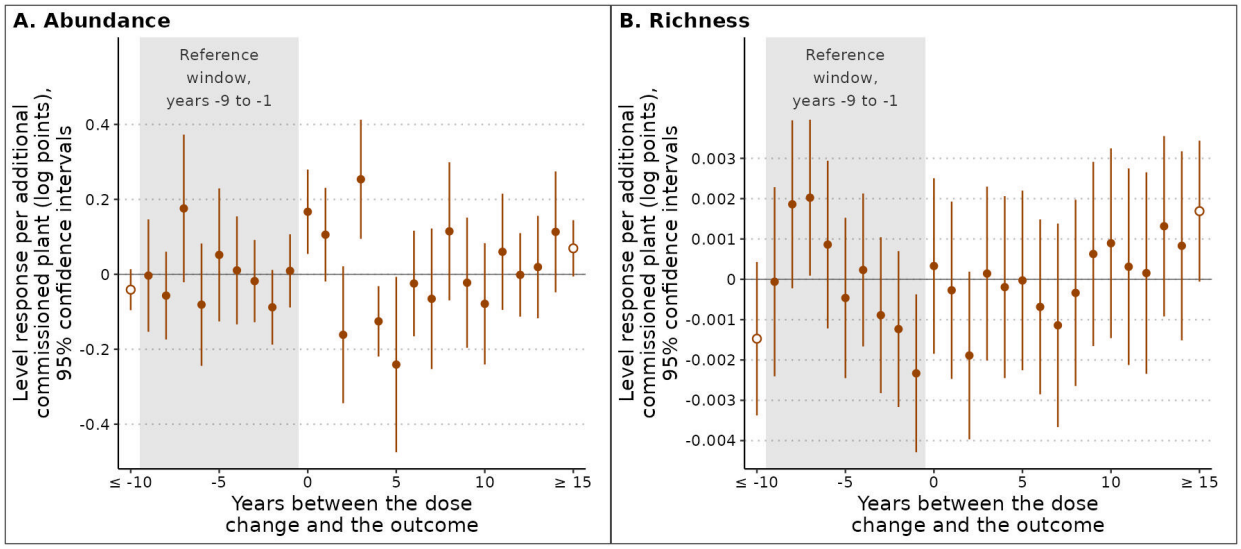


Figure 12: Response of all birds' abundance and richness to an additional commissioned plant upstream, by year relative to its commissioning, Christmas Bird Count circles, 1940 to 2024.

Notes: Event study of equation (4) with the plant count in place of the dose, at the main specification, on 85,934 surveys: panel A, abundance, and panel B, richness, both in log points per plant. Each coefficient is the response to one more commissioned plant upstream that many years earlier, negative years being commissionings still to come. The pre-trend test has a p-value of 0.21 for abundance and 0.12 for richness, and the window averages around which the coefficients are normalized are −0.009 (SE 0.050) and 0.00233 (SE 0.00100). Otherwise as in Figure 11.

tests (p-values 0.06 to 0.74; Appendix Figures B9 and B10).[46] Oxygen demand's event-study coefficients rise late, by about 0.3 to 0.4 standard deviations at years 10 to 15, about 1.4 to 1.8 milligrams per liter. The earliest cohorts alone identify those years, on circles whose set of readings is not held fixed. The rise may therefore reflect which circles and readings remain in those years as much as an increase in oxygen demand below the plants.

Split by the plants' effluent before their grants, the treatment effect on water quality per unit of dose differs between the two halves for suspended solids and fecal coliform, which the comparison of the two halves in Section 5.1.2 requires. For suspended solids, the reading fell by 0.035 standard deviations per unit of dose in the circles below plants with dirtier effluent (SE 0.007), about 2.9 milligrams per liter, and rose by 0.017 in the cleaner half (SE 0.003), about 1.4 milligrams per liter, a difference with a p-value below 0.001. For fecal coliform, split by the plants' treatment level (Section 3.6), the halves differ at a p-value of 0.041, and for oxygen demand (p-value 0.36) and dissolved oxygen (p-value 0.59) they do not.[47] With the commissioning indicator in place of the dose, dissolved oxygen also differs between the halves, falling 0.25 standard deviations in the dirtier half, about 3 percentage points of saturation, and rising 0.26 in the cleaner, about 3.5 (p-value 0.004). Figure 13 draws the event studies on the commissioning indicator for each half. Fecal coliform in the dirtier half falls to about half a standard deviation below its reference by the tenth year, roughly half as much fecal coliform, and to 0.7 to 0.8 below at years 14 and 15, roughly 65 to 70% less. The cleaner half stays flat, and no series rejects in its pre-trend test (p-values 0.19 to 0.72).

Table 4 shows that, for all birds, the comparisons of Section 5.1.2 disagree in their static forms. With predicted cleaning from the plants' cut in suspended solids beside the dose, the dose's coefficient on the effective number of species is unchanged, 0.0063 against 0.0056 per unit dose alone. Split by the plants' suspended-solids effluent, predicted cleaning is distinguishable from zero on none of the five statistics at the 5% level (Table 4 and Appendix Table B11). Split by the plants' treatment level and measured on fecal coliform, the whole static coefficient is attributed to predicted cleaning, 0.61 standard deviations of the

[46] Of the nitrogen analytes, total nitrogen falls by about a quarter of a within-circle standard deviation from the fourth year after commissioning, with its pre-trend test not rejecting (p-value 0.86), although its static estimate is positive (Appendix Figure B11). Under staggered timing a static estimate of a response that moves over the years after commissioning weights those years in ways it does not control, and I therefore report the event study. Ammonia rises with its pre-trend test rejecting (p-value 0.04), and nitrate plus nitrite rises with the dose but not with commissioning. Lead does not move (0.041, SE 0.045).

[47] On suspended solids, the reading rises with the dose in the cleaner half, and on fecal coliform it rises in the dirtier half, by 0.060 standard deviations per unit of dose (SE 0.026), about 9% more fecal coliform, against 0.007 (SE 0.005), about 1%, in the cleaner half. A plant's larger cut in suspended solids predicts cleaner suspended solids at the circles it doses, by 0.0035 standard deviations per milligram per liter of cut per unit of dose (SE 0.0011), about 0.3 milligrams per liter. Its cut in oxygen demand is not distinguishable from zero (−0.009, SE 0.006, p-value 0.11).

Table 3: Least-squares estimates of the effect of commissioning and of the dose on water-quality readings, by analyte and by the plants' effluent before their grants

Panel A. All treated circles

	(1) Effect of commissioning, standard deviations	(2) Pre-trend test p-value	(3) Circle-years
Regulated			
Biochemical oxygen demand	0.047	0.62	27,449
	(0.050)		
Total suspended solids	−0.034	0.31	66,153
	(0.043)		
Fecal coliform, log	0.00073	0.74	45,643
	(0.034)		
Dissolved oxygen saturation	0.15**	0.06	36,401
	(0.066)		
Partly regulated			
Total phosphorus	−0.15**	0.20	57,804
	(0.066)		
Nitrogen			
Ammonia	0.082	0.04	53,583
	(0.062)		
Total nitrogen	0.053	0.86	47,188
	(0.045)		
Nitrate plus nitrite	0.038	0.13	43,889
	(0.069)		
Falsification			
Lead (Pb)	0.041		43,859
	(0.045)		

Panel B. Treated circles split at the median of their plants' effluent before their grants, by analyte

	Per unit of dose			Commissioning		
	(4) Dirtier half	(5) Cleaner half	(6) p-value, halves equal	(7) Dirtier half	(8) Cleaner half	(9) p-value, halves equal
Biochemical oxygen demand	0.048**	0.016	0.36	0.071	−0.024	0.47
	(0.020)	(0.030)		(0.13)	(0.092)	
Circle-years		8,617			8,617	
Total suspended solids	−0.035***	0.017***	<0.001	−0.14	−0.034	0.29
	(0.0071)	(0.0031)		(0.10)	(0.095)	
Circle-years		23,840			23,840	
Fecal coliform, log	0.060**	0.0071	0.04	−0.23	0.039	0.09
	(0.026)	(0.0051)		(0.16)	(0.077)	
Circle-years		14,820			14,820	
Dissolved oxygen saturation	−0.043	0.0044	0.59	−0.25	0.26**	0.004
	(0.089)	(0.021)		(0.15)	(0.13)	
Circle-years		12,940			12,940	
Circle and year effects			Yes			
Effort controls			Yes			
Weather controls			Yes			
Attribute-by-year trends			Yes			

Notes: Each estimate is from a static least-squares regression of an analyte's within-circle standard score, the reading less the mean of the circle's readings over their standard deviation, with circle and year effects and the controls of equation (7): effort, winter weather, and baseline housing density and cropland share interacted with year indicators. Panel A regresses the score on the commissioning indicator Post_{ct}, equal to one from the year of the circle's first commissioning, on every circle-year with a reading, and gives the p-value of the event study's pre-trend test on the same circle-years. Lead has no event study. Panel B splits the treated circles within each bin of commissioning cohorts at the median of their plants' effluent before the grants (Section 3.6): oxygen demand and dissolved oxygen on the plants' oxygen demand, suspended solids on their suspended solids, and fecal coliform on the share of the dose from plants below secondary treatment. It regresses the score on the dose in force and on the commissioning indicator in each half. The p-value tests that the two halves' coefficients are equal. Fecal coliform enters in logarithms. Standard errors clustered by circle are in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels. Appendix Figures B9 to B11 and Figure 13 give the event studies.

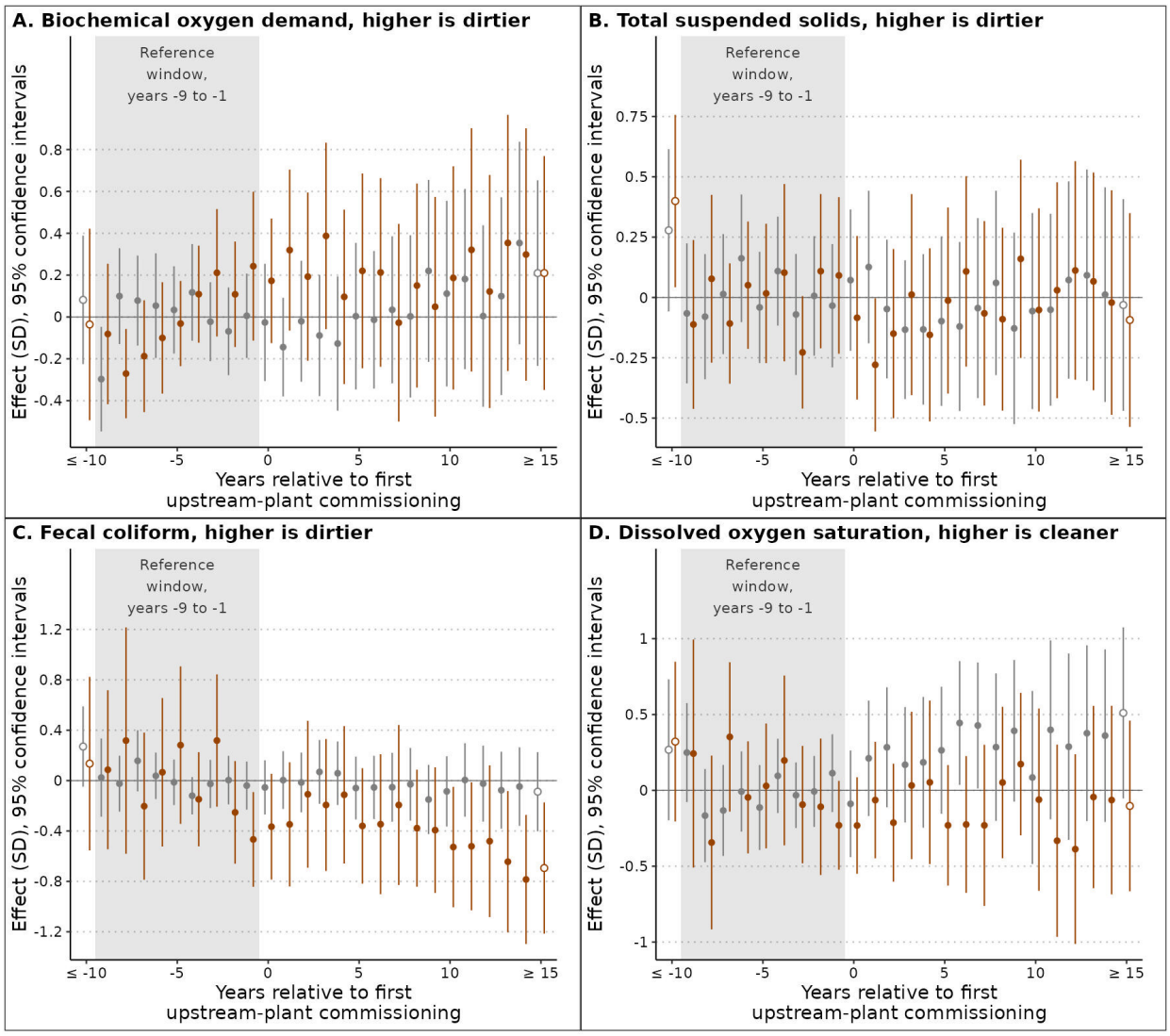


Figure 13: The four regulated analytes by year relative to the first commissioning of a grant-funded plant upstream, in the circles whose plants discharged dirtier and cleaner effluent before their grants.

Notes: Event study of each analyte's within-circle standard score on commissioning, by least squares with circle and year effects and the controls of equation (7), estimated separately in the dirtier and the cleaner half of the treated circles, the halves of Table 3, Panel B, each against the circles with no plant upstream. Each coefficient is measured against the average of years −9 to −1, the shaded window (Miller, 2023). The first and last points pool every year 10 or more before and 15 or more after commissioning and are drawn hollow. Within each year the cleaner half's point is drawn to the left of the year and the dirtier half's to the right. Whiskers are 95% confidence intervals from standard errors clustered by circle. The cleaner and the dirtier half hold 5,318 and 4,215 circle-years for oxygen demand, 14,561 and 12,391 for suspended solids, 12,865 and 3,750 for fecal coliform, and 8,072 and 6,987 for dissolved oxygen. In the same order, the pre-trend test's p-values are 0.34 and 0.40, 0.53 and 0.35, 0.63 and 0.19, and 0.22 and 0.72, and the window averages around which each series is normalized are −0.01 (SE 0.10) and −0.24 (SE 0.18), 0.03 (SE 0.13) and −0.09 (SE 0.17), 0.039 (SE 0.097) and 0.47 (SE 0.19), and −0.11 (SE 0.13) and 0.23 (SE 0.15). Fecal coliform enters in logarithms.

effective number per standard deviation of predicted cleaning (SE 0.24), and the dose's own coefficient falls to zero. Under the assumptions of Section 5.1.2 the comparisons estimate one direct effect. Thus, either the analytes do not reflect one change in water quality or one comparison's assumption fails. In event time every comparison rejects in its pre-trend test (p-values below 10^{-8}), and the capacity comparison lacks its premise, since fecal coliform fell by 0.45 standard deviations (SE 0.10), about half as much fecal coliform, in the circles whose largest commissioned share came from a plant that only added capacity. The data therefore do not identify how much of the dose's effect on the birds operates through water quality.[48]

The tests of Section 5.1.3 find selection on gains in the dose, no placebo among the commissionings that only added capacity, and a channel at the plant itself. First, on selection on gains, the design flow and the river's flow enter as the share requires: entered separately, their logarithms carry equal and opposite coefficients (p-values 0.62 on abundance and 0.72 on richness for their sum). Among the circles below one plant, those taking a larger share of its flow do not respond more (abundance 0.019, SE 0.041, over 628 plants and 1,818 plant-circle pairs).[49] But a circle's birds before the program predict the dose it later received. The logarithm of the dose a circle ends with rises by 0.54 per standard deviation of its abundance before the program and falls by 0.30 per standard deviation of that abundance's trend over 1951 to 1972 (joint p-value below 0.001). Neither the grant dollars awarded upstream nor their eligible cost is predicted (joint p-values 0.38 and 0.49). This selection on gains in the dose itself is why I report no slope of the dose-response curve.

Second, on the technology margin, a circle's birds before the program do not predict a pond, wetland, or land-treatment plant later built upstream of it (p-values 0.09 or above on each statistic), but the commissionings that only added capacity are no placebo. Among the circles whose largest commissioned share came from such a plant, fecal coliform fell by 0.45 standard deviations (SE 0.10), about half as much fecal coliform, and abundance rose by 0.68 log points (SE 0.36, pre-trend p-value 0.06). The circles dosed by such plants alone are too few to interpret. Neither therefore estimates the direct effect.

Third, on the plant as a place, a plant inside a circle that does not drain to the circle's outlet still raises its birds, by 0.31 log points of abundance (SE 0.14) and 0.087 standard

[48]The online supplement reports the association between the birds and water quality across circles, with a correction for measurement error in the readings. The association is descriptive and does not estimate the birds' response to water quality.

[49]The dose built from the river alone, with one common design flow for every plant, fails its pre-trend tests in both dynamic forms. In the distributed lag, for all birds, the test rejects on richness and the three diversity statistics (p-values below 0.001) and does not reject only on abundance (p-value 0.96). Circles with larger and smaller doses of this kind were already diverging before commissioning, and its estimates after commissioning cannot be told apart from that divergence.

Table 4: Two-way fixed-effects estimates of the effect of the dose on all birds, with and without predicted cleaning

	(1) From the plants' cut in suspended solids	(2) From the halves split on the plants' suspended solids before the grants	(3) From the halves split on the plants' treatment level, on fecal coliform
Panel A. Abundance, Poisson quasi-likelihood, in log points			
Dose alone	0.0067*	0.0068*	0.0082
	(0.0039)	(0.0038)	(0.0050)
Dose, with predicted cleaning	0.0061	0.0073	0.020*
	(0.0039)	(0.0067)	(0.012)
Predicted cleaning	−0.012*	0.017	−1.5
	(0.0064)	(0.13)	(1.9)
Surveys	79,120	82,810	81,950
Panel B. Richness, Poisson quasi-likelihood, in log points			
Dose alone	−0.00025***	−0.00021**	−0.00017
	(0.000089)	(0.00010)	(0.00014)
Dose, with predicted cleaning	−0.00025***	−0.00036	−0.0011**
	(0.000088)	(0.00023)	(0.00042)
Predicted cleaning	−0.00013	−0.0051	0.12**
	(0.00036)	(0.0061)	(0.057)
Surveys	79,120	82,810	81,950
Panel C. Effective number of species, least squares, in standard deviations			
Dose alone	0.0056***	0.0055***	0.0044**
	(0.0015)	(0.0015)	(0.0020)
Dose, with predicted cleaning	0.0063***	0.0045**	−0.00028
	(0.0012)	(0.0021)	(0.0028)
Predicted cleaning	0.012	−0.039	0.61***
	(0.011)	(0.065)	(0.24)
Surveys	79,120	82,810	81,950
Circle and year effects	Yes	Yes	Yes
Effort controls	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
Attribute-by-year trends	Yes	Yes	Yes

Notes: Each column is one static regression of the statistic, for all birds, on the dose in force and predicted cleaning, equation (8), with circle and year effects and the controls of Table 1. The row for the dose alone is the same regression without predicted cleaning, on the same surveys. The coefficients on the dose are per unit of dose, and those on predicted cleaning per within-circle standard deviation of the predicted reading of the analyte. Column (1) builds predicted cleaning from the plants' cut in suspended solids, and columns (2) and (3) from the two halves' estimated effects on water quality, in suspended solids and in fecal coliform (Table 3, Panel B). Standard errors clustered by circle are in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels. The effective number of common species and the entropy are in Appendix Table B11. The event-study versions of the three comparisons reject their pre-trend tests and are not shown.

deviations of the effective number of species (SE 0.033) per plant. A plant with no route to surface water raises them by 0.28 (SE 0.21) and 0.098 (SE 0.053), though these event studies are thin and their abundance pre-trend tests reject (p-value 0.02). In freezing circles aerated lagoons and stabilization ponds do not differ in abundance (−0.45, SE 0.51), outside the circle the response does not fall with distance over land (0.026 per kilometer, SE 0.035), and the comparison of pond plants with conventional plants outside the circle is degenerate. A channel at the plant other than water quality therefore cannot be ruled out, and none of these tests separates birds drawn to the plants' ponds from observers drawn to them.[50]

Finally, the submerged-plant grazers are the one category whose abundance rises in both forms. Their estimate of 0.351 log points, about 42%, means about 800 more grazers per survey at the treated circles' average of 1,900 before commissioning and about 27 at the median of 64, and with the extreme counts dropped 0.247 log points, about 28%, means about 380 at the average of 1,338 and about 18 at the median. Scaled to the continent as in Section 4.2.1, the additional grazers with commissioning come to 7.9 million (95% interval 1.8 to 12.5 million) on the full counts and 6.1 million (0.1 to 10.7 million) with the extreme counts dropped, of a population of 33.1 million across 17 species. Like the water-linked birds' figure, it is an upper bound.[51]

6 Discussion

6.1 Benefits and Costs

How do these effects bear on the program's benefit-cost record? Keiser and Shapiro (2019a) compare what each grant added to housing values, which they take as its measured benefits, with its costs: local and federal capital expenditures and operating and maintenance costs over the 30 years in which they estimate the grants affect water pollution. Counting owner-occupied homes up to 25 miles from the river, the ratio of the grants' effect on home values to their costs is 0.26, and their confidence regions reject a ratio of one but not a ratio of

[50] Higher-dose circles also carry a larger summed flow share from the works the grants did not fund, a mean share of 0.089 in the highest dose quartile against 0.006 in the lowest, though the share of circles with any such works is flat across the quartiles (0.32 to 0.37). The dose groups also differ before treatment: the largest standardized difference between them, 1.5, is in the flow of the largest river, and 36 of 65 exceed 0.25, as the dose's dependence on the river implies.

[51] For all birds the main estimate's interval spans zero, and the increase from the regression with a coefficient for every cohort and year comes from one circle's blackbird roost. Thus, I report no total for all birds. With the extreme counts dropped, the same calculation gives 638.5 million birds (109.3 to 1,119.5 million), a figure that rests on how the largest counts are treated.

zero.[52] Scored as annual flows over 1962 to 2001 rather than as present values, the grants' ratio is 0.24, $3.9 billion of benefits against $16.3 billion of costs a year (Keiser et al., 2019), and across twenty benefit-cost analyses of US water-quality policy the median ratio is 0.37. Adding the infant health that Flynn and Marcus (2025) find downstream of the plants that had to upgrade raises the measured ratio to as much as 0.45 (Flynn and Smith, 2022). These estimates leave ecological consequences out, as Section 2 notes.

Keiser and Shapiro (2019a) list the reasons their ratio may understate the true one. Among them are nonuse or existence values, since "a person may value a clean river even if they never visit or live near that river", and, among the general equilibrium effects, ecology: "The health of many aquatic species (so indirectly, the benefit people derive from a river) may depend nonlinearly on the area of clean water" (Keiser and Shapiro, 2019a).[53] They conclude that the grants' benefits exceed their costs "if these unmeasured components of willingness to pay are three or more times the components of willingness to pay that we measure" (Keiser and Shapiro, 2019a). Ecology and nonuse values are therefore categories of their own, outside the measured ratio rather than corrections to it.

The birds downstream of the plants are one item of that list, and this paper quantifies them in their own units without monetizing them, as federal guidance asks for an effect that can be quantified but not monetized (Office of Management and Budget, 2023a, p. 46).[54] Across all birds, the number of species recorded per survey rose about 2% after commissioning, and the increase concentrates among the water-linked species, whose number rose about 5%, about one more species at the average survey and one more every two surveys at the median. With the extreme counts dropped, the number of water-linked birds rose about 14%, about 830 more at the average survey and about 85 at the median, while the full counts give an estimate whose interval spans zero. Compared with the counterfactual scenario in which no plant upstream of a circle is commissioned with a grant, the surveys recorded about 3% more water-linked species from 1967 to 2024 (95% interval 1 to 5), about 870 more entries of such a species on the surveys' lists across all circles each year. Table 5

[52] Their costs are discounted at 7.85% a year and are in 2014 dollars, for grants given from 1962 to 2000. The ratio is 0.06 for owned homes within one mile of the river and 0.22 when rental units are added, and against federal capital costs alone it is 0.8 to 0.9 (Keiser and Shapiro, 2019a). Dividing the ratios by the 45 cents of each grant dollar that reached municipal sewerage capital puts the return on the dollars the plants received at about one, with an interval that rejects a ratio below 0.62 (Flynn and Smith, 2022).

[53] The others are incomplete information about water pollution, local fees that may have lowered home values, a radius of 25 miles that misses the longest recreational trips, and pass-through estimates from 198 cities. Commercial fisheries, industrial water supplies and drinking water are further benefits that housing values may not measure (Keiser and Shapiro, 2019a).

[54] The guidance is the 2023 revision of Circular A-4, which Executive Order 14192 of January 31, 2025 (sec. 6(b), 90 Federal Register 9065) ordered revoked, with the 2003 version reinstated. I follow the 2023 revision for its practice on effects that are not monetized and for its discount rate of 2.0%, with the rates of 3 and 7% of the 2003 version as alternatives.

sets these beside the measured ratios, the costs, and the categories left unmeasured, without summing them, since adding monetized and non-monetized effects risks double-counting.

Beside discounted net benefits, the circular on federal programs encourages other summary effectiveness measures, its example being "the number of injuries prevented per dollar of cost" (Office of Management and Budget, 2023b, p. 5). The counterpart here is the federal grant dollars a year per additional species entry and per additional counted bird. Charging the whole grant to the birds makes it an upper bound on what an entry or a bird cost, and never a benefit per dollar. Charged that way, the cost per entry is also the value per entry at which the birds alone would repay the grants, the threshold the regulatory guidance asks for when benefits are not monetized: "How small could the value of the non-monetized benefits be ... before the regulation would yield zero net benefits" (Office of Management and Budget, 2023a, p. 47). The three-or-more-times threshold of Keiser and Shapiro (2019a) is their own, computed on their costs at their discount rate of 7.85%, so I form no ratio from their costs and the costs behind this paper's figures.

6.2 Limitations

These findings have three kinds of limits. The first concerns internal validity. Richness's pre-treatment coefficients slope upward before commissioning. Thus, the post-period mean partly continues a trend already under way: subtracting that pre-trend leaves about half of the water-linked species' rise, 0.050 log points (SE 0.026) of 0.092 (SE 0.017). Where next year's survey effort follows this year's richness, I report the richness pre-trend test without reading it as evidence (Section 4.2.2). The ways of separating the plants' effect through the water from their other channels, the structures a plant adds and the warm effluent it discharges, disagree. Thus, the data do not identify how much of the effect operates through water quality (Section 5.2). And the comparison with distant untreated circles tests spillovers near the treated circles, not a draw of birds across a whole flyway.

The second concerns external validity. The estimates describe winter surveys at the circles that volunteers chose to survey, and the circles are not placed by a sampling design: their locations "tend to be biased toward areas of high human population density and areas of high bird density and diversity" (Meehan et al., 2019).[55] They describe plants commissioned from 1967 to 1999 under a program whose grants ended with fiscal year 1990, when capitalization grants to State revolving funds replaced them, and birds within 25 river miles downstream of a plant.

The third concerns measurement. Birds counted are not a population: each survey is

[55] Analyses of the CBC "are limited by non-random selection of circle locations and the lack of standardization in counting methods" (Link et al., 2006).

Table 5: Benefits and costs of the construction grants: measured in dollars, quantified in birds, and not quantified

	Monetized, billions of 2014 dollars		**Non-monetized**, water-linked species		
	Housing values	Infant health	Birds per survey	Species per survey	Entries a year
Benefits average, median	89	in the ratio	+14% 830 and 85	+5% 1 and 1 every 2	+3% 870 against no grants
Costs	341 present value, 650 to 870 over the projects' lives				
Benefits over costs	0.24 to 0.26	up to 0.45, with housing			
Source	Keiser and Shapiro (2019a), Keiser et al. (2019)	Flynn and Smith (2022)	This paper		
	Not quantified				
	Ecosystems beyond the birds	Nonuse values	Recreation beyond 25 miles	Fisheries and water supplies	

Notes: The housing values, the costs of 341, and the ratio of 0.26 are present values in billions of 2014 dollars at 7.85% over a 30-year life, from Table VI, column 2, of Keiser and Shapiro (2019a). The ratio of 0.24 is annual flows over 1962 to 2001 (Keiser et al., 2019). The lifetime totals are what the grant-funded projects cost, in 2014 dollars (Keiser and Shapiro, 2019a) and in 2017 dollars (Keiser and Shapiro, 2019b). The health ratio adds the infant health that Flynn and Marcus (2025) find downstream of the plants that had to upgrade (Flynn and Smith, 2022). In the non-monetized columns, each percent is $100(e^b - 1)$ of the post-period mean b of the main specification for the water-linked species pooled, on the surveys of the treated circles in the nine years before commissioning (18 species and 6,021 birds at the average survey, 13 and 620 at the median), with the extreme counts dropped, the records above the 97.5th percentile of the year's distinct count values. On the full counts the birds' estimate is −17% with an interval that spans zero. An entry is one species on one survey's list, summed over all circles each year from 1967 to 2024 against the counterfactual scenario in which no plant upstream of a circle is commissioned with a grant. The monetized and non-monetized columns are not summed, since federal guidance warns against double-counting when the two kinds of effect are mixed (Office of Management and Budget, 2023a, p. 46).

one day's count within a circle 15 miles across and carries its observers' effort, and the continental figures of Sections 4.2.1 and 5.2 treat the surveyed circles as standing for the continent. Thus, they overstate the continental change unless the birds at the circles not surveyed changed as those at the surveyed circles did and a local increase is an increase in the population. An entry of a species on a survey's list is not a distinct species. The dose measures a plant's design flow, its capacity at the end of its assumed life, against the river's mean annual flow, not the effluent the plant discharged in a given year, and water quality is read only where a station sampled, in fewer circle-years than the birds were surveyed.

6.3 What Future Work Could Ask

Four questions follow from these estimates and lie beyond them. The first is which species increase and which decrease, and why. The rise in richness here is concentrated among the species tied to water, and whether the composition of a community, by what its species eat and where they forage, moves where its size and evenness do not is a question of its own.

The second is what the increases are worth to the people who watch birds, as revealed by the trips they take and the sites they choose. Where a benefit cannot yet be monetized, federal guidance finds it "helpful to outline the data collection or analysis that would enable quantification or monetization" (Office of Management and Budget, 2023a, p. 45), and birders' trips to the sites downstream of the plants are such data. The third is which plants and places produced the largest gains, so that a dollar of the revolving funds that replaced the grants could be placed where it does the most for birds. The fourth is whether the industrial dischargers, which the Act regulated by permit rather than by grant, changed the birds below their outfalls as the grant-funded plants did.

7 Conclusion

North America's birds have declined for half a century, and valuing an environmental policy for the ecosystem services it provides starts from measuring what the policy did for them. This paper estimates how the treatment plants built with the construction grants of the Clean Water Act changed the birds counted downstream of them, using the staggered commissioning of the plants over three decades and the winter surveys of volunteer birders.

After a plant was commissioned, the surveys downstream recorded more species, while the diversity of the community did not change. The increase concentrated among the species tied to water. Their numbers also rose once the largest flocks are set aside, and the grazers of submerged plants grew in number with those flocks kept as well. These increases did not

scale with how much of the river a plant's effluent made up. The water itself grew cleaner, with more dissolved oxygen and less phosphorus downstream of the plants.

Part of the rise in species continues a trend under way before the plants, and the data do not yet separate the effect of cleaner water from the plants' other channels. Still, the plants brought a measurable increase in the birds downstream, and because it followed the plants rather than their dilution, a valuation that judges a treatment program's ecological return by the water alone would miss it. What the increase is worth to the people who watch birds, and which plants and places produced it, are questions for future work.

A Identification Results

This appendix states and proves the two results on which the estimates rest: what the coefficients of equation (1) identify, and how the dose's effect of Section 5 decomposes into a direct effect and an effect through water quality.

A.1 Event-Study Coefficients

Write $\tau_k(D_{cs} : s \le t)$ for the proportional effect of commissioning on a treated circle in year t, k years after its first commissioning: a function of its event time and of its dose path D_{cs} up to year t, otherwise unrestricted, and zero before the first commissioning. Equation (1) fits one coefficient per event time to every treated circle, whatever its dose, and however many plants followed the first, and Proposition A.1 gives that coefficient's value. Its four conditions are the assumptions of Section 4.1. The conditional mean assumption, with the year effects of equation (1) in the index, is the index version of parallel trends: a common growth factor for every circle's expected count without commissioning. Strict exogeneity of the dose path and staggered adoption with no anticipation are the no anticipation and strict exogeneity of timing stated there, and the fourth condition, an observation window mean-independent of the counts, allows for an unbalanced panel.

Proposition A.1 (the event-study coefficients)**.** *Let the expected count of circle c in year t follow the fixed-effects Poisson index of equation (1), and let the index carry one coefficient per event time, $\tau_k(D_{cs} : s \le t) = \beta_k$ for every treated circle at $k \ge 0$, with the pre-treatment coefficients β_k, $k < 0$, on the indicators of the years before the first commissioning, which is the restriction that equation (1) imposes. Assume the conditional mean assumption of Wooldridge (1999) (CM), strict exogeneity of the dose path conditional on the circle effect (Wooldridge, 1997) (SE-D), staggered adoption with no anticipation (Callaway et al., 2025) (STAG-T), and an observation window mean-independent of the counts given the circle's effect and covariates (OBS). Then:*

(a) (the value for one cohort) *with the index saturated by cohort and event time, the coefficient of cohort a at event time k is the logarithm of the abundance-weighted mean, over the cohort's circles, of their proportional effects under their own paths,*

$$\exp\left(\beta^0_{ak}\right) = \frac{\sum_{c:G_c=a} \phi_c \overbrace{\exp\left(\tau_k(D_{cs} : s \le a+k)\right)}^{\text{each circle's response to its commissioning, through every channel}}}{\underbrace{\sum_{c:G_c=a} \phi_c}_{\text{the cohort's abundance}}}, \tag{A.1}$$

identified under these four conditions alone. The coefficient is the ratio-of-means version of the dose-aggregated event-study parameter of Callaway et al. (2025) at event time $e = k$, the object that the remark in Appendix C of their paper says a binarized event study targets.

(b) (the pooled coefficient) *the single β_k of equation (1) across cohorts equals the common value of equation (A.1) when that value does not vary with a, and otherwise, to first order, a weighted average of the cohorts' values with the weights the proof gives.*

Proof. To show part (a), write the index of equation (1) saturated by cohort and event time as $x'_{ct}\boldsymbol{\beta}$, a year effect λ_t and one coefficient β_{aj} for each cohort a and event time $j \geq 0$, with the covariates set aside. Write τ_{cs} for circle c's proportional effect in year s, $\tau_{s-a}(D_{cs'} : s' \leq s)$ for $s \geq a$ and zero before, so that under the conditional mean assumption the expected count is $\mu_{cs} = \phi_c \exp(\lambda_s + \tau_{cs})$, with $n_c = \sum_s \mu_{cs}$ the circle's expected total. The fixed-effects Poisson conditions out ϕ_c, and its population first-order condition for each column of the index is equation (3.4) of Wooldridge (1999),

$$\sum_c \sum_t x_{ct}\Big(\mu_{ct} - p_{ct}(\boldsymbol{\beta})\, n_c\Big) = 0, \qquad p_{ct}(\boldsymbol{\beta}) = \frac{\exp(x'_{ct}\boldsymbol{\beta})}{\sum_s \exp(x'_{cs}\boldsymbol{\beta})}, \tag{A.2}$$

where p_{ct} is the share of the circle's expected total that the index places in year t. Every circle of cohort a carries the same index in every year. Thus, the denominator of p_{ct} is one value for the cohort, $Q_a(\boldsymbol{\beta}) \equiv \sum_{s<a} e^{\lambda_s} + \sum_{s \geq a} e^{\lambda_s + \beta_{a,s-a}}$, and for the column of cohort a at event time k, the indicator $\mathbf{1}(G_c = a)\mathbf{1}(t = a + k)$, equation (A.2) sets the cohort's expected total in year $a + k$ equal to its fitted total:

$$\sum_{c:G_c=a} \phi_c\, e^{\lambda_{a+k} + \tau_{c,a+k}} = \frac{e^{\lambda_{a+k} + \beta_{ak}}}{Q_a(\boldsymbol{\beta})} \sum_{c:G_c=a} n_c, \qquad \sum_{c:G_c=a} n_c = \sum_{c:G_c=a} \phi_c \Big(\sum_{s<a} e^{\lambda_s} + \sum_{s \geq a} e^{\lambda_s + \tau_{cs}} \Big), \tag{A.3}$$

where the second display writes out the cohort's expected total with $\tau_{cs} = 0$ before commissioning. The year-effect columns' conditions, less the cell conditions of the cohorts treated by year t, involve the circles with no commissioned plant in year t alone. Thus, λ_t is the growth of those circles' expected counts, which no anticipation and strict exogeneity of the dose path make the treated circles' counterfactual growth. Now substitute the candidate $e^{\beta_{aj}} = \sum_{c:G_c=a} \phi_c e^{\tau_{c,a+j}} / \sum_{c:G_c=a} \phi_c$ for every $j \geq 0$ into Q_a:

$$Q_a(\boldsymbol{\beta}) \sum_{c:G_c=a} \phi_c = \sum_{c:G_c=a} \phi_c \sum_{s<a} e^{\lambda_s} + \sum_{s \geq a} e^{\lambda_s} \sum_{c:G_c=a} \phi_c e^{\tau_{cs}} = \sum_{c:G_c=a} n_c, \tag{A.4}$$

by equation (A.3)'s second display. Thus, the right side of equation (A.3)'s first display equals $e^{\lambda_{a+k} + \beta_{ak}} \sum_c \phi_c = e^{\lambda_{a+k}} \sum_c \phi_c e^{\tau_{c,a+k}}$, which is its left side. The candidate solves every cell condition of the cohort at once, it is the unique solution because the quasi-log-

likelihood is concave in $\boldsymbol{\beta}$ (Wooldridge, 1999), and at $j = k$ it is equation (A.1). The dose-aggregated event-study parameter of Callaway et al. (2025) averages each treated unit's effect at event time e over the units treated e periods earlier, and on the ratio scale that average is abundance-weighted, which gives the reading.

To show part (b), note that equation (1) fits one column per event time across cohorts. When the cohort values of equation (A.1) are equal, the common value solves every cohort's first-order condition at once, and the pooled regression returns it exactly, by the partition of the first-order conditions applied cohort by cohort. When they differ, the index is misspecified, and the limit of the regression is the projection of the true index on the model's columns, weighted by the zero-effect means $\mu^{00}_{ct} = \phi_c \exp(\lambda_t)$ and with the circle effects partialled by weighted demeaning,

$$\beta^*_k \approx \sum_a \omega_{ak} \beta^0_{ak}, \qquad \omega_{ak} \propto \sum_{c:G_c=a} \mu^{00}_{c,a+k} \tilde{D}_{c,a+k} D_{c,a+k}, \tag{A.5}$$

where $D_{c,a+k}$ is the event-time column and $\tilde{D}_{c,a+k}$ that column less its μ^{00}-weighted mean over the circle's years. The display is the first-order expansion of the same population score, equation (3.4) of Wooldridge (1999), around the zero-effect means, with the weights summed within each cohort, and no assumption beyond the four enters. □

In words, the coefficient at event time k is the abundance-weighted average, over the circles treated k years earlier, of their proportional responses to commissioning through every channel it opens, and not the effect at any one dose or through any one channel (part (a)). Across cohorts, one coefficient is exact when the cohorts respond alike and a weighted average otherwise (part (b)). The comparison of Section 4.1 follows from the same first-order conditions: the year effects are learned from the zero-dose cells of each year, which are the never-treated circles and those not yet treated. The baseline-by-year terms $\mathcal{T}_{ct}$ of equation (1) interact the circle's housing density, its cropland share, the plants the grants did not fund, and the grant plants' prior treatment level with year indicators, as Section 4.1 defines them. With those terms the year effects are learned from the zero-dose circles at every value of the baselines the data hold, and the identifying condition is parallel trends conditional on the baselines and linear in them on the index.

A.2 Direct Effect and the Effect Through Water Quality

Write M_{ct} for the analyte's reading at circle c in year t, on the within-circle standard score of Section 5.1, oriented so that more is cleaner. Proposition A.2 writes the reduced-form slope of equation (6) as the sum of the direct effect and the effect through water quality,

establishes that the data cannot split the sum, and bounds the second term under one sign restriction.

Proposition A.2 (the decomposition of the dose's effect, and the bound)**.** *Assume the four conditions of Proposition A.1 for the count's index (CM, SE-D, STAG-T, and OBS), and parallel trends for the reading as an outcome of its own. Within one half ς of the split by the plants' effluent before their grants, assume the two structural forms, one for the reading and one for the count,*

$$M_{ct} = \alpha_c^M + \lambda_t^M + a_\varsigma D_{ct}, \qquad \mathrm{E}\Big(N_{ct} \mid \cdot\Big) = \phi_c \exp\Big(\lambda_t + \beta_D^{\mathrm{dir}} D_{ct} + \beta_M^{\mathrm{wq}} M_{ct}\Big), \tag{A.6}$$

with the reading oriented so that more is cleaner, its slope $a_\varsigma \equiv \beta_\varsigma^M$ positive, which is monotonicity (MONO), and no variation in the reading beyond its circle and year effects and the dose. Then:

(i) (the decomposition) *the count's reduced form on the dose has the slope of equation (6),*

$$\tau_\varsigma = \underbrace{\beta_D^{\mathrm{dir}}}_{\text{the direct effect}} + \underbrace{\beta_M^{\mathrm{wq}} a_\varsigma}_{\text{the effect through the water-quality channel}},$$

and both τ_ς and a_ς are identified. Write w_ς for the second term.

(ii) (the split is not identified) *every pair of slopes with the same sum generates the same joint law of the dose, the reading, and the expected counts, once the circle effects and the year effects are adjusted.*

(iii) (the bound under the sign restriction) *if the direct effect and the effect through the water-quality channel each carry the sign of τ_ς (SIGN), the identified set for w_ς is the closed interval between 0 and τ_ς, and for the birds' response to water quality it is the closed interval between 0 and $\tau_\varsigma / a_\varsigma$. Both sets are sharp, and the ratio $\tau_\varsigma / a_\varsigma$ equals the birds' response to water quality exactly when the direct effect is zero.*

(iv) (the proportional scale) *at dose d the proportional effect is the exponential of $\tau_\varsigma d$ minus one, and its two parts multiply rather than add,*

$$1 + \tau_\varsigma^{prop}(d) = e^{\beta_D^{\mathrm{dir}} d} \cdot e^{w_\varsigma d}, \tag{A.7}$$

and under the sign restriction the effect through the water-quality channel, the exponential of $w_\varsigma d$ minus one, lies between 0 and $\tau_\varsigma^{prop}(d)$.

Proof. To show part (i), substitute the first form of equation (A.6) into the second:

$$\lambda_t + \beta_D^{\mathrm{dir}} D_{ct} + \beta_M^{\mathrm{wq}} M_{ct} = \underbrace{\Big(\lambda_t + \beta_M^{\mathrm{wq}} \lambda_t^M\Big)}_{\text{a year effect}} + \underbrace{\beta_M^{\mathrm{wq}} \alpha_c^M}_{\text{a circle effect}} + \Big(\beta_D^{\mathrm{dir}} + \beta_M^{\mathrm{wq}} a_\varsigma\Big) D_{ct}, \tag{A.8}$$

where the equality collects terms (algebra). The circle term multiplies ϕ_c and the year term replaces λ_t, so that the conditional mean of the count is the multiplicative form of equation (1) with the coefficient τ_ς of equation (6) on the dose. That coefficient is identified under the four conditions with the circle effect conditioned out (Wooldridge, 1999). The slope a_ς is the treatment effect on water quality, the reading's regression on the dose with circle and year effects, identified under the reading's own parallel trends.

To show part (ii), fix any pair $(\tilde{\beta}^{\text{dir}}, \tilde{\beta}^{\text{wq}})$ with $\tilde{\beta}^{\text{dir}} + \tilde{\beta}^{\text{wq}} a_\varsigma = \tau_\varsigma$, and adjust the effects by the difference of the water slopes, $\Delta \equiv \beta_M^{\text{wq}} - \tilde{\beta}^{\text{wq}}$:

$$\tilde{\phi}_c = \phi_c e^{\Delta \alpha_c^M}, \qquad \tilde{\lambda}_t = \lambda_t + \Delta \lambda_t^M. \tag{A.9}$$

The logarithm of the adjusted pair's conditional mean is then

$$\begin{aligned} \ln \tilde{\phi}_c + \tilde{\lambda}_t + \tilde{\beta}^{\text{dir}} D_{ct} + \tilde{\beta}^{\text{wq}} M_{ct} &= \ln \phi_c + \Delta \alpha_c^M + \lambda_t + \Delta \lambda_t^M + \tilde{\beta}^{\text{dir}} D_{ct} \\ &\quad + \tilde{\beta}^{\text{wq}} \left(\alpha_c^M + \lambda_t^M + a_\varsigma D_{ct} \right) \\ &= \ln \phi_c + \beta_M^{\text{wq}} \alpha_c^M + \lambda_t + \beta_M^{\text{wq}} \lambda_t^M + \left(\tilde{\beta}^{\text{dir}} + \tilde{\beta}^{\text{wq}} a_\varsigma \right) D_{ct} \\ &= \ln \phi_c + \beta_M^{\text{wq}} \alpha_c^M + \lambda_t + \beta_M^{\text{wq}} \lambda_t^M + \tau_\varsigma D_{ct}, \end{aligned} \tag{A.10}$$

where the first equality substitutes the adjusted effects and the first form of equation (A.6), the second collects terms ($\Delta + \tilde{\beta}^{\text{wq}} = \beta_M^{\text{wq}}$), and the third uses the pair's sum. The true pair's logarithm, expanded the same way, is the same final line, with the true sum $\beta_D^{\text{dir}} + \beta_M^{\text{wq}} a_\varsigma = \tau_\varsigma$ in its last term. Thus, the expected counts agree at every circle and year. The dose and the reading do not involve either slope, and the joint law of the observables is therefore the same. The hypothesis of no water variation beyond the effects and the dose is what lets the effects absorb Δ.

To show part (iii), take $\tau_\varsigma > 0$. The case $\tau_\varsigma < 0$ follows by reflecting every sign, and $\tau_\varsigma = 0$ forces both parts to zero under the sign restriction. The sign restriction gives $\beta_D^{\text{dir}} \geq 0$ and $w_\varsigma \geq 0$, and equation (6) gives $w_\varsigma = \tau_\varsigma - \beta_D^{\text{dir}} \leq \tau_\varsigma$, so that w_ς lies in the interval from 0 to τ_ς. Dividing by $a_\varsigma > 0$, which monotonicity supplies, puts the water slope in the interval from 0 to $\tau_\varsigma / a_\varsigma$. For sharpness, fix any w^* in the first interval. The pair $(\tau_\varsigma - w^*, w^*/a_\varsigma)$ satisfies the sign restriction and has the sum τ_ς, so that by part (ii) it generates the observed law: every point of the interval is the truth of some world that produces the data. Finally,

$$\frac{\tau_\varsigma}{a_\varsigma} - \beta_M^{\text{wq}} = \frac{\beta_D^{\text{dir}}}{a_\varsigma}, \tag{A.11}$$

by equation (6) divided by a_ς, which is nonnegative under the sign restriction and zero exactly when the direct part is.

To show part (iv), exponentiate equation (6) at a dose step d: $e^{\tau_\varsigma d} = e^{(\beta_D^{\text{dir}} + w_\varsigma)d} = e^{\beta_D^{\text{dir}} d} \cdot e^{w_\varsigma d}$, by the exponential's product rule. The proportional effect is the left side minus one, by its definition as the exponential of the index change minus one, which gives equation (A.7). Under the sign restriction with $\tau_\varsigma > 0$, the direct factor is at least one, so that the water factor is at most $e^{\tau_\varsigma d}$, and $w_\varsigma \geq 0$ makes the water factor at least one. Subtracting one throughout gives the stated interval. □

The sign restriction says that neither channel works against the other. The restriction is weaker than exclusion, which sets the direct effect to zero, and its nearest counterpart is the support restriction on an instrument's direct effect of Conley et al. (2012), here the support from zero to the whole effect with the dose as the instrument. It fails wherever a channel works against the cleaning: construction on a riparian site, a sewer extension that enables development, or a neighboring circle's spillover of the opposite sign. The data do not test it.

The proposition is static, with one dose step and no lags, and it assumes that the reading varies only through its circle and year effects and the dose. With heterogeneous slopes the bound holds for each cohort and event time of the saturated fit, whose weights are nonnegative, and a pooled coefficient whose weights can be negative can leave it.

B Additional Tables and Figures

B.1 Tables

Table B1: What a treatment plant changes in the water, and how each change reaches the birds downstream

Constituent	Change in the water	Change in the birds' food	Response of the birds
Oxygen-demanding material	Oxygen levels increase	Invertebrates recolonize; pollution-sensitive taxa replace first-order opportunists	Birds may benefit
Suspended solids	The water's clarity rises	Submerged vegetation returns; waterfowl diets follow it	The breeding waterbird community grows
Nutrients	Nutrients removed	The food sustained by the outfalls is withdrawn; submerged vegetation returns	Scaup, goldeneye and gulls decline; grazing waterbirds track their food

Table B2: Event-study estimates of the effect of grant-funded commissioning on every species category and statistic, Christmas Bird Count surveys, 1940 to 2024

	Main specification					Extreme counts dropped		Surveys	
	(1) Abundance log points	(2) Richness log points	(3) Effective number of species, SD	(4) Effective number of common species, SD	(5) Entropy SD	(6) Abundance log points	(7) Richness log points	Abundance, richness	Diversity statistics
Panel A. All birds									
All birds	−0.15	0.015**	−0.0083	−0.016	−0.031	0.11**	0.015**	85,934	85,934
	(0.19) [0.07]	(0.0076) [0.36]	(0.028) [0.008]	(0.031) [0.03]	(0.032) [<0.001]	(0.049) [0.05]	(0.0076) [0.38]		
Panel B. The water-linked species and their feeding guilds									
Water-linked species, pooled	−0.18	0.045***	0.0067	0.0042	0.033	0.13**	0.045***	85,534	81,992
	(0.26) [0.06]	(0.015) [0.27]	(0.029) [0.91]	(0.030) [0.89]	(0.029) [0.86]	(0.061) [0.20]	(0.015) [0.29]		
Visual fish-eaters	0.019	0.033*	−0.0091	−0.0070	0.0043	−0.017	0.033*	84,359	74,818
	(0.11) [0.20]	(0.019) [0.97]	(0.029) [0.31]	(0.034) [0.29]	(0.026) [0.18]	(0.11) [0.18]	(0.019) [0.97]		
Marsh invertebrate-eaters	−0.13	0.020	0.017	0.013	0.019	0.091	0.021	83,840	71,818
	(0.33) [0.02]	(0.016) [0.10]	(0.033) [0.80]	(0.036) [0.68]	(0.031) [0.91]	(0.098) [0.10]	(0.016) [0.14]		
Submerged-plant grazers	0.35**	0.060***	−0.011	−0.011	0.021	0.25**	0.060***	82,765	69,163
	(0.14) [0.31]	(0.022) [0.16]	(0.039) [0.73]	(0.041) [0.78]	(0.034) [0.49]	(0.12) [0.33]	(0.022) [0.16]		
Bottom-sediment feeders	−0.50***	0.078***	0.050	0.031	0.050	0.20	0.079***	82,732	65,197
	(0.099) [0.06]	(0.023) [0.52]	(0.035) [0.41]	(0.037) [0.32]	(0.033) [0.34]	(0.14) [0.26]	(0.023) [0.51]		
Crustacean-eaters	0.40**	0.051	−0.018	−0.012	−0.023	0.37**	0.050	28,074	13,809
	(0.18) [<0.001]	(0.038) [0.97]	(0.052) [0.43]	(0.051) [0.60]	(0.052) [0.38]	(0.16) [<0.001]	(0.038) [0.97]		
Scavenging gulls	−0.041	0.063**	0.038	0.042	0.037	−0.043	0.063**	74,434	52,472
	(0.15) [0.65]	(0.026) [0.40]	(0.037) [0.15]	(0.043) [0.27]	(0.039) [0.10]	(0.11) [0.75]	(0.026) [0.40]		
Scavenging crow	0.35	0.13**				0.32	0.13**	22,529	
	(0.24) [0.003]	(0.056) [0.95]				(0.24) [0.003]	(0.056) [0.95]		
Aerial insect-eaters	0.56**	0.020	−0.11*	−0.10	−0.099	0.40***	0.0084	13,834	6,037
	(0.26) [0.002]	(0.074) [0.53]	(0.066) [0.47]	(0.069) [0.65]	(0.061) [0.39]	(0.13) [0.005]	(0.074) [0.51]		
Panel C. The general groups									
Waterbirds	0.19	0.054***	0.0027	0.0090	0.052*	0.11	0.054***	82,090	69,659
	(0.12) [0.48]	(0.019) [0.36]	(0.035) [0.63]	(0.041) [0.60]	(0.030) [0.32]	(0.092) [0.34]	(0.019) [0.37]		
Waterfowl	0.18	0.054***	−0.030	−0.049	−0.012	0.15	0.054***	85,135	79,746
	(0.17) [0.17]	(0.019) [0.49]	(0.033) [0.15]	(0.035) [0.06]	(0.032) [0.12]	(0.11) [0.24]	(0.019) [0.49]		
Shorebirds	−0.10	0.014	0.029	0.039	0.010	−0.11	0.014	70,677	50,082
	(0.12) [0.20]	(0.020) [0.45]	(0.026) [0.51]	(0.029) [0.53]	(0.023) [0.49]	(0.12) [0.08]	(0.020) [0.45]		
Landbirds	−0.0059	0.0072	0.030	0.026	0.019	0.080	0.0073	85,934	85,934
	(0.24) [0.13]	(0.0068) [0.56]	(0.031) [0.46]	(0.032) [0.35]	(0.034) [0.34]	(0.056) [0.03]	(0.0068) [0.58]		
Circle and year effects				Yes					
Effort controls				Yes					
Weather controls				Yes					
Attribute-by-year trends				Yes					

Notes: One event study of equation (1) per category and statistic, as in Table 1: abundance and richness by Poisson quasi-likelihood in log points, and the diversity statistics, standardized within each category, by least squares in standard deviations. Each estimate is the average of the event-time coefficients over years 0 to 14 less their average over years −9 to −1, with its standard error in parentheses and, in brackets, the p-value of the pre-trend test, the Wald test that every pre-treatment coefficient is zero in the regression that omits year −1. ***, **, and * denote significance at the 1, 5, and 10% levels. Extreme counts are dropped for abundance and richness only. The diversity statistics are undefined where a survey records no bird of the category, and the scavenging crow, a single species, has none.

Table B3: Treated circles and surveys at each year relative to the first commissioning of a grant-funded plant upstream, all birds

	All surveys of treated circles			Main specification's sample	
Year relative to first commissioning	Treated circles	Treated surveys	Share of treated surveys (%)	Treated circles	Treated surveys
−10 or earlier	1,504	12,371	16.4	1,392	11,794
−9	740	740	1	728	728
−8	778	778	1	767	767
−7	792	792	1	781	781
−6	833	833	1.1	820	820
−5	853	853	1.1	839	839
−4	876	876	1.2	861	861
−3	884	884	1.2	869	869
−2	902	902	1.2	889	889
−1	914	914	1.2	899	899
0	934	934	1.2	918	918
1	944	944	1.3	929	929
2	974	974	1.3	959	959
3	988	988	1.3	973	973
4	1,004	1,004	1.3	988	988
5	1,015	1,015	1.3	1,000	1,000
6	1,042	1,042	1.4	1,026	1,026
7	1,048	1,048	1.4	1,031	1,031
8	1,060	1,060	1.4	1,043	1,043
9	1,074	1,074	1.4	1,053	1,053
10	1,108	1,108	1.5	1,086	1,086
11	1,105	1,105	1.5	1,084	1,084
12	1,127	1,127	1.5	1,105	1,105
13	1,136	1,136	1.5	1,114	1,114
14	1,144	1,144	1.5	1,124	1,124
15 or later	1,753	39,839	52.8	1,721	37,510

Notes: The first three columns count the treated circles and their surveys in the panel at each year relative to the first commissioning, and the share of the treated circles' surveys that the year holds. The last two count them in the main specification's estimation sample. The first and last rows count every year 10 or more before and 15 or more after the first commissioning, which the endpoints pool.

Table B4: Change in the event study's post-period mean when two sets of circles are dropped or kept, by statistic, in standard errors of the main estimate

	Abundance	Richness	Effective number of species	Effective number of common species	Entropy	Circles changed
Panel A. Treated circles dosed only by plants with no surface discharge dropped						
All birds	−0.12	−0.25	−0.48	−0.55	−0.24	190
Water-linked species, pooled	−0.16	−0.17	−0.27	−0.37	−0.20	190
Visual fish-eaters	0.58	−0.01	−0.25	−0.17	−0.22	190
Marsh invertebrate-eaters	−0.15	−0.09	0.33	0.32	0.10	190
Submerged-plant grazers	0.08	−0.20	−0.20	−0.12	−0.18	190
Bottom-sediment feeders	0.00	−0.34	−0.53	−0.41	−0.56	190
Crustacean-eaters	0.29	−0.06	0.05	0.05	0.11	145
Scavenging gulls	0.19	−0.15	−0.45	−0.36	−0.50	190
Scavenging crow	−0.31	0.41				27
Aerial insect-eaters	0.15	−0.16	−0.05	−0.06	−0.02	135
Waterbirds	0.19	−0.03	−0.10	−0.06	−0.04	190
Waterfowl	0.15	−0.33	−0.60	−0.53	−0.48	190
Shorebirds	−0.07	0.00	0.42	0.42	0.06	189
Landbirds	−0.12	−0.18	−0.18	−0.29	0.00	190
Panel B. Never-treated circles holding a plant whose path misses their outlet kept						
All birds	−0.18	0.35	0.36	0.22	0.48	29
Water-linked species, pooled	−0.11	0.27	−0.07	−0.04	0.00	29
Visual fish-eaters	0.12	0.43	−0.31	−0.44	−0.04	29
Marsh invertebrate-eaters	−0.14	0.27	−0.17	−0.21	−0.04	29
Submerged-plant grazers	−0.20	−0.02	−0.07	−0.14	−0.09	29
Bottom-sediment feeders	0.61	0.33	0.18	0.21	0.15	29
Crustacean-eaters	−0.40	0.33	0.36	0.33	0.37	21
Scavenging gulls	0.12	−0.19	−0.70	−0.64	−0.80	29
Scavenging crow	−0.26	0.30				10
Aerial insect-eaters	0.05	0.31	0.13	0.12	0.17	16
Waterbirds	−0.21	0.11	−0.25	−0.34	−0.14	29
Waterfowl	−0.31	0.27	0.30	0.34	0.28	29
Shorebirds	−0.08	0.28	0.08	−0.03	0.32	29
Landbirds	−0.17	0.39	0.14	0.04	0.07	29

Notes: Each cell is the estimate on the changed sample less the main estimate of Appendix Table B2, divided by the main estimate's standard error. The last column counts the treated circles dropped in Panel A and the never-treated circles added in Panel B, before the circles with a single survey leave the regression. The rows follow the order of the paper's groups. The scavenging crow, a single species, has no diversity statistics. Over the 67 cells, the largest change in absolute value is 0.60 standard errors in Panel A and 0.80 in Panel B, and the medians are 0.18 and 0.21.

Table B5: Event-study estimates of the effect of commissioning on whether a circle-year holds a survey and whether a group's diversity statistics are defined

	Share of all circle-years (%)	Pre-trend test p-value, years −9 to −2	Pre-trend test p-value, every pre-treatment coefficient	Effect, years 0 to 14, percentage points	Standard error
A survey taken	23.4	0.22	<0.001	−0.76	(0.76)
Diversity statistics defined: all birds	23.4	0.22	<0.001	−0.76	(0.76)
waterbirds	18.5	0.06	<0.001	2.7	(0.67)
waterfowl	21.3	0.07	<0.001	1.2	(0.72)
shorebirds	12.5	0.22	<0.001	2.2	(0.58)
landbirds	23.4	0.22	<0.001	−0.75	(0.76)

Notes: Each row is an event study, on the full grid of 443,080 circle-years (every circle of the panel in every year, with zeros before a circle's first survey and after its last), of an indicator that the circle-year enters the sample named, with circle and year effects. The effect is the average of the event-time coefficients over years 0 to 14 less their average over years −9 to −1, in percentage points, with its standard error clustered by circle in parentheses. The first pre-trend test covers the pre-treatment coefficients of years −9 to −2 and the second every one of them, the pooled endpoint included. The second rejects on the far coefficients, which record when circles are first surveyed. A group's diversity statistics are defined only where the group is present. Thus, its row measures the group's presence as well as the circle's participation.

Table B6: Joint tests that the event-study coefficients of the balanced sample of circles equal those of the full sample, by species group and statistic

	Joint test p-value, the 25 event-time coefficients					Abundance, Balanced circles		
	Abundance	Richness	Effective number of species	Effective number of common species	Entropy	difference	All	Treated
Panel A. All birds								
All birds	<0.001	0.37	0.21	0.44	0.07	0.21	969	366
						(0.18)		
Panel B. The water-linked species and their feeding guilds								
Water-linked species, pooled	<0.001	0.42	0.50	0.75	0.70	0.19	946	366
						(0.23)		
Visual fish-eaters	0.34	0.28	0.88	0.89	0.83	0.097	901	364
						(0.081)		
Marsh invertebrate-eaters	<0.001	0.30	0.63	0.61	0.40	0.47	898	365
						(0.33)		
Submerged-plant grazers	<0.001	0.87	0.92	0.92	0.95	−0.18	861	361
						(0.12)		
Bottom-sediment feeders	0.32	0.13	0.72	0.91	0.59	−0.015	855	363
						(0.089)		
Crustacean-eaters	<0.001	0.86	0.15	0.09	0.23	−0.0074	308	120
						(0.14)		
Scavenging gulls	0.06	0.92	0.95	0.92	0.94	−0.0057	682	349
						(0.060)		
Scavenging crow	0.85	0.75				−0.19	224	127
						(0.13)		
Aerial insect-eaters	0.003	0.87	0.97	0.97	0.97	0.12	164	60
						(0.23)		
Panel C. The general groups								
Waterbirds	0.85	0.31	0.50	0.31	0.65	−0.063	849	360
						(0.051)		
Waterfowl	0.28	0.73	0.19	0.46	0.06	−0.15	924	366
						(0.10)		
Shorebirds	0.003	0.65	0.49	0.64	0.57	0.077	752	318
						(0.082)		
Landbirds	<0.001	0.51	0.40	0.31	0.33	0.23	969	366
						(0.24)		
Circle and year effects	Yes	Yes	Yes	Yes	Yes			
Effort controls	Yes	Yes	Yes	Yes	Yes			
Weather controls	Yes	Yes	Yes	Yes	Yes			
Attribute-by-year trends	Yes	Yes	Yes	Yes	Yes			

Notes: The balanced sample keeps the treated circles surveyed in every year from ten before the first commissioning to fifteen after, and every never-treated circle. Each p-value is the joint Wald test that the 25 event-time coefficients of the main specification fitted on the balanced sample equal those fitted on the full sample, with the covariance from the two regressions' stacked scores, clustered by circle. The abundance column gives the difference in the post-period mean, the balanced sample's less the full sample's, in log points, with its standard error in parentheses. For all birds a circle bootstrap of 19 resamples gives a standard error of 0.140. The regressions carry the circle and year effects and the controls of Table 1. The rows follow the order of the paper's groups, and the scavenging crow, a single species, has no diversity statistics.

Table B7: Pre-treatment characteristics of the circles in the estimation sample by treatment group, as standardized differences from the never-treated circles

Characteristic	Never treated, mean	Never treated, standard deviation	All treated, mean	All treated	Not yet treated, by cohort	Treated, no measured dose	Dose quartile 1	Dose quartile 2	Dose quartile 3	Dose quartile 4
Abundance, log	7.8	1.6	8.4	0.41	0.069	0.35	0.28	0.30	0.44	0.60
Richness, species	49	30	54	0.16	-0.029	0.44	0.032	-0.024	0.12	0.33
Effective number of species	12	6.7	12	-0.047	-0.10	0.41	-0.14	-0.13	-0.12	0.013
Party-hours, log	3.1	0.65	3.4	0.46	0.067	0.52	0.18	0.39	0.53	0.66
Field counters, log	2.0	0.74	2.4	0.53	0.083	0.49	0.23	0.47	0.62	0.78
Winter temperature, degrees Celsius	1.1	8.0	-0.058	-0.16	-0.093	0.053	-0.23	-0.33	-0.21	0.019
Winter precipitation, millimeters	115	79	152	0.46	0.047	0.37	0.48	0.45	0.49	0.47
Housing density, pre-treatment	18	50	237	0.36	0.14	0.51	0.46	0.44	0.77	0.52
Cropland share, pre-treatment	0.12	0.15	0.22	0.58	-0.011	0.34	0.58	0.54	0.68	0.59
Land in farms, percent, pre-treatment	51	29	47	-0.14	-0.0063	-0.23	-0.047	-0.14	-0.11	-0.23
Latitude, degrees	38	5.7	39	0.35	0.15	0.20	0.38	0.45	0.37	0.24
Largest river's mean annual flow, log	5.4	2.6	7.0	0.69	0.12	0.16	1.5	0.85	0.54	0.33
Largest river's drainage area, log	7.4	2.3	8.3	0.44	0.10	0.016	1.2	0.57	0.22	0.093
Biochemical oxygen demand, log	0.90	0.83	1.1	0.26	0.097	0.43	-0.093	0.078	0.31	0.59
Total suspended solids, log	3.7	1.5	3.7	0.0036	0.028	0.012	-0.029	-0.082	-0.066	0.15
Fecal coliform, log	5.9	2.2	6.8	0.40	0.060	0.0085	0.21	0.37	0.48	0.61
Dissolved oxygen, percent saturation	92	11	89	-0.21	-0.0091	-0.18	-0.19	-0.11	-0.032	-0.66
Circles	284		1481	1481	1479	111	317	332	338	383
Pre-treatment surveys	2681		19002	19002	18976	1191	3597	4099	4707	5408

Notes: The first three columns give the never-treated circles' mean and standard deviation and the treated circles' mean. Every other column gives a standardized difference: the group's mean minus the comparison circles' mean, over the square root of the mean of the two variances. The comparison circles are the never-treated circles, except in the not-yet-treated column, which compares each of 26 commissioning cohorts with the 1,453 circles first commissioned later, over the years before the cohort's commissioning, each cohort weighted by its size. The bird statistics, effort, and climate are averaged over each circle's surveys before its first commissioning, and for the never-treated circles over the surveys before 1980, the treated circles' median commissioning year. The water rows average the annual means of the circle's eight-digit watershed over the same years, in logarithms for oxygen demand, suspended solids, and fecal coliform. Land use is read at the last census before each treated circle's commissioning, and before 1972 for the never-treated circles. The dose quartiles are of the dose in force in 2024 among the treated circles with a measured dose. The sample is the main specification's estimation sample for all birds' abundance, 85,934 surveys on 2,975 circles, and the circle rows count the circles with a survey in each window.

Table B8: Baseline water quality of the circles' watersheds before commissioning by treatment group, in physical units

Characteristic	Never treated	All treated	Treated, no measured dose	Dose quartile 1	Dose quartile 2	Dose quartile 3	Dose quartile 4
Biochemical oxygen demand, median, milligrams per liter	2.2	2.8	3.0	2.3	2.5	2.9	3.6
Biochemical oxygen demand, share of circles with a value	0.53	0.66	0.49	0.70	0.64	0.68	0.67
Total suspended solids, median, milligrams per liter	28	32	32	32	31	32	38
Total suspended solids, share of circles with a value	0.59	0.69	0.59	0.74	0.69	0.72	0.67
Fecal coliform, median, most probable number per 100 milliliters	443	1050	601	648	1001	1200	1668
Fecal coliform, share of circles with a value	0.65	0.78	0.72	0.80	0.74	0.81	0.81
Dissolved oxygen, median, percent saturation	92	87	87	88	90	88	85
Dissolved oxygen, share of circles with a value	0.18	0.21	0.23	0.26	0.20	0.20	0.19
Circles	284	1481	111	317	332	338	383

Notes: Each value is the median across the circles of the group. A circle's value is the mean of the annual means of its eight-digit watershed before its first commissioning, and for a never-treated circle before 1980. The share counts the circles of each group whose watershed was measured in those years. Eight watershed means of zero enter the medians here and are left out of the logarithms of Table B7. The sample and the groups are those of Table B7.

Table B9: Event-study estimates of the effect of grant-funded commissioning on the three diversity statistics downstream, by species group

	(1) All birds	(2) Water-linked species	(3) Waterbirds	(4) Waterfowl	(5) Shorebirds	(6) Landbirds
Panel A. Least squares, in standard deviations of the statistic over the group's surveys						
Effective number of species	−0.0083	0.0067	0.0027	−0.030	0.029	0.030
	(0.028)	(0.029)	(0.035)	(0.033)	(0.026)	(0.031)
Effective number of common species	−0.016	0.0042	0.0090	−0.049	0.039	0.026
	(0.031)	(0.030)	(0.041)	(0.035)	(0.029)	(0.032)
Entropy	−0.031	0.033	0.052*	−0.012	0.010	0.019
	(0.032)	(0.029)	(0.030)	(0.032)	(0.023)	(0.034)
Surveys	85,934	81,992	69,659	79,746	50,082	85,934
Panel B. Pre-trend test p-values						
Effective number of species	0.008	0.91	0.63	0.15	0.51	0.46
Effective number of common species	0.03	0.89	0.60	0.06	0.53	0.35
Entropy	<0.001	0.86	0.32	0.12	0.49	0.34
Panel C. The effects in the statistic's own units						
Effective number of species, in species	−0.075	0.029	0.0065	−0.068	0.059	0.20
	(0.26)	(0.12)	(0.084)	(0.075)	(0.054)	(0.20)
Effective number of common species, in species	−0.088	0.012	0.014	−0.082	0.057	0.12
	(0.17)	(0.083)	(0.065)	(0.058)	(0.043)	(0.14)
Entropy, in nats	−0.020	0.025	0.034*	−0.0074	0.0066	0.012
	(0.021)	(0.022)	(0.020)	(0.020)	(0.014)	(0.021)
Circle and year effects	Yes	Yes	Yes	Yes	Yes	Yes
Effort controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
Attribute-by-year trends	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each estimate is the average of the event-study coefficients of equation (1) over years 0 to 14 after commissioning, less their average over years −9 to −1, by least squares on each statistic standardized within its group, in standard deviations of the statistic over the group's surveys. The regressions carry the circle and year effects and the controls of Table 1. Standard errors clustered by circle are in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels. The diversity statistics are undefined in a survey that records no bird of the group. Thus, the surveys are fewer than in Table 1. Panel B gives the p-value of the Wald test that every pre-treatment coefficient, the pooled endpoint included, is zero in the regression that omits year −1. Panel C multiplies each estimate and its standard error by the standard deviation of the statistic over the group's surveys: the two effective numbers are in species and the entropy in nats.

Table B10: Distributed-lag estimates of the effect of the dose and of the number of commissioned grant plants upstream on the three diversity statistics, by species group

	Effective number of species		Effective number of common species		Entropy	
	(1)	(2)	(3)	(4)	(5)	(6)
	Dose step	Plant added	Dose step	Plant added	Dose step	Plant added
	Response at event time 14, standard deviations					
Panel A. All birds						
All birds	0.011***	−0.010	0.015***	−0.011	0.0057	−0.0099
	(0.0021)	(0.0077)	(0.0029)	(0.0088)	(0.0037)	(0.011)
Pre-trend test p-value	<0.001	0.10	<0.001	0.27	<0.001	0.54
Surveys	78,136					
Panel B. The water-linked species						
Water-linked species, pooled	0.010***	0.0053	0.013***	0.0046	0.0016	0.0038
	(0.0032)	(0.0068)	(0.0043)	(0.0081)	(0.0018)	(0.0083)
Pre-trend test p-value	<0.001	0.03	<0.001	0.15	0.03	0.75
Surveys	73,651					
Submerged-plant grazers	0.0065***	0.048***	0.0050**	0.046***	0.0043**	0.049***
	(0.0023)	(0.0098)	(0.0024)	(0.010)	(0.0020)	(0.0093)
Pre-trend test p-value	<0.001	0.14	<0.001	0.24	<0.001	0.22
Surveys	59,785					
Panel C. The general groups						
Waterbirds	0.0020	−0.0031	0.0011	−0.0056	−0.0012	−0.0043
	(0.0021)	(0.0060)	(0.0023)	(0.0074)	(0.0021)	(0.0078)
Pre-trend test p-value	<0.001	0.08	<0.001	0.22	0.01	0.32
Surveys	60,739					
Waterfowl	0.013***	0.0053	0.011***	0.0023	0.0098***	0.0056
	(0.0022)	(0.0063)	(0.0030)	(0.0074)	(0.0018)	(0.0068)
Pre-trend test p-value	0.002	0.04	0.003	0.11	<0.001	0.14
Surveys	71,215					
Shorebirds	0.0018	−0.0033	−0.00022	−0.0018	0.0015	−0.0014
	(0.0024)	(0.0082)	(0.0029)	(0.011)	(0.0018)	(0.011)
Pre-trend test p-value	<0.001	0.09	<0.001	0.11	0.12	0.55
Surveys	41,249					
Landbirds	0.0033	−0.016*	0.0029	−0.021**	0.00037	−0.023**
	(0.0036)	(0.0089)	(0.0048)	(0.010)	(0.0024)	(0.010)
Pre-trend test p-value	<0.001	0.95	<0.001	0.79	<0.001	0.95
Surveys	78,136					
Circle and year effects	Yes	Yes	Yes	Yes	Yes	Yes
Effort controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
Attribute-by-year trends	Yes	Yes	Yes	Yes	Yes	Yes

Notes: As in Table 2, for the three diversity statistics, estimated in first differences by least squares on each statistic standardized within its group. Each estimate is the sum of the yearly increments through the fourteenth year after the dose rises by one, or after one more plant is commissioned, in standard deviations. Standard errors clustered by circle are in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels. The diversity statistics are undefined where a survey records no bird of the group. Thus, the regressions hold fewer surveys than those of Table 2.

Table B11: Two-way fixed-effects estimates of the effect of the dose on all birds' effective number of common species and entropy, with and without predicted cleaning

	(1) From the plants' cut in suspended solids	(2) From the halves split on the plants' suspended solids before the grants	(3) From the halves split on the plants' treatment level, on fecal coliform
Panel A. Effective number of common species, least squares, in standard deviations			
Dose alone	0.0054**	0.0053**	0.0034
	(0.0022)	(0.0022)	(0.0026)
Dose, with predicted cleaning	0.0067***	0.0044*	−0.0017
	(0.0014)	(0.0023)	(0.0032)
Predicted cleaning	0.023	−0.037	0.66***
	(0.015)	(0.076)	(0.19)
Surveys	79,120	82,810	81,950
Panel B. Entropy, least squares, in standard deviations			
Dose alone	0.0019*	0.0018	0.0018
	(0.0011)	(0.0011)	(0.0017)
Dose, with predicted cleaning	0.0029**	0.0041**	−0.0061*
	(0.0014)	(0.0016)	(0.0033)
Predicted cleaning	0.017**	0.090*	1.0**
	(0.0080)	(0.047)	(0.41)
Surveys	79,120	82,810	81,950
Circle and year effects	Yes	Yes	Yes
Effort controls	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
Attribute-by-year trends	Yes	Yes	Yes

Notes: As in Table 4, for the effective number of common species and the entropy, by least squares on each statistic standardized over all birds' surveys, in standard deviations. Standard errors clustered by circle are in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels.

Table B12: Ecosystem services the literature reports for the feeding guilds of five water-linked species

Species	Feeding guild	Ecosystem services of its group
Mallard	Submerged-plant grazer	Hunting, seed dispersal to new and restored wetlands, early detection of avian influenza
Canvasback	Submerged-plant grazer	Hunting, grazing that can raise submerged-plant diversity, a diet that tracked the loss of submerged plants
Lesser Scaup	Bottom-sediment feeder	Hunting, predation that reduces invasive zebra mussels, whose U.S. damages and control cost about $1 billion a year
Ring-billed Gull	Scavenging gull	Disposal of waste at dumps, removal of nutrients from dumps and treatment plants, control of farm pests
Great Blue Heron	Visual fish-eater	Removal of fish before die-offs as waters fall, predation on crayfish pests, a receptor species for risk assessments of contaminants

Notes: The paper classifies each water-linked species into one feeding guild (Section 3.2). The services are those Green and Elmberg (2014) report by group of species (dabbling and diving ducks, gulls, and herons), several as likely and most from studies outside North America, with the hunting of the Mallard and the Lesser Scaup from the 2022 National Survey (U.S. Fish and Wildlife Service (2023)), the Canvasback's diet from Perry and Deller (1996), and the heron as a receptor species from Seston et al. (2009). None is measured for these species in this paper.

Table B13: Treatment technologies: how each works, its lifetime cost, and what the literature reports for birds

Technology	Category	How it works	Lifetime	Lifetime cost	Benefits for birds
Facultative lagoon	I, secondary	Basins 8 to 16 feet deep in which aeration aids digestion	20 years	$0.42, Aroostook-Prestile, Maine	Oxygen rises, plants and invertebrates recolonize, and birds may benefit (Pringle and Burton, 2017)
Biotower	I, secondary	Bacteria grown on a plastic medium degrade the organic matter	20 years	$0.39, Tolleson, Arizona	
Activated sludge with a pressure filter	I, secondary	Bacteria degrade the organic matter. The filter, added in case limits tightened, proved unnecessary	20 years	$1.01, Valley Forge, Pennsylvania, grant-funded	
Nitrification and phosphorus removal	IIB, advanced	Bacteria convert ammonia to nitrate, and filters remove phosphorus			Submerged plants, food for grazing waterbirds, returned to the upper Potomac in 1983, with better treatment and unusual weather (Dennison et al., 1993)
Constructed wetland	Innovative or alternative	Effluent flows through a planted wetland		$166,000 to build, against $500,000 to upgrade conventionally, Monterey, Virginia	Habitat, with 412 bird species recorded at 28 treatment wetlands (Knight et al., 2001)

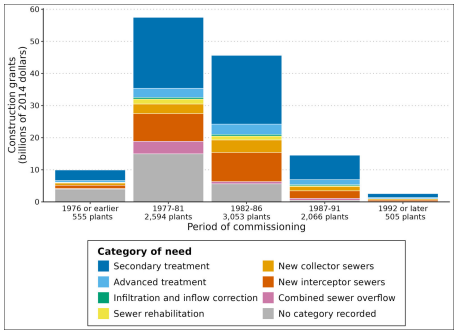
Notes: Lifetime cost is capital plus operating costs over the assumed 20-year life, discounted at 10%, in 1983 dollars per 1,000 gallons, for the plant named (Congressional Budget Office, 1985). Each category is a class of need in the 1982 Needs Survey (U.S. Environmental Protection Agency, 1982), and the benefit reported for the first secondary technology applies to all three. The advanced treatment is Blue Plains, Washington, D.C. (Carter and Rybicki, 1986; Ruhl and Rybicki, 2010), and the wetland's costs are Monterey, Virginia's (U.S. General Accounting Office, 1994).

Table B14: Grant-funded plants by period of commissioning

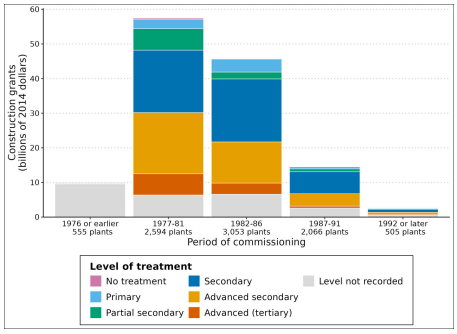
	1976 or earlier	1977-81	1982-86	1987-91	1992 or later
Plants	556	2,594	3,053	2,066	505
Years from the first construction grant to commissioning					
Median	2.1	3.1	3.5	3.0	3.9
25th to 75th percentile	1.4 to 3.1	2.1 to 4.3	2.4 to 5.0	2.0 to 4.8	2.6 to 6.0
Share taking more than 5 years (percent)	6.3	16	26	23	35
Share first awarded before October 1984 (percent)	100	100	99	42	13
Median design flow (million gallons a day)	0.30	0.50	0.43	0.30	0.19
Median construction grants (millions of 2014 dollars)	2.1	4.3	3.8	2.4	2.1
Share holding a Step 4, 7, or 9 grant (percent)	15	10	33	47	62
Circles first treated in the period	466	1,341	664	276	47

Notes: Every plant with a completed, funded construction grant of Step 3, 4, 7, or 9, by the period of its commissioning (Section 3.3). The years from the first construction grant to commissioning run from that grant's award to the completion date that sets commissioning. A grant awarded before October 1984 carried the 75% federal share (Section 2). Design flow is read from the Needs Survey wave the dose uses, and a plant's grants are its funded construction grants in 2014 dollars, each deflated by the consumer price index of its award year. The last row counts the circles whose first upstream plant was commissioned in the period.

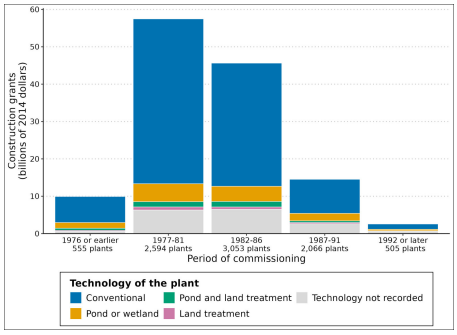
B.2 Figures



(a) By category of need



(b) By the plant's level of treatment



(c) By the plant's technology

Figure B1: Construction grants by period of commissioning, in billions of 2014 dollars, by category of need, the plant's level of treatment, and its technology.

Notes: Each column is labeled with its plants. Panel (a) splits each grant's dollars across the Needs Survey's categories of need in proportion to its eligible costs by category and leaves out the innovative and alternative technology amounts, which overlap the categories. Of the dollars whose category is recorded, 61% paid for treatment and 39% for sewers. Panel (b) takes the plant's level of treatment from the Needs Survey wave that supplies its design flow, which the 1973 and 1976 waves do not record. Two-thirds of the dollars, and 84% of those whose plant has a level recorded, went to plants at secondary treatment or beyond. Panel (c)'s technology is recorded for about 80% of the plants. Conventional plants received about $95 billion of the $130 billion, and the 3,104 plants built around ponds, wetlands, or land treatment about $19 billion.

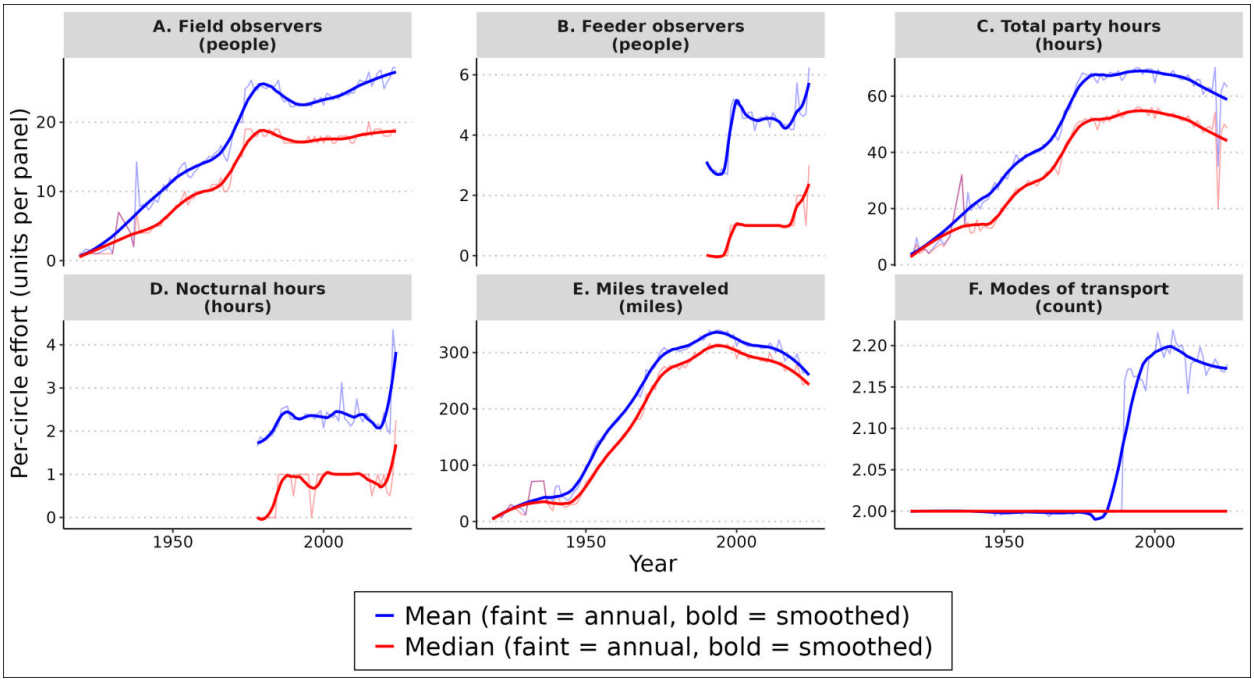


Figure B2: Survey effort per circle in each year, 1919 to 2024: the mean and the median of each measure, smoothed, with the annual values faint beneath.

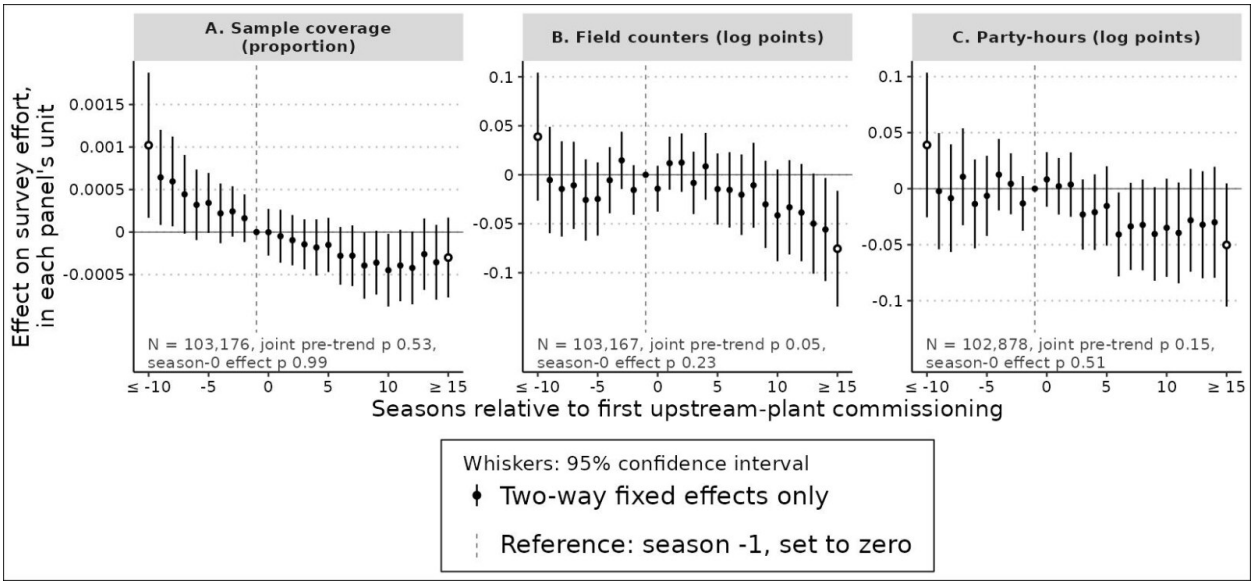


Figure B3: Survey effort by year relative to the first commissioning of a grant-funded plant upstream: sample coverage, field observers and party-hours.

Notes: Event study of each measure of effort on commissioning with circle and year effects only, year −1 the reference, set to zero (dashed line). Field observers and party-hours are in log points. Sample coverage, the estimated share of a survey's birds that belong to the species it detected, is a proportion. The first and last points pool every year 10 or more before and 15 or more after the first commissioning. They are estimated differently from the interior coefficients and are drawn hollow, and neither enters the window. Whiskers are 95% confidence intervals from standard errors clustered by circle. Each panel prints the number of surveys, the p-value of the joint pre-trend test, and the p-value of the effect in the year of the first commissioning.

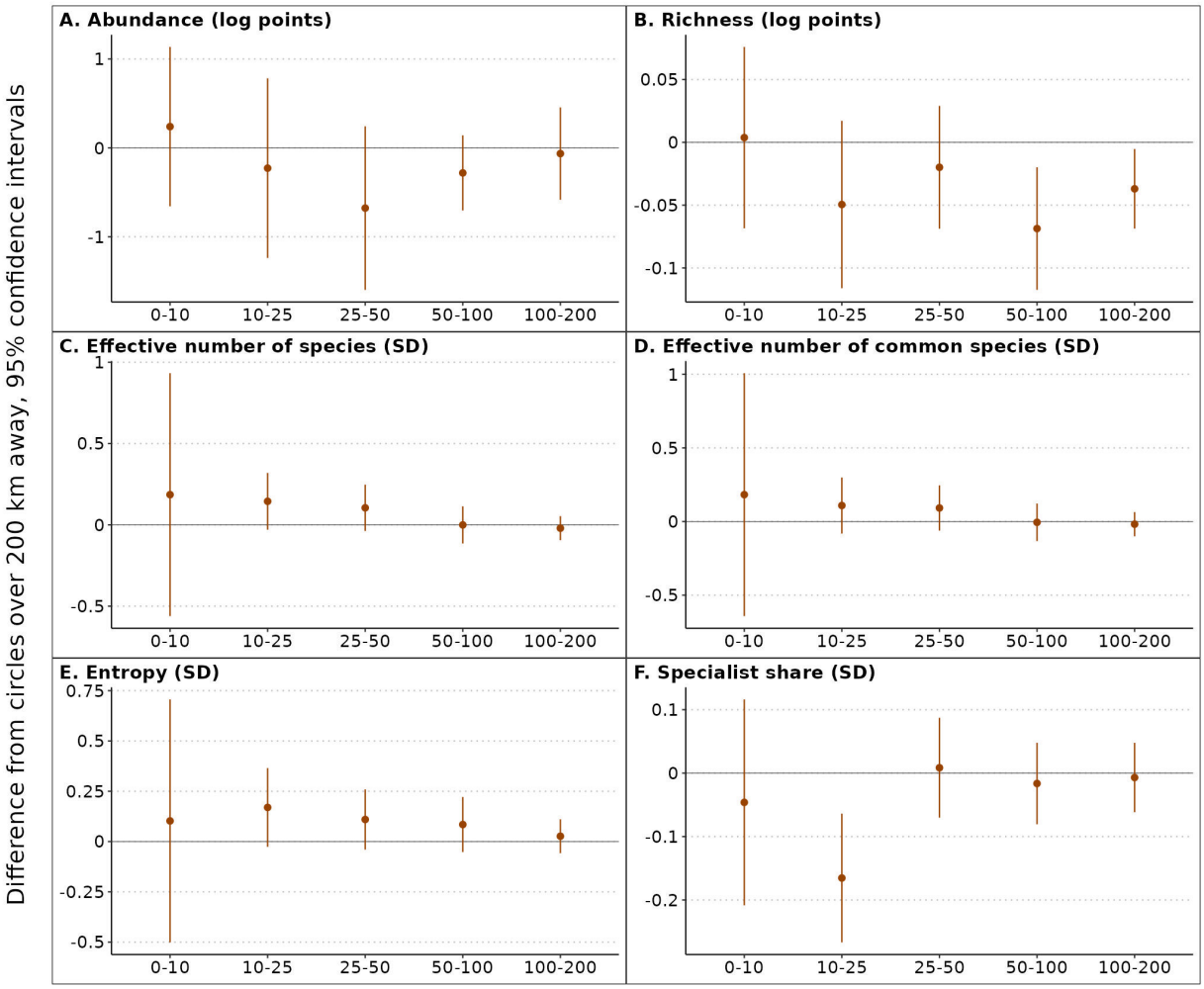


Figure B4: Circles with no grant-funded plant upstream: each statistic by distance to the nearest treated circle, relative to circles more than 200 kilometers away.

Notes: On the surveys of circles with no grant-funded plant upstream, 26,240 for abundance and richness and 26,237 for the three diversity statistics, each statistic is regressed on indicators for the distance to the nearest circle treated in that year, with circle and year effects, the bin beyond 200 kilometers the reference: abundance and richness by Poisson quasi-likelihood, in log points, and the standardized statistics by least squares, in standard deviations. Whiskers are 95% confidence intervals from standard errors clustered by circle. The joint test that the bins within 25 kilometers are zero has a p-value of 0.68 for abundance, 0.31 for richness, 0.26 for the effective number of species, 0.52 for the effective number of common species, and 0.24 for the entropy. Panel F is not used in this paper.

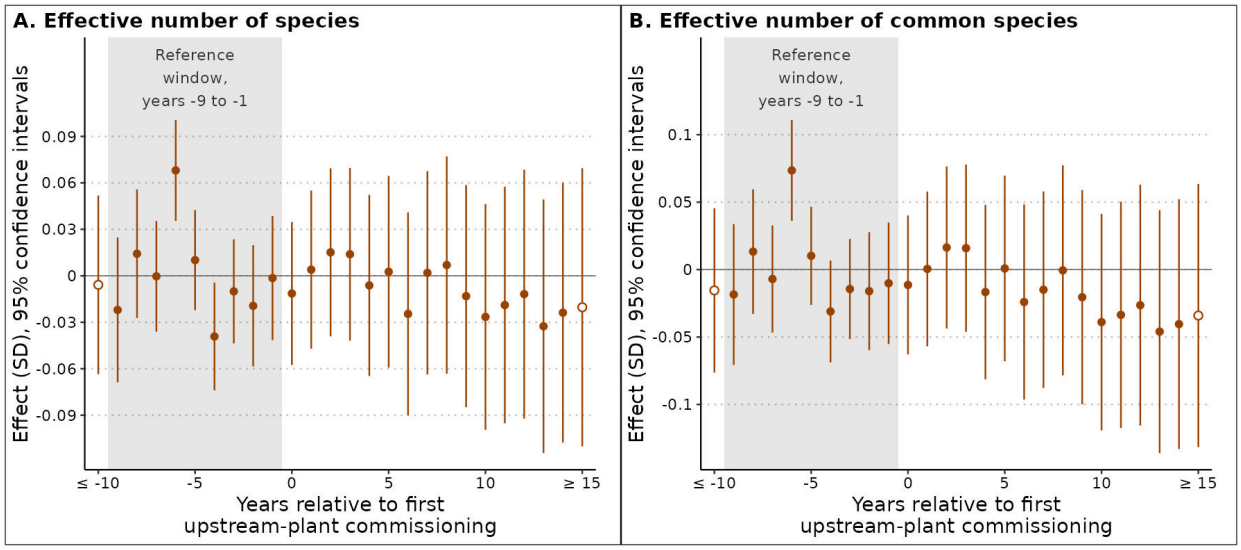
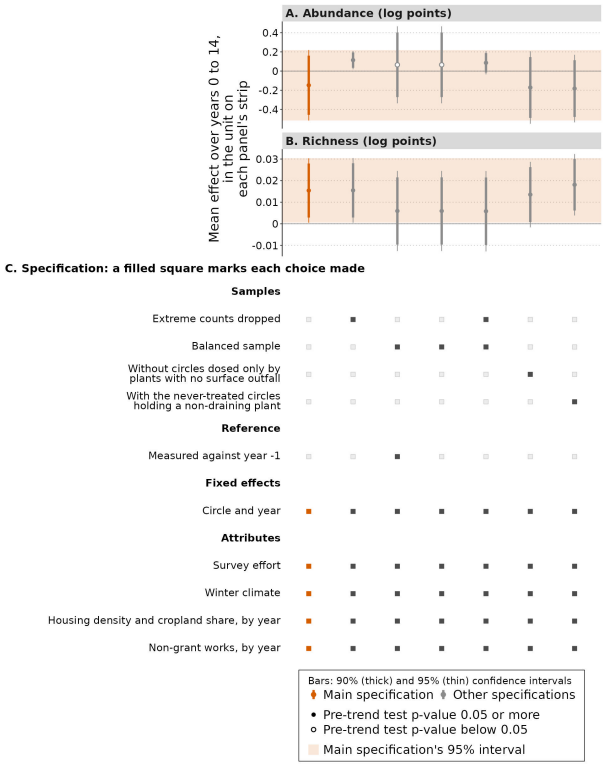
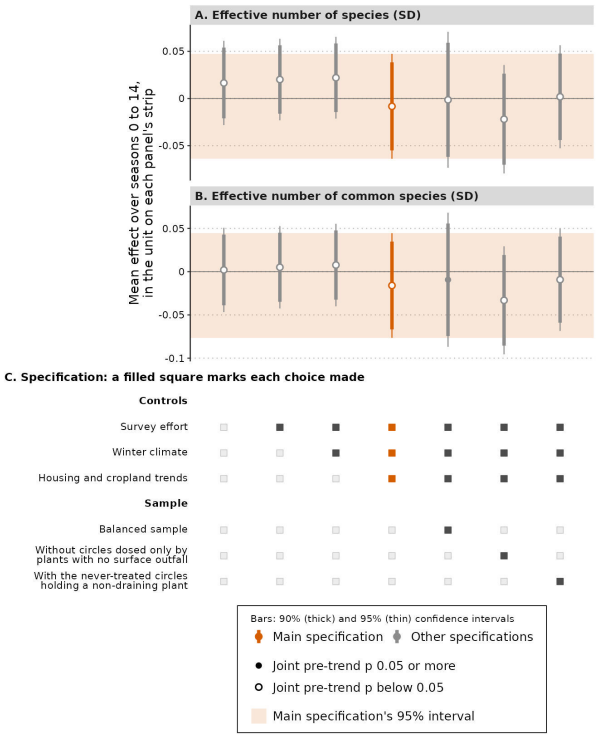


Figure B5: The effective numbers of species and of common species of all birds by year relative to the first commissioning of a grant-funded plant upstream, Christmas Bird Count circles, 1940 to 2024.

Notes: Event study of equation (1) at the main specification, on 85,934 surveys at 2,975 circles, each statistic standardized and fitted by least squares with circle and year effects and the controls of Section 4.1: panel A, the effective number of species, and panel B, the effective number of common species, both in standard deviations. Each coefficient is measured against the average of years −9 to −1, the shaded window (Miller, 2023). Thus, year −1 carries its own interval. The first and last points pool every year 10 or more before and 15 or more after the first commissioning. They are estimated differently from the interior coefficients and are drawn hollow. Whiskers are 95% confidence intervals from standard errors clustered by circle. The pre-trend test, a Wald test that every pre-treatment coefficient is zero in the regression that omits year −1, has a p-value of 0.008 for the effective number of species and 0.03 for the effective number of common species. The window averages around which the coefficients are normalized are 0.001 (SE 0.020) and 0.010 (SE 0.023). The entropy, which Appendix Table B9 reports, is not drawn.



(a) Abundance and richness



(b) Effective numbers of species and of common species

Figure B6: Specification chart for all birds: the event study's average effect over years 0 to 14 under each specification.

Notes: Each column is one specification of equation (1), its choices marked by the filled squares beneath. The point is the mean of the event-time coefficients over years 0 to 14 less their mean over years −9 to −1; the bars are 90% and 95% confidence intervals from standard errors clustered by circle, and the shaded band is the main specification's 95% interval. A hollow point rejects the joint pre-trend test at the 5% level. With the extreme counts dropped, every record above the 97.5th percentile of its year's distinct count values is left out, for abundance and richness only. The entropy is not drawn. A specification with fewer than 100 treated circles is not drawn.

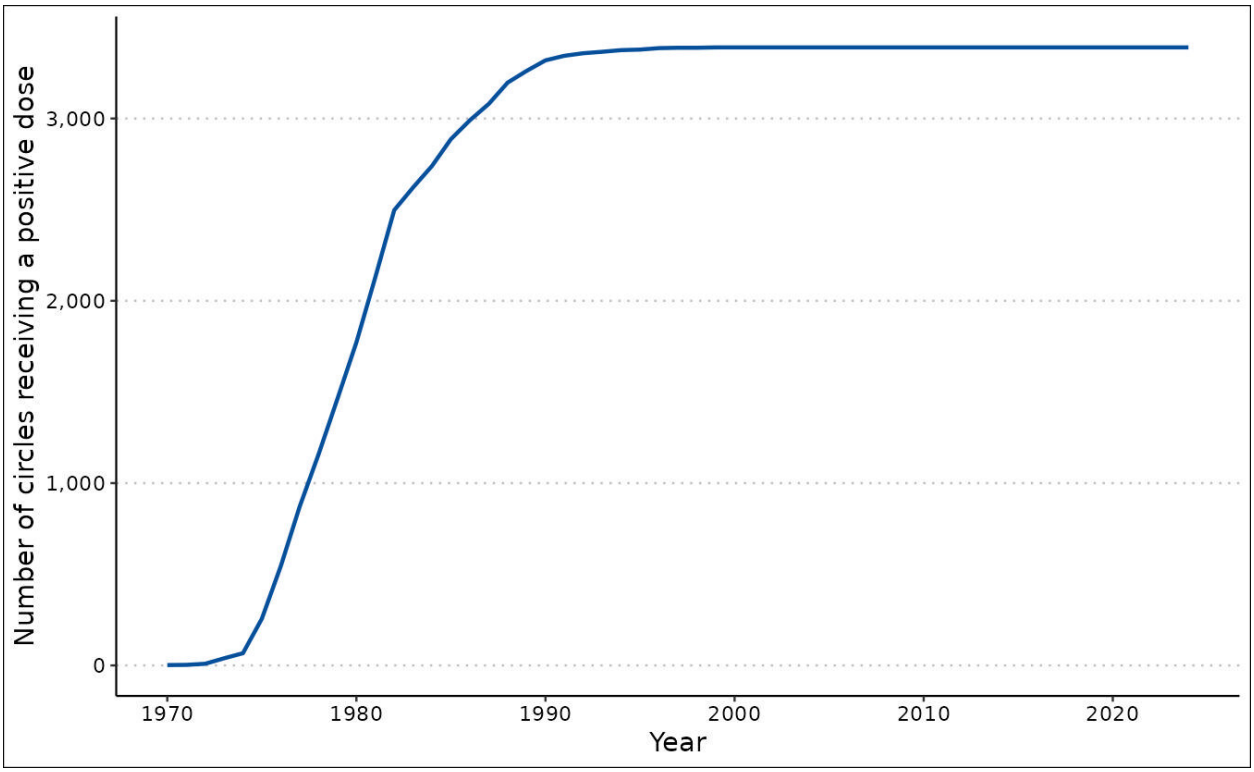


Figure B7: The number of circles receiving a positive dose from a commissioned grant-funded plant within 25 river miles upstream, by year, 1970 to 2024.

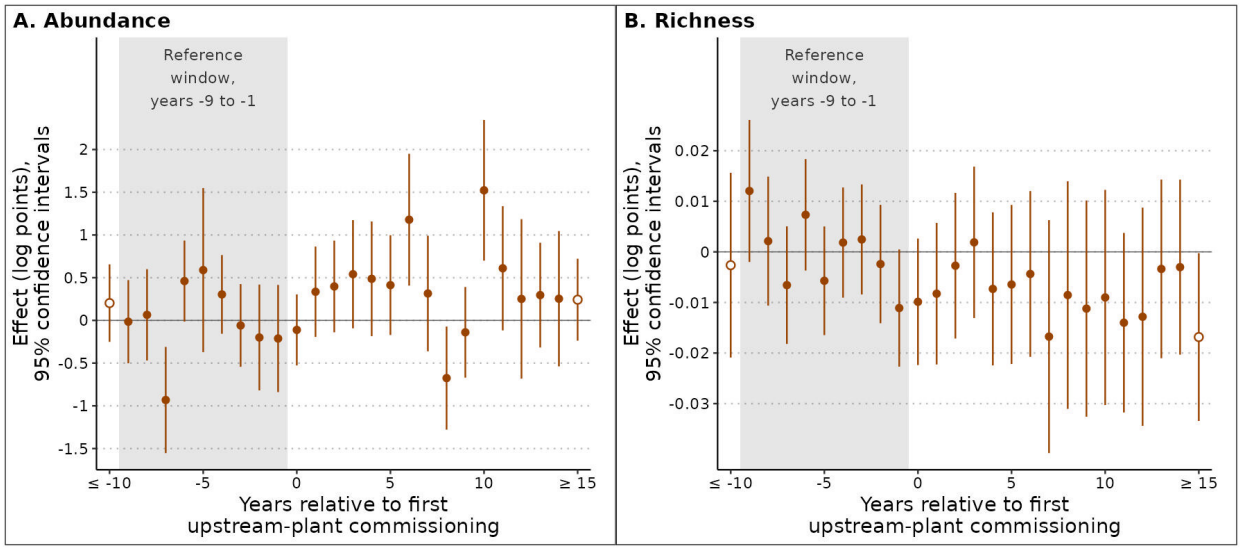


Figure B8: All birds' abundance and richness in the circles with a higher dose at commissioning against those with a lower one, by year relative to the first commissioning of a grant-funded plant upstream.

Notes: The treated circles are split at the median of their dose at first commissioning within each bin of commissioning cohorts, and the event study of equation (5) at the main specification is estimated for the higher half against the lower, on 71,835 surveys at 2,665 circles, 1,333 in the higher half and 1,332 in the lower, by Poisson quasi-likelihood with circle effects, bin-by-year effects, and the controls of Section 4.1. Both statistics are in log points. Each coefficient is measured against the average of years −9 to −1, the shaded window (Miller, 2023). The first and last points pool every year 10 or more before and 15 or more after the first commissioning and are drawn hollow. Whiskers are 95% confidence intervals from standard errors clustered by circle. The pre-trend test, the joint test that the higher half's pre-treatment coefficients are zero, has a p-value of 0.02 for abundance and 0.25 for richness, and the window averages around which the coefficients are normalized are 0.21 (SE 0.32) and 0.0111 (SE 0.0059).

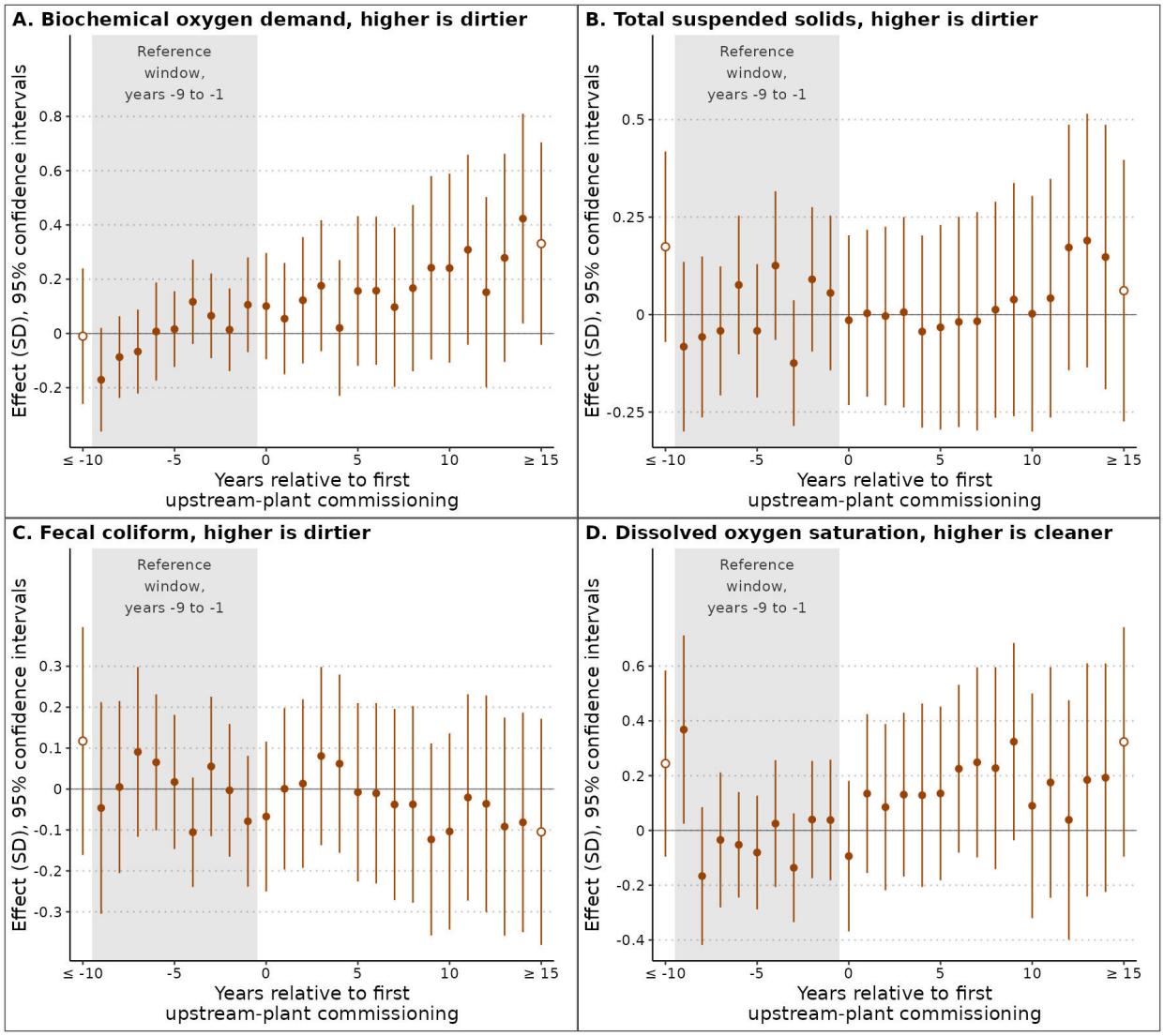


Figure B9: The regulated analytes by year relative to the first commissioning of a grant-funded plant upstream.

Notes: Event study of each analyte's within-circle standard score on commissioning, by least squares with circle and year effects and the controls of equation (7), on the circle-years with a reading: 9,232 at 953 circles for oxygen demand, 26,288 at 1,957 for suspended solids, 16,314 at 1,297 for fecal coliform, and 14,478 at 1,215 for dissolved oxygen. Each coefficient is measured against the average of years −9 to −1, the shaded window (Miller, 2023). The first and last points pool every year 10 or more before and 15 or more after commissioning and are drawn hollow. Whiskers are 95% confidence intervals from standard errors clustered by circle. In the same order, the pre-trend test's p-values are 0.62, 0.31, 0.74, and 0.06, and the window averages around which the coefficients are normalized are −0.106 (SE 0.089), −0.06 (SE 0.10), 0.079 (SE 0.082), and −0.04 (SE 0.11). Fecal coliform enters in logarithms.

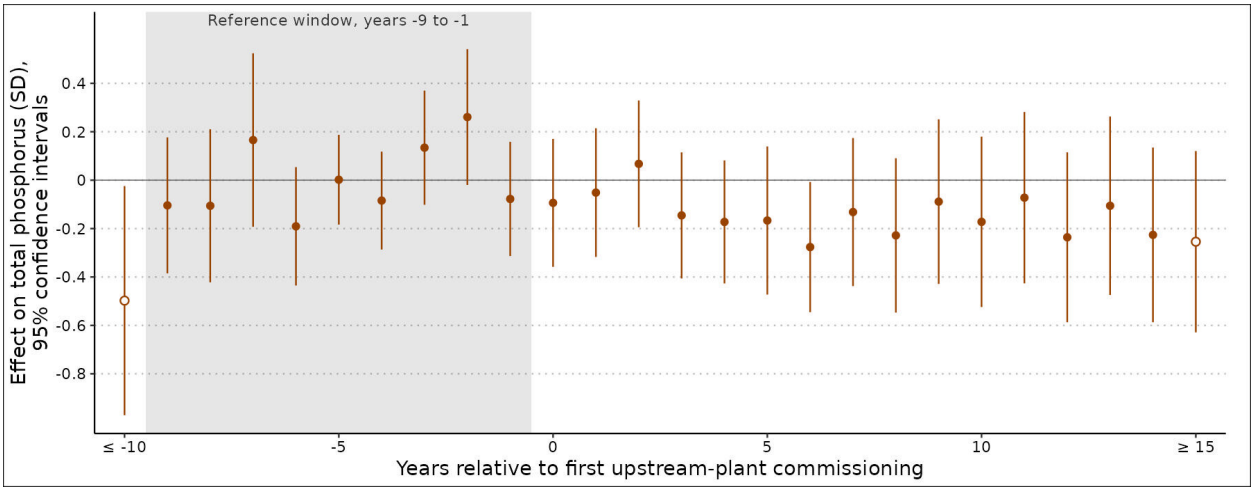


Figure B10: Total phosphorus by year relative to the first commissioning of a grant-funded plant upstream.

Notes: As in Appendix Figure B9, on 24,026 circle-years at 1,860 circles, with a pre-trend p-value of 0.20 and a window average of 0.08 (SE 0.12).

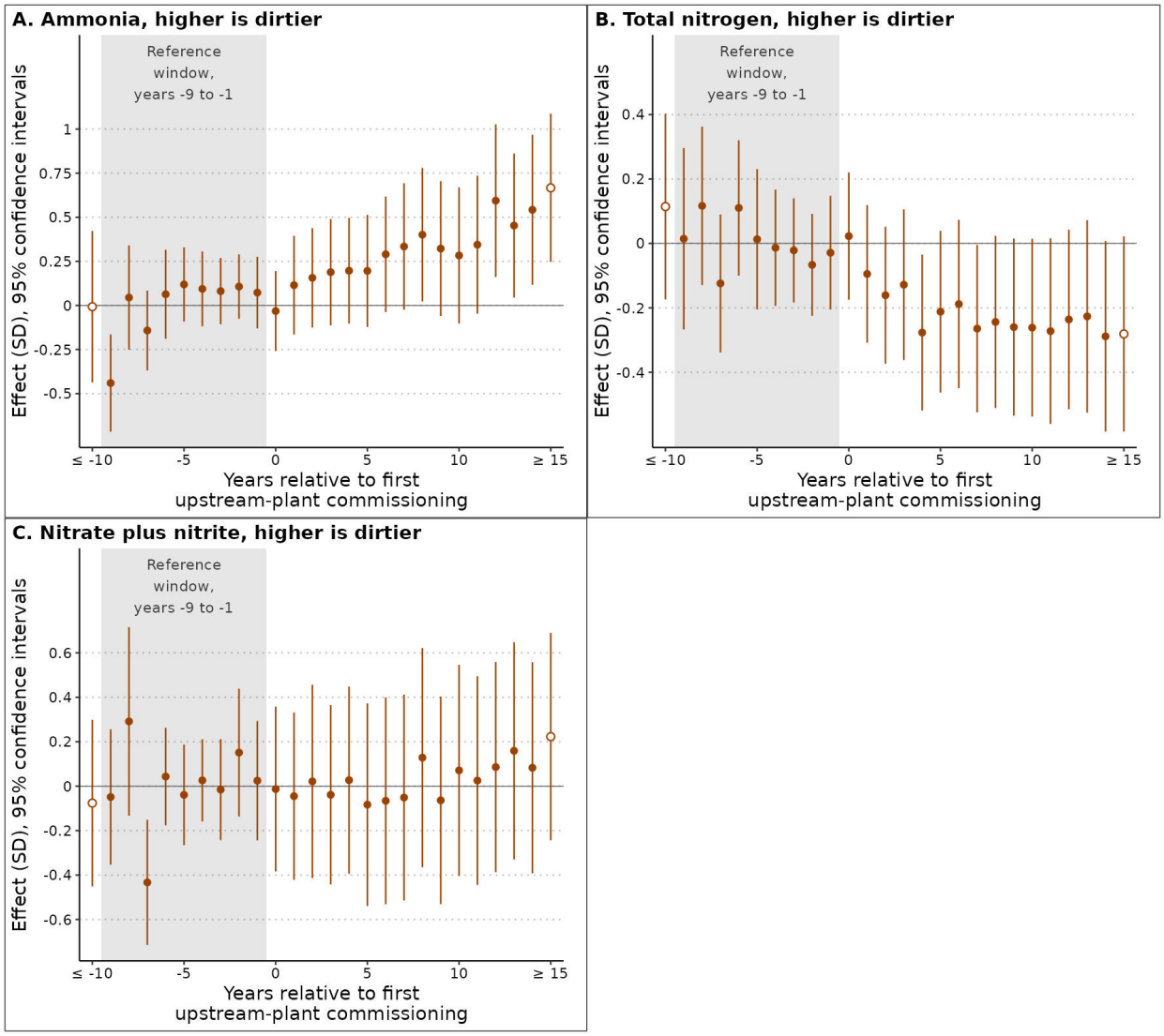


Figure B11: Ammonia, total nitrogen and nitrate plus nitrite by year relative to the first commissioning of a grant-funded plant upstream.

Notes: As in Appendix Figure B9, on 20,865 circle-years at 1,828 circles for ammonia, 18,867 at 1,686 for total nitrogen, and 17,334 at 1,770 for nitrate plus nitrite, with pre-trend p-values of 0.04, 0.86, and 0.13 and window averages of −0.07 (SE 0.10), 0.029 (SE 0.090), and −0.02 (SE 0.14) in the same order.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics 138*(1), 1–35.

Adams, C. M., D. L. Winkelman, and R. M. Fitzpatrick (2023). Impact of wastewater treatment plant effluent on the winter thermal regime of two urban colorado south platte tributaries. *Frontiers in Environmental Science 11*, 1120412.

Anza, I., D. Vidal, C. Laguna, S. Díaz-Sánchez, S. Sánchez, Á. Chicote, M. Florín, and R. Mateo (2014). Eutrophication and bacterial pathogens as risk factors for avian botulism outbreaks in wetlands receiving effluents from urban wastewater treatment plants. *Applied and Environmental Microbiology 80*(14), 4251–4259.

Baylis, K., A. Garcia, and R. Heilmayr (2026). Causal inference for biodiversity conservation. *Review of Environmental Economics and Policy 20*(1), 98–124.

BirdLife International and NatureServe (2024). Bird species distribution maps of the world, version 2024.2. BirdLife International, Cambridge, UK. Request-gated; not a public download. Species range maps requested via datazone.birdlife.org/species/requestdis. Key Biodiversity Area polygons are a separate request via keybiodiversityareas.org/kba-data/request.

Borja, A., J. Franco, and V. Pérez (2000). A marine biotic index to establish the ecological quality of soft-bottom benthos within european estuarine and coastal environments. *Marine Pollution Bulletin 40*(12), 1100–1114.

Boyd, J. and S. Banzhaf (2007). What are ecosystem services? the need for standardized environmental accounting units. *Ecological Economics 63*(2-3), 616–626.

Butts, K. (2026, February). Difference-in-differences with spatial spillovers. Working paper, University of Arkansas. Revision of arXiv:2105.03737; no DOI.

Callaway, B., A. Goodman-Bacon, and P. H. C. Sant'Anna (2025). Difference-in-differences with a continuous treatment. NBER Working Paper 32117, National Bureau of Economic Research. arXiv:2107.02637, v8 (December 31, 2025).

Callaway, B. and P. H. C. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230.

Campbell, L. H. (1978). Patterns of distribution and behaviour of flocks of seaducks wintering at Leith and Musselburgh, Scotland. *Biological Conservation 14*(2), 111–124.

Campbell, L. H. (1984). The impact of changes in sewage treatment on seaducks wintering in the Firth of Forth, Scotland. *Biological Conservation 28*(2), 173–180.

Carter, V. and N. Rybicki (1986). Resurgence of submersed aquatic macrophytes in the tidal Potomac River, Maryland, Virginia, and the District of Columbia. *Estuaries 9*(4B), 368–375.

Chao, A. and L. Jost (2012). Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology 93*(12), 2533–2547.

Chen, J. and J. Roth (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics 139*(2), 891–938.

Chen, Q., J. Li, Y. Liang, W. Zhao, Y. Cheng, and T. Huang (2026). Ecological costs of energy infrastructure: Evidence from migratory bird tracking. *Journal of Environmental Economics and Management 138*, 103352. The byline prints the authors surname first (Chen Qiangyuan, Li Jianmin, Liang Yuanning, Zhao Weiping, Cheng Yunqian, Huang Tian).

Congressional Budget Office (1985, June). Efficient investments in wastewater treatment plants. CBO Study 85-CBO-016, Congressional Budget Office, Washington, D.C. Also available via HathiTrust: https://catalog.hathitrust.org/Record/001818130 and Google Play Books: https://play.google.com/store/books/details?id=nDeALSwnxaEC.

Conley, T. G., C. B. Hansen, and P. E. Rossi (2012). Plausibly exogenous. *Review of Economics and Statistics 94*(1), 260–272.

Contreras, A. and R. Hoyer (2022). The art of the Christmas Bird Count. National Audubon Society, Christmas Bird Count compiler resources. Revised 2022.

Copeland, C. (2016). Clean water act: A summary of the law. Technical Report RL30030, Congressional Research Service, Washington, D.C.

Daily, G. C., S. Polasky, J. Goldstein, P. M. Kareiva, H. A. Mooney, L. Pejchar, T. H. Ricketts, J. Salzman, and R. Shallenberger (2009). Ecosystem services in decision making: time to deliver. *Frontiers in Ecology and the Environment 7*(1), 21–28.

Dasgupta, P. (2021). *The Economics of Biodiversity: The Dasgupta Review*. London: HM Treasury.

De Graaf, R. M., N. G. Tilghman, and S. H. Anderson (1985). Foraging guilds of north american birds. *Environmental Management 9*(6), 493–536.

Dennison, W. C., R. J. Orth, K. A. Moore, J. C. Stevenson, V. Carter, S. Kollar, P. W. Bergstrom, and R. A. Batiuk (1993). Assessing water quality with submersed aquatic vegetation. *BioScience 43*(2), 86–94.

Ehalt Macedo, H., B. Lehner, J. Nicell, G. Grill, J. Li, A. Limtong, and R. Shakya (2022). Distribution and characteristics of wastewater treatment plants within the global river network. *Earth System Science Data 14*(2), 559–577.

Fan, C. (2025). Dam Thy Neighbor: International relations and the externalities of impounding rivers. Job Market Paper, University of Chicago Harris School of Public Policy.

Flynn, P. and M. Marcus (2025). A watershed moment: The Clean Water Act and birth weight. *Journal of Human Resources 60*(6), 1915–1936.

Flynn, P. and T. Smith (2022). Rivers, lakes and revenue streams: The heterogeneous effects of Clean Water Act grants on local spending. *Journal of Public Economics 212*, 104711.

Fox, A. D., H. E. Jørgensen, E. Jeppesen, T. L. Lauridsen, M. Søndergaard, K. Fugl, P. Myssen, T. J. S. Balsby, and P. Clausen (2020). Relationships between breeding waterbird abundance, diversity, and clear water status after the restoration of two shallow nutrient-rich Danish lakes. *Aquatic Conservation: Marine and Freshwater Ecosystems 30*(2), 237–245.

Frank, E. and A. Sudarshan (2024). The social costs of keystone species collapse: Evidence from the decline of vultures in India. *American Economic Review 114*(10), 3007–3040.

Fuller, R. J. and D. E. Glue (1980). Sewage works as bird habitats in Britain. *Biological Conservation 17*(3), 165–181.

Green, A. J. and J. Elmberg (2014). Ecosystem services provided by waterbirds. *Biological Reviews 89*(1), 105–122.

Greening, H. and A. Janicki (2006). Toward reversal of eutrophic conditions in a subtropical estuary: water quality and seagrass response to nitrogen loading reductions in Tampa Bay, Florida, USA. *Environmental Management 38*(2), 163–178.

Haines, M., P. Fishback, and P. Rhode (2014). United states agriculture data, 1840–2012. Inter-university Consortium for Political and Social Research (ICPSR) 35206. Version 4.

Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut (2023). ERA5 monthly averaged data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Accessed: 2026-01-09.

Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology 54*(2), 427–432.

Hu, B., Y. Wang, J. Lin, and L. Geng (2026). Can rural tourism policy and development influence local bird diversity? A quasi-field experiment. *Current Issues in Tourism*, 1–15. Advance online publication, August 4, 2026; no volume or issue assigned.

Jayalath, T. A., P. Lloyd-Smith, and M. Becker (2023). Biodiversity benefits of birdwatching using citizen science data and individualized recreational demand models. *Environmental and Resource Economics 86*(1-2), 83–107.

Jerch, R. L. (2026). Municipal governments under the Clean Water Act. *Journal of the Association of Environmental and Resource Economists 13*(3), 791–830.

Katovich, E. (2024). Quantifying the effects of energy infrastructure on bird populations and biodiversity. *Environmental Science and Technology 58*(1), 323–332.

Keiser, D. A., C. L. Kling, and J. S. Shapiro (2019). The low but uncertain measured benefits of US water quality policy. *Proceedings of the National Academy of Sciences 116*(12), 5262–5269.

Keiser, D. A. and J. S. Shapiro (2019a). Consequences of the Clean Water Act and the demand for water quality. *The Quarterly Journal of Economics 134*(1), 349–396.

Keiser, D. A. and J. S. Shapiro (2019b). US water pollution regulation over the past half century: Burning waters to crystal springs? *Journal of Economic Perspectives 33*(4), 51–75.

Knight, R. L. (1997). Wildlife habitat and public use benefits of treatment wetlands. *Water Science and Technology 35*(5), 35–43.

Knight, R. L., R. A. Clarke, Jr., and R. K. Bastian (2001). Surface flow (SF) treatment wetlands as a habitat for wildlife and humans. *Water Science and Technology 44*(11-12), 27–37.

Kolstoe, S. and T. A. Cameron (2017). The non-market value of birding sites and the marginal value of additional species: Biodiversity in a random utility model of site choice by eBird members. *Ecological Economics 137*, 1–12.

Lepori, F. (2022). Stream pollution causes aggregation of wintering insectivorous birds through increased aquatic emergence. *Frontiers in Ecology and Evolution 10*, 926529.

Lewis, D. J., D. M. Kling, S. J. Dundas, and D. K. Lew (2022). Estimating the value of threatened species abundance dynamics. *Journal of Environmental Economics and Management 113*, 102639.

Liang, Y., I. Rudik, and E. Y. Zou (2025). Economic activity and biodiversity in the United States. *Journal of Public Economics 250*, 105482.

Liang, Y., I. Rudik, E. Y. Zou, A. Johnston, A. D. Rodewald, and C. L. Kling (2020). Conservation cobenefits from air pollution regulation: Evidence from birds. *Proceedings of the National Academy of Sciences 117*(49), 30900–30906.

Link, W. A., J. R. Sauer, and D. K. Niven (2006). A hierarchical model for regional analysis of population change using Christmas Bird Count data, with application to the American Black Duck. *The Condor 108*(1), 13–24.

Loyn, R. H., D. I. Rogers, R. J. Swindley, P. W. Menkhorst, K. Stamation, S. Haynes, H. Graham, G. Hepworth, and W. K. Steele (2023). Waterfowl populations decline with nutrient reduction and increase with nutrient restoration: 20 years of adaptive management at a Ramsar-listed wastewater treatment plant. *Hydrobiologia 850*(19), 4127–4147.

Madhok, R. (2025). Infrastructure, institutions, and the conservation of biodiversity in india. *Journal of the Association of Environmental and Resource Economists 12*(6), 1705–1745.

Markley, S. N., S. R. Holloway, T. J. Hafley, and M. E. Hauer (2022). Housing unit and urbanization estimates for the continental U.S. in consistent tract boundaries, 1940–2019. *Scientific Data 9*.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Rea (2012). NHDPlus version 2: User guide. U.S. Environmental Protection Agency, Office of Water. National Hydrography Dataset Plus, Version 2 (Data Model 2.1); guide revised 13 March 2019.

Meehan, T. D., N. L. Michel, and H. Rue (2019). Spatial modeling of Audubon Christmas Bird Counts reveals fine-scale patterns and drivers of relative abundance trends. *Ecosphere 10*(4), e02707.

Meier, S. and E. Strobl (2026). The impact of wildfire smoke on local avian biodiversity. *Journal of Environmental Economics and Management 139*, 103379.

Melstrom, R. T. and C. Reeling (2026). Modeling recreation and tourism demand using crowdsourced data: An application to visitation statistics from the ebird project. *Tourism Management 112*, 105276.

Meng, L., P. Liu, Y. Zhou, and Y. Mei (2025). Blaming the wind? The impact of wind turbine on bird biodiversity. *Journal of Development Economics 172*, 103402.

Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives 37*(2), 203–230.

National Audubon Society (2024). Compiler's manual: How to plan and conduct a Christmas Bird Count. Technical report, National Audubon Society.

National Audubon Society (2025a). Christmas Bird Count (CBC) raw data README file. National Audubon Society, distributed with the raw-data request. Version 2.1, May 19, 2025.

National Audubon Society (2025b). Christmas Bird Count Results. http://www.christmasbirdcount.org. Accessed: October 28, 2025.

Noack, F., D. Engist, and A. Larsen (2025). The value of biodiversity: Evidence from migratory birds. Working paper, SSRN.

Noack, F., A. Larsen, J. Kamp, and C. Levers (2022). A bird's eye view of farm size and biodiversity: The ecological legacy of the iron curtain. *American Journal of Agricultural Economics 104*(4), 1460–1484. Published online November 16, 2021.

Noby, K., I. M. A. Heitkönig, E. Bulte, and M. Naguib (2025). Sustainability and avian biodiversity tensions in wastewater management in arid zones. *Biological Conservation 305*, 111046.

Noordhuis, R., D. T. van der Molen, and M. S. van den Berg (2002). Response of herbivorous water-birds to the return of Chara in Lake Veluwemeer, The Netherlands. *Aquatic Botany 72*(3-4), 349–367.

Office of Management and Budget (2023a, November). Circular no. A-4: Regulatory analysis. Technical report, Executive Office of the President, Washington, D.C.

Office of Management and Budget (2023b, November). Circular no. A-94: Guidelines and discount rates for benefit-cost analysis of federal programs. Technical report, Executive Office of the President, Washington, D.C. Revised November 9, 2023.

Office of the Federal Register (1999). Title 40, part 35: State and local assistance. Code of Federal Regulations, July 1, 1999 edition. Subparts E (the 1978 construction grants rule), I (the 1984 rule), J (delegation to states) and K (state revolving funds); the printed edition that still carries Subpart E.

Office of the Federal Register (2026). Title 40, part 133: Secondary treatment regulation. Electronic Code of Federal Regulations, current as of October 2026.

Olander, L. P., R. J. Johnston, H. Tallis, J. Kagan, L. A. Maguire, S. Polasky, D. Urban, J. Boyd, L. Wainger, and M. Palmer (2018). Benefit relevant indicators: Ecosystem services measures that link ecological and social outcomes. *Ecological Indicators 85*, 1262–1272.

Olmstead, S. M. (2010). The economics of water quality. *Review of Environmental Economics and Policy*.

Palmer, C., B. Groom, L. Sileci, and S. Langton (2026). Biodiversity–food trade-offs when agricultural land is spared from production. *American Journal of Agricultural Economics 108*(1), 254–284. Published online February 9, 2025.

Partners in Flight (2024). Avian conservation assessment database, version 2024. Partners in Flight / Bird Conservancy of the Rockies. Database available at https://pif.birdconservancy.org/ACAD, accessed July 6, 2026.

Perry, M. C. and A. S. Deller (1996). Review of factors affecting the distribution and abundance of waterfowl in shallow-water habitats of Chesapeake Bay. *Estuaries 19*(2A), 272–278.

Polasky, S., C. Costello, and A. Solow (2005). The economics of biodiversity. In K.-G. Mäler and J. R. Vincent (Eds.), *Handbook of Environmental Economics, Volume 3*, pp. 1517–1560. Elsevier.

Pringle, H. E. K. and N. H. K. Burton (2017, August). Improving understanding of the possible relationship between improving freshwater and coastal water quality and bird interest on designated sites: phase 1 review. BTO Research Report 696, British Trust for Ornithology.

Ramseur, J. L. (2016, June). Allocation of wastewater treatment assistance: Background and formula history. CRS Report RL31073, Congressional Research Service, Washington, D.C.

Ramseur, J. L. (2023, March). Clean water state revolving fund allotment formula: Background and options. CRS Report R47474, Congressional Research Service, Washington, D.C.

Ramseur, J. L. (2025, June). Wastewater infrastructure funding: Background and affordability issues. CRS Report R48565, Congressional Research Service, Washington, D.C.

Raven, S. J. and J. C. Coulson (2001). Effects of cleaning a tidal river of sewage on gull numbers: a before-and-after study of the River Tyne, northeast England. *Bird Study 48*(1), 48–58.

Rosenberg, K. V., A. M. Dokter, P. J. Blancher, J. R. Sauer, A. C. Smith, P. A. Smith, J. C. Stanton, A. Panjabi, L. Helft, M. Parr, and P. P. Marra (2019). Decline of the North American avifauna. *Science 366*(6461), 120–124.

Ruhl, H. A. and N. B. Rybicki (2010). Long-term reductions in anthropogenic nutrients link to improvements in Chesapeake Bay habitat. *Proceedings of the National Academy of Sciences 107*(38), 16566–16570.

Seston, R. M., M. J. Zwiernik, T. B. Fredricks, S. J. Coefield, D. L. Tazelaar, D. W. Hamman, J. D. Paulson, and J. P. Giesy (2009). Utilizing the great blue heron (Ardea herodias) in ecological risk assessments of bioaccumulative contaminants. *Environmental Monitoring and Assessment 157*(1-4), 199–210.

Shapiro, J. S. (2022). Pollution trends and US environmental policy: Lessons from the past half century. *Review of Environmental Economics and Policy 16*(1), 42–61.

Solow, A., S. Polasky, and J. Broadus (1993). On the measurement of biological diversity. *Journal of Environmental Economics and Management 24*(1), 60–68.

Stiefel, H. J. (1994, February). Municipal wastewater treatment: Privatization and compliance. Policy Study 175, Reason Foundation, Los Angeles.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling (2009). ebird: A citizen-based bird observation network in the biological sciences. *Biological conservation 142*(10), 2282–2292.

Taylor, C. A. and H. Druckenmiller (2022). Wetlands, flooding, and the Clean Water Act. *American Economic Review 112*(4), 1334–1363.

Tobias, J. A., C. Sheard, A. L. Pigot, A. J. M. Devenish, et al. (2022). AVONET: morphological, ecological and geographical data for all birds. *Ecology Letters 25*(3), 581–597.

U.S. Census Bureau (2020). QuickFacts: Land area in square miles, 2020, for New York County, New York, Boston city, Massachusetts, and Dallas city, Texas. https://www.census.gov/quickfacts. Accessed October 26, 2026.

U.S. Code (2018). Title 33, chapter 26: Water pollution prevention and control. United States Code, 2018 edition, 33 U.S.C. §§ 1251 to 1389.

U.S. Department of the Interior, U.S. Fish and Wildlife Service (2023, sep). 2022 national survey of fishing, hunting, and wildlife-associated recreation. Technical report, U.S. Fish and Wildlife Service, Washington, D.C. Title page dated September 2023 (the file was posted October 2023); the year field read 2022 until 2026-10-03, when the title page was read. Fielded by NORC for the Service under Multistate Conservation Grant F20AP00134; the first survey after the Census Bureau's withdrawal, with a new methodology the report says is not directly comparable to earlier surveys. Program page: https://www.fws.gov/program/national-survey-fishing-hunting-and-wildlife-associated-recreation-fhwar.

U.S. Environmental Protection Agency (1979, October). National municipal policy and strategy for construction grants, NPDES permits, and enforcement under the Clean Water Act. Technical Report EPA 430/9-79-502, U.S. Environmental Protection Agency, Washington, D.C. NEPIS document P1005D0R.

U.S. Environmental Protection Agency (1980). Handbook of procedures: Construction grants program for municipal wastewater treatment works. Second edition, U.S. Environmental Protection Agency, Office of Water Program Operations, Municipal Construction Division, Washington, D.C. NEPIS document 9101S3HC, catalogued as EPA 832-R-80-114; the edition Flynn and Smith (2022) cite.

U.S. Environmental Protection Agency (1982, December). 1982 needs survey: Cost estimates for construction of publicly-owned wastewater treatment facilities. Technical Report EPA 430/9-82-009, U.S. Environmental Protection Agency, Office of Water Program Operations, Washington, D.C. Dated December 31, 1982; NEPIS document 2000GEHT.

U.S. Environmental Protection Agency (1984). National municipal policy on publicly-owned treatment works. 49 Fed. Reg. 3832 (January 30, 1984).

U.S. Environmental Protection Agency (1985, February). 1984 needs survey: Report to Congress. assessment of needed publicly owned wastewater treatment facilities in the United States. Technical Report EPA 430/9-84-011, U.S. Environmental Protection Agency, Office of Municipal Pollution Control, Washington, D.C. NEPIS document 2000GGUV.

U.S. Environmental Protection Agency (1988, January). Initial guidance for state revolving funds. Technical Report EPA 832-B-88-102, U.S. Environmental Protection Agency, Office of Municipal Pollution Control, Washington, D.C.

U.S. Environmental Protection Agency (2017). Grants information and control system data fields. Codebook distributed with the construction-grant record.

U.S. General Accounting Office (1980, November). Costly wastewater treatment plants fail to perform as expected. Report to the Congress by the Comptroller General CED-81-9, U.S. General Accounting Office, Washington, D.C.

U.S. General Accounting Office (1994, September). Water pollution: Information on the use of alternative wastewater treatment systems. Technical Report GAO/RCED-94-109, U.S. General Accounting Office, Washington, D.C.

U.S. Geological Survey (2026). North American breeding bird survey. Eastern Ecological Science Center, Laurel, Maryland, https://www.usgs.gov/centers/eesc. The survey's overview page, accessed October 26, 2026.

van Kippersluis, H. and C. A. Rietveld (2018). Beyond plausibly exogenous. *The Econometrics Journal 21*(3), 316–331.

Wardle, A. R. (2025). The recreational value of rare species: Causal evidence from the cassia crossbill. *Journal of the Association of Environmental and Resource Economists 12*(5), 1133–1164.

Wilman, H., J. Belmaker, J. Simpson, C. de la Rosa, M. M. Rivadeneira, and W. Jetz (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology 95*(7), 2027.

Wooldridge, J. M. (1997). Multiplicative panel data models without the strict exogeneity assumption. *Econometric Theory 13*(5), 667–678.

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics 90*(1), 77–97.

Wooldridge, J. M. (2023). Simple approaches to nonlinear difference-in-differences with panel data. *Econometrics Journal 26*(3), C31–C66.