THE WATER QUALITY IMPACTS OF CRITICAL HABITAT DESIGNATION FOR ENDANGERED SPECIES

by

Taurey Rosenhahn Carr

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Economics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2023

DEDICATION

I dedicate this thesis to the many amazing family members, teachers, professors, and others who helped me reach this goal. Mom, Dad, Jordan, thank you for setting the best example for me to follow. I can't express how important you all are in everything that I do. Fallyn, thank you for supporting me in so many ways throughout my education. Buster, thank you for your helpful feedback and consistency. Señora Spurlock, gracias por mostrarme la importancia de la lengua. Jason Dayton, thank you for showing me that science and art are complimentary, not opposing. Thank you Thorsten Janus and Sasha Skiba for your instruction and for encouraging me to pursue graduate education. J, thank you for challenging me to think better and for always keeping your door open. Steve Newbold, thank you for helping me to become a better researcher and for giving so much of your time to my development. Thank you to so many more—I am here because of you.

ACKNOWLEDGEMENTS

I would like to thank Dr. Melissa LoPalo, my thesis chair, for her constant support and invaluable advice throughout this process. I could not have received better guidance during my research—thank you immensely. Thank you as well to Dr. Nick Hagerty and Dr. Justin Gallagher for your time and assistance. You both helped me to become a better researcher. Thank you Sadiq, Jadon, Laura, Brock, Katie, Olivia, Justin, Hannah, and Trevor. Your humor, intelligence, and thoughtfulness have contributed greatly to my experience in graduate school.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

Table Page

LIST OF FIGURES

GLOSSARY

BOD	Biochemical oxygen demand
CDC	Centers for Disease Control and Prevention
CH	Critical habitat
CWA	Clean Water Act of 1972
DiD	Difference-in-differences
DO	Dissolved oxygen
EPA	Environmental Protection Agency
ESA	Endangered Species Act of 1973
FWS	Fish and Wildlife Service
HUC	Hydrologic unit code
MOU	Memorandum of understanding
MPN	Most probable number
NMFS	National Marine Fisheries Service (also referred to as NOAA Fisheries)
NOAA	National Oceanic and Atmospheric Administration (see NMFS)
NPDES	National Pollutant Discharge Elimination System (an EPA water quality permitting program established by the CWA)
TSS	Total suspended solids
TWFE	Two-way fixed effects
USGS	United States Geological Survey
WQP	Water Quality Portal

ABSTRACT

The Endangered Species Act of 1973 [ESA] is well-known by environmental economists for its extensive provisions that create a variety of impacts on housing, land development, timber harvesting, etc. However, the ESA's impact on water quality has not been formally studied despite being discussed extensively by federal agencies that administer the Act. I estimate the causal effect of critical habitat designation, an ESA provision that regulates land use, on a range of water quality outcomes. Using administrative data on water quality from 1970-2018, I employ event study and difference-in-differences [DiD] empirical models to evaluate temporal and spatial changes in water quality resulting from plausibly exogenous variation in critical habitat designations. I find null results for most water quality outcomes and mixed evidence of a decrease in pH after designations occur. However, pooled DiD results find no evidence of average declines in pH in the years following designation. Slight declines in pH from the event-study results are concentrated partially in urban areas and primarily around critical habitat designations involving fish species. Results provide some evidence that fish designations may result in more significant water quality impacts after designation across pH and additional outcomes than all designations on average. These results add to a body of research that questions if other species conservation provisions may lead to more efficient outcomes than critical habitat designation.

INTRODUCTION

In the 50 years since the Endangered Species Act of 1973 [ESA] was enacted, proponents and critics have debated its merits from both ideological and economic perspectives. Even prior to the passage of the ESA, Krutilla (1967) noted that endangered species represent a classic market failure by creating externalities due to their divergent social and private values. While the ESA—deemed perhaps the most extensive environmental law in U.S. history (Brown & Shogren, 1998)—has aimed to solve this market failure, it has also repeatedly stoked concerns over private property rights, the high potential costs of far-reaching provisions to protect endangered species, and the value of biodiversity protection. Accordingly, economists have contributed to the debate by quantifying various ESA costs and benefits, documenting behavioral responses to ESA policies, and highlighting the efficiency trade-offs of various policy tools aimed at species conservation. Still, many effects of the ESA remain unknown, including its impact on water quality. This is somewhat surprising given that the ESA explicitly aims to "resolve water resource issues in concert with conservation of endangered species" (Endangered Species Act of 1973). Critical habitat designation, a key ESA land use provision, potentially alters water quality by influencing land development, which has been shown to negatively impact water quality. While the main link between critical habitat and water quality is through land development, critical habitat may also alter water pollution regulations or complement recovery activities that improve water quality. Ultimately, given that water quality improvement is a key goal of the ESA, it is essential from a public policy perspective to improve understanding of the causal relationship between the two.

This paper measures the causal effect of critical habitat designation on eleven measures of water quality. I use event study and difference-in-differences [DiD] frameworks to leverage variation in protections for ESA species from geographical boundaries and across time. In the main specification, I utilize cross-sectional comparisons between treated and control water stations within the same state and calendar year as well as changes in treated stations before and after designation. The main results indicate that, for most water quality outcomes, critical habitat designation does not significantly alter water quality. Standard error estimates for post-event year indicators suggest that these results are precise null findings, signaling that critical habitat designation does not lead to any large water quality impacts. pH experiences statistically significant but small declines 6-8 years after designations occur. It is difficult to determine if a decrease in pH represents an increase or decrease in water quality since pH values that are lower or higher than 6.5 or 8, respectively, may adversely impact species (U.S. EPA, 2017); however, given the mean pH for the sample period of roughly 7.5 units and the magnitude of the effect, the result likely does not indicate a meaningful increase or decrease in water quality. The pooled DiD results, which compare all pre-designation years to the entire post-designation period instead of evaluating year-over-year treatment effects, support this conclusion since they find no significant decrease in pH caused by the designation during the post-period. These results suggest there may be a lagged but non-lasting effect of critical habitat designation on pH, hinting that mechanisms that impact pH may take many years to produce notable effects after designation and may not result in long-run effects.

Additionally, I explore patterns in the sampling of water stations to examine the influence of sampling location and frequency in the unbalanced water quality panel on dissolved oxygen and pH results. I find that both dissolved oxygen and pH are less likely to be sampled in critical habitat areas after designations are published even while controlling for baseline levels of sampling at each station. To provide confidence that this under-sampling of treated stations does not drive the main results, I estimate the main specification using a sub-sample of treated stations that are sampled in all event years. The results remain qualitatively consistent compared with the main specification.

One potential concern with the baseline results is that water stations close to critical habitat could receive spillover water quality effects from nearby designations. If this is the case, the presence of spillovers could attenuate the estimated treatment effects towards zero because some control stations would be partially treated, making my analysis less likely to identify a significant treatment effect. To account for this potential concern, I extend the model by removing stations most likely to receive spillover effects from designations and find no notable changes in the main results. I also employ alternate cross-sectional comparisons using only control stations in watersheds immediately surrounding critical habitat designations. In this alternate specification, pH results are noisier six years after designation but remain significant after eight years. Results for dissolved oxygen change slightly to indicate decreases cased by critical habitat of up to 2.6% saturation eight years after designation, but this effect is inconsistent across specifications.

Finally, I explore the effects of critical habitat on pH separately for urban and rural stations and for aquatic and fish species designations. I find that the mild pH effects are concentrated in urban water monitoring stations and particularly among fish designations. One potential mechanism could be that critical habitat causes declines in road construction, which is often concentrated more in urban areas than rural areas and produces pollutants that are alkaline (potentially explaining mild surface water acidification in critical habitat areas). Still, additional analysis is required to confirm this speculation about potential mechanisms. I also analyze results from pooled DiD specifications for other main outcomes in fish designations and find that total phosphorus and total suspended solids experience significant increases at 5% and decreases at 10% significance levels after designation, respectively. Since fish designations are often more directly concerned with water quality improvements than other designations, it is unsurprising that these designations produce more significant results than all designations combined. However, these effects remain small in magnitude.

This paper to provides four main contributions to the current literature. First, it explores a potential effect of the ESA that has not been formally studied. Providing an estimate of the water quality impacts of critical habitat allows for a more complete cost-benefit analysis of the designation. Second, this study provides additional evidence for the causal link between ESA land use policy and water quality, informing a more general question underlying the analysis: how does human activity adversely impact water systems, and do current management approaches lead to better social outcomes? Third, the study generates a detailed, 49-year panel of water quality observations with data previously used sparingly in economics research. This data is merged with weather and ESA critical habitat boundaries data to form a novel dataset which can be used in further studies or merged with other national-level geospatial data to explore various economic research questions related to water quality. Lastly, this paper examines the effects of critical habitat designation at a broad national scale, in contrast to many current studies that look at single designations or designation cohorts[1], to generate estimates with relevance beyond a single geographic area or species designation.

[1] A cohort is a group of critical habitat designations that occur in the same year.

BACKGROUND

The Endangered Species Act of 1973

The ESA establishes the authority of the U.S. Fish and Wildlife Service [FWS] and the National Oceanic and Atmospheric Administration's [NOAA] National Marine Fisheries Service [NMFS] to list (or unlist) species as threatened or endangered and provides protections and support for the recovery of these species. Listing decisions generally begin with a third party petition to list a species or an inquiry from one of the federal agencies, and a species is listed based on the discretion of the agencies according to the "best scientific and commercial data available" (Endangered Species Act of 1973). Species can be listed if they are affected by any of the following factors: "the present or threatened destruction, modification, or curtailment of its habitat or range; overutilization for commercial, recreational, scientific, or educational purposes; disease or predation; the inadequacy of existing regulatory mechanisms; or other natural or man-made factors affecting its continued existence". When a species is listed, it receives ESA protections from "takings."[1] Listed species often receive critical habitat and recovery plan actions under the ESA that take effect in the years following listing.

Section 7 of the ESA outlines the provision of critical habitat for listed species, which are areas deemed "essential" to a species' conservation and managed under the authority of the FWS and NMFS. The multi-step process for creating designations is conducted as follows: (a) a species is listed in the federal register; (b) a determination of critical habitat is made based on physical and biological necessities for species' conservation; (c) a critical habitat proposal is made publicly, followed by a period of public comment and potential modifications to the habitat; and (d) a final designation is announced (U.S. FWS, 2017).

[1] "Taking" refers to "any action which injures or kills an endangered creature, or significantly adversely modifies the habitat of an endangered species" (Innes, Polasky, & Tschirhart, 1998).

Critical habitat may not be designated for a listed species in cases where habitat designation is not "prudent" or "determinable" (Endangered Species Act of 1973). At the beginning of 2023, over 800 unique final critical habitat designations covering just over 168,000 square miles and 36,060 river miles of the U.S. and surrounding oceans—around 4% of total U.S. land area—were active (U.S. FWS, 2022).

Critical habitat designation is not the only tool afforded to the federal government by the ESA for species protection, but it is notable aside from other protections for listed species under the act.[2] Primarily, critical habitat areas delineate geographic boundaries that impact development activities that encounter a federal nexus[3] "and are likely to destroy or adversely modify critical habitat" in habitat areas (U.S. FWS, 2017). Specifically, any of these activities must undergo a section 7 consultation with the FWS or NMFS. The consultation process begins with an informal consultation, a review of the proposed development project—this includes a review of actions proposed by the project and a determination of whether the activity will likely adversely impact listed species and/or their habitat. If the FWS or NMFS determines that the activity may adversely impact listed species, a formal consultation is conducted to determine the specific impacts of the project on listed species and their habitat. The process ends in a biological opinion that includes whether the agency has ensured its actions will not adversely impact the species, proposed conservation measures, and other actions that minimize the risk of adverse impacts. The purpose of the consultation process is to ensure any activities approved or conducted by a federal agency do not adversely impact listed species and their habitat. For landowners and developers, this process could increase development costs by adding additional time to projects or requiring alterations to proposed development.

The requirement for developers to consult with the FWS or NMFS is the primary

[2]Main provisions for protecting listed species include the establishment of critical habitat, recovery plans, and prohibitions on "taking" of listed species.

[3]A federal nexus includes any activity which requires funding or permitting from a federal agency.

mechanism by which critical habitat protects species beyond other ESA provisions; however, designation may also elicit a number of indirect effects within habitat boundaries. First, the ESA's prohibition on the "take" of a species, which includes adverse modification of species' habitat, may deter development activities on designated land which could present a potential risk to the species and, consequently, the landowner. This may be especially true if landowners on critical habitat properly assess the potential costs of "taking" a species, which are significant according to Brown and Shogren (1998). While "taking" a listed species is illegal both inside and outside of critical habitat boundaries, critical habitat may augment the perceived risk of development for landowners within habitat boundaries after designation for two reasons: there may be an increased presence of the endangered species compared to non-designated areas, and the "take" prohibition includes adverse modification of habitat which may be less relevant in non-designated areas. This heightened perceived risk of "taking" in critical habitat areas is highlighted in the U.S. Supreme Court case *Babbitt v. Sweet Home Chapter of Communities for a Great Oregon (1995)*, in which a group of logging companies, landowners, and timber workers sued the Secretary of the Interior to challenge the inclusion of adverse habitat modification in the definition of "take." The plaintiffs had forgone logging and commercial activities due to concerns of adverse modification of nearby critical habitat for the Red-cockaded woodpecker and Northern spotted owl. Similarly, the risk of "takings" in critical habitat has been shown to provide an adverse incentive for landowners concerned with potential future designation to preemptively destroy Red-cockaded woodpecker habitat (Wood & Watkins, 2020). In general, anecdotal evidence suggests that perception of "take" prohibitions are heightened by critical habitat.

There are two additional indirect effects that critical habitat produces beyond other ESA provisions. First, critical habitat designations may provide a legal pathway for individuals or organizations with conservation interests to halt development activities (Friedman, 2022; Radford, 2022; Wernik, 2022). Lastly, designations may depress land values or make land parcels less attractive to developers who may otherwise carry out land development projects (Auffhammer, Duru, Rubin, & Sunding, 2020; Melstrom, 2021). If demand for undeveloped land decreases due to critical habitat, this could lower the overall amount of development in critical habitat boundaries. Developers would substitute development activities to comparably more attractive parcels outside critical habitat boundaries. Landowners could also alter existing land use behavior in critical habitat areas if existing activities present a heightened risk to listed species due to designation. For instance, if a landowner previously conducted logging operations on land that becomes designated as critical habitat, they may decide to abate this behavior in the future if logging could negatively impact a listed species (resulting in legal exposure to the landowner), even if no section 7 consultations are required to carry out the activity. Still, the ESA does not explicitly regulate existing land use or land use that does not encounter a federal nexus, so these hypothetical impacts require empirical evidence to determine how landowner decisions are indirectly impacted by critical habitat. Ultimately, however, designation makes development more costly for landowners through increased likelihood of section 7 consultations and potential risks of legal exposure for adversely modifying critical habitat.

The complex implications of designation have led policymakers to consider the economic costs of critical habitat. While the ESA originally forbade economic considerations from factoring into decisions about declaring critical habitat, amendments made to the bill in 1978 required economic considerations be made during the designation process (U.S. FWS, n.d.).[4] Additionally, many economics studies have specifically evaluated the effects of the critical habitat provision of the ESA due to its high perceived costs. Still, quantifying the effects of critical habitat has proven difficult even for the FWS and NMFS. For instance, Plantinga, Helvoigt, and Walker (2014) evaluate a number of case studies on the agencies' calculation of economic costs of critical habitat designations and find that while the processes

[4]Economic considerations are not made during the listing process.

used to evaluate costs are sound, there exists great uncertainty surrounding precision of the estimated costs.[5] Additionally, the authors report that economic considerations did not play a significant role in any designation exemptions that were evaluated. Thus, the distribution of the costs and benefits of these policies remain relatively unknown.

Water Quality and Links to the ESA

Water pollution has been among the United States' top environmental concerns for decades. Surveys suggest Americans are more concerned about water quality than other threats such as climate change, biodiversity loss, air pollution, etc. (Gallup, 2022). The Clean Water Act [CWA] was enacted in 1972 to address these concerns. Since then, the CWA has improved many water quality measures, but more than half of U.S. surface waters continue to violate environmental standards (Keiser & Shapiro, 2019a, 2019b). The main mechanisms for improving water quality in the CWA mainly target point sources of pollution;[6] however, the act largely evades regulation of non-point pollution sources that are impacted by land use decisions.[7] Thus, policies aimed at improving water quality through land use may provide unique advantages compared to current U.S. regulations.

Not only does ESA critical habitat provide an example of such a land use policy, but the ESA and the CWA share similarities in the specific goals they aim to accomplish. Namely, the ESA acknowledges that water quality improvement is essential to endangered species protection. Embedded in section 2 of the ESA is a call for regulators to leverage species protection as a mechanism for assisting in efforts to address water issues; it states: "federal agencies shall cooperate with state and local agencies to resolve water resource issues in

[5]The types of "costs" evaluated by policy-makers have changed over time and currently include modifications to potential development projects and administrative costs of managing critical habitat

[6]Point sources are easily identifiable, singular sources of pollution, while non-point sources come from larger areas from which there is no discernible, single source of pollution discharge. Examples of pollution include wastewater discharge from a municipal treatment plant (point source) and agricultural runoff (non-point source).

[7]Keiser and Shapiro (2019a) note that agriculture and other non-point sources are affected by land use.

concert with conservation of endangered species" (Endangered Species Act of 1973). The link between the ESA and water quality is also embedded in section 3(5)(A)(i) of the Act, which defines critical habitat as areas occupied by a listed species "on which are found those physical or biological features essential to the conservation of the species and which may require special management considerations or protection" (Endangered Species Act of 1973).[8] As water is a key physical and biological element to any ecosystem, its importance for species conservation is readily apparent. Indeed, numerous final rulings on critical habitat decisions list water or water quality among a handful of "primary constituent elements" that define the essential features of a species' habitat. For instance, the final critical habitat ruling for the Fluted Kidneyshell lists five "principal constituent elements" necessary to the species' conservation, and water quality with low pollutants and satisfactory temperature, pH, dissolved oxygen, hardness, and turbidity levels is listed as one key element (U.S. FWS, 2013). Thus, water quality outcomes are not merely unintended side effects of critical habitat designations; rather, they are necessary considerations of designation itself.

Understanding the link between the ESA and water quality is important for many reasons. First, the relationship may highlight whether the ESA provides any useful policy levers to enhance water quality. Second, understanding the water quality effects of the ESA may provide insight into the effectiveness of the act's provisions. As indicated by Langpap, Kerkvliet, and Shogren (2018), the effectiveness of the ESA remains unclear for several reasons. If the ESA improves water quality, this could suggest that it positively impacts ecosystems. Alternatively, if the ESA is found to harm water quality, this finding would shed light on perverse incentives to develop land before final critical habitat designations are published, some of which are documented in the literature. Finally, the link between the ESA and water quality contributes to debate on the key question from the past half-century:

[8]Critical habitat is also defined in this section as areas not occupied by a listed species but determined to be necessary to the conservation of the species.

are the high costs of the ESA worth it? While water quality improvement is not the primary goal of the ESA, it does relate closely with biodiversity protection.

While the motivation exists to explore this relationship, the question remains: even if water quality and critical habitat designations are associated, is there reason to suspect that designations may cause water quality outcomes to change? Academics and federal agencies provide reasons to believe so. Nelson, Withey, Pennington, and Lawler (2017) mention that the ESA may create costs to landowners through restrictions on water use, but similar mechanisms may also impact the quality of water resources in critical habitat boundaries. For instance, the authors mention the possibility of water withdrawals in critical habitat boundaries receiving additional regulation; it could also be possible for critical habitat designations to impact (or at least delay) the success of receiving federal permits for water pollution discharge. Additionally, federal agencies that administer the CWA and the ESA acknowledge the overlap between water quality and endangered species protection, as evidenced by a memorandum of understanding [MOU] published by the agencies (U.S. EPA, 2001). The document highlights manners in which the ESA may impact water quality, including through enforcement of water discharge permitting standards that align with critical habitat protection. The MOU indicates that water quality is a known and intended aspect of critical habitat protection, and it leaves open the possibility that critical habitat may alter water quality in designated areas. Furthermore, effects of critical habitat designation on land use may also impact water quality. I explore these possibilities in the next section.

CONCEPTUAL FRAMEWORK

There are several causal channels through which critical habitat may impact water quality. Since critical habitat is, at its core, a land use policy, the most likely mechanism is through land development. However, other potential mechanisms include the designation's effect on pollution regulation and recovery activities in habitat areas. More broadly, critical habitat is known to evoke a range of behavioral responses, and there are numerous agents whose responses to designation may impact water quality, including private landowners and corporations, federal regulators, third-party environmental consultants, and conservation activists.

Given the variety of potential mechanisms and the many possible directions of potential effects, empirical investigation is needed to provide evidence on the likely degree to which any mechanism described below is present. While the most likely impact of critical habitat would be an increase in water quality after designation, if there is an effect at all, this hypothesis requires additional empirical evidence. It should also be noted that the mechanisms above may differentially target point-source and non point-source pollution. For instance, if the primary mechanism at play is land development, critical habitat designation may be effective at targeting non-point source pollution. Alternatively, if critical habitat primarily impacts water quality by augmenting water quality regulation, its effect is likely concentrated among point-sources of pollution. The data used in this analysis is not well-equipped to distinguish the differential impacts of critical habitat on point and non-point source pollution, so further investigation beyond the scope of this study of the first-stage relationship between critical habitat and these mechanisms would be needed to make such a determination.

Land Development as a Potential Mechanism

The most likely mechanism through which critical habitat may impact water quality is through changes in land development or land use. Studies are inconclusive about the exact impact of critical habitat on land development, so there are three theoretical impacts that could occur. First, if critical habitat decreases the overall rate of development, similar to the finding by Zabel and Paterson (2006), literature linking land development and water quality would indicate that water quality should increase following designation. Second, if, critical habitat increases the rate of development, perhaps through perverse incentives highlighted in studies from List, Margolis, and Osgood (2006); Lueck and Michael (2003), a decrease in water quality around the time of designation would be expected. Third, if designation does not significantly impact land development, as suggested by the results of Nelson et al. (2017), there would be no identifiable effect on water quality from this mechanism.

To further examine these three potential behavioral response of developers, first consider the net present value of future earnings for a private landowner that must choose between developing or preserving an empty land parcel. As an example, suppose the landowner considers whether or not to convert a parcel with natural forest cover to agricultural land that can be used as an input in the production of crops. The landowner considers the potential benefits of development—primarily, the discounted future earnings from producing crops on the converted parcel and from selling harvested timber. Additionally, they consider the costs of development—the labor required to convert the parcel, the additional machinery that must be purchased to convert the land and farm crops, etc. Assume that the land conversion produces an environmental externality which decreases water quality within the parcel but does not factor into the private costs considered by the landowner. When the private landowner makes this choice outside of critical habitat boundaries, they will choose to develop the parcel if the future revenues from conversion exceed the costs of development. However, development incurs additional costs when the choice is made inside of critical habitat boundaries. Namely, critical habitat introduces both the risk of "taking" an endangered species by destroying its habitat, and it introduces potential increased costs by triggering a section 7 consultation if any development requires federal agency action. Such consultations may require alterations to the development project that could require additional capital, or they could increase the completion time of the development project, which may decrease future revenues or result in higher labor costs.

Linking this example back to the potential impacts of critical habitat on land development, there are three potential outcomes depending on additional assumptions about the timing of the decision with respect to designation and the influence of critical habitat on the costs of development. Assuming the landowner does not make their choice until designation occurs and that the costs imposed by critical habitat are large (either the landowner perceives the risks of "takings" as large or the project would trigger a section 7 consultation), we would expect development to decrease within habitat boundaries. In this scenario, it becomes more likely that the landowner chooses not to develop due to higher costs. In the aggregate, lower rates of development produce fewer externalities, which improves water quality. Alternatively, assume the landowner makes their choice prior to designations taking effect, but the costs imposed by critical habitat remain large. Now, the landowner may mitigate some costs by developing the parcel before critical habitat is officially designated. In this way, the landowner may avoid destruction of critical habitat because it is not yet designated, and they may avoid the need for section 7 consultations. Given this new assumption about the timing of landowner decisions, it becomes more likely that the landowner develops earlier since the costs increase over time. Under this scenario, water quality would decrease around the time of designation due to increases in preemptive development. Lastly, assume that the costs imposed by critical habitat are not sufficiently large enough to outweigh the benefits of development. Perhaps landowners do not perceive the risks of "taking" as large, or perhaps their actions do not require approval from other federal agencies. Regardless of the timing of the development decision, we would expect to see no identifiable change in development, and consequently water quality, between parcels inside and outside of critical habitat since the choice remains unchanged by designation.

Increased Water Quality Regulation as a Potential Mechanism

Critical habitat may also improve water quality by increasing regulatory scrutiny in designated areas. Specifically, section 7 consultations required by the ESA may provide additional opportunities for water quality regulators to enforce water quality standards that protect listed species. Consultation between water quality regulators and the FWS or NMFS often results in biological opinions on existing water quality standards that determine if standards sufficiently abate adverse impacts on aquatic species listed under the ESA. For instance, Carlson, Foo, Burns, and Asner (2022) note that section 7 consultations under the ESA can be leveraged to protect aquatic ecosystems housing coral reef species.

Anecdotal evidence provides additional support that these ESA provisions are leveraged by regulators. For instance, the EPA considers critical habitat when reviewing and granting Clean Water Act NPDES pollution permits (U.S. EPA, 2001). NPDES permits are required to lawfully discharge any pollutant into U.S. waters through a point source, and they set standards for acceptable levels of pollutants allowed under the CWA. Rosan (2000) notes the relationship between critical habitat and NPDES permitting, finding that the EPA regularly completes section 7 consultations with the FWS and NMFS during triennial reviews of state water quality standards to ensure that these standards comply with species conservation goals in critical habitat areas. Under this process, individual water quality permit applications deemed by the FWS or NMFS to present adverse impacts to listed species are reviewed by the EPA and can even be contested using existing CWA regulatory power. Furthermore, state agencies may also adopt stronger water quality regulations in response to critical habitat designation. For example, in Oregon's Willamette Basin, specific water quality standards that prohibit measurable increases in temperature caused by human activity exist only for river segments with threatened and endangered species (Rosan, 2000). Additionally, the NMFS completed a section 7 consultation and issued a biological opinion on the Oregon Department of Environmental Quality's water quality standards relative to potential effects on listed salmon species and their critical habitat. In general, these procedures highlight the coordination between the EPA—the primary federal water regulatory body—and agencies that administer the ESA. In places with critical habitat, water quality regulation may consequently be enforced with additional effort and scrutiny than in non-designated areas.

Still, while numerous examples of regulatory cooperation between ESA agencies and water quality regulators exist, Rosan (2000) also indicates that current practices may not exercise enough regulatory authority to adequately protect listed species. Thus, it remains possible that this mechanism does not support a strong enough causal relationship between critical habitat and water quality regulation to produce a noticeable effect.

To further evaluate this potential mechanism, consider the example of wastewater facilities regulated by the EPA through a CWA discharge permit. Assume that the stringency and/or enforcement of water quality standards outlined in the CWA permit depends on the vulnerability of nearby aquatic species. That is, assume the EPA is more likely to introduce stronger standards if aquatic ecosystems nearby are sensitive due to the presence of threatened or endangered species. Additionally, assume the EPA allocates resources equally among all permits it regulates, such that it has equal information about aquatic ecosystems in each permit location. Critical habitat for aquatic species acts both as a signal that water quality impacts vulnerable species and as a regulatory obligation for the EPA to fulfill through consultation. Outside of critical habitat areas, the EPA regulates discharge permits according to CWA guidelines. In critical habitat areas, in addition to these same CWA regulations, the EPA is also required to consult the FWS or NMFS about water quality standards and permitting to ensure that its actions do not harm listed aquatic species. Two potential outcomes emerge. If section 7 consultations regularly identify lapses in water quality standards that harm listed species, and if these lapses are subsequently corrected and enforced by the water quality regulators, wastewater facilities in critical habitat areas are allowed to discharge less pollutants than they otherwise would be. Water quality would subsequently improve more inside critical habitat areas than outside. If, however, section 7 consultations do not results in substantial changes in water quality standards, no change would exist. This could occur for several reasons. Perhaps water quality standards are already sufficiently high to protect listed species in most cases. Instead, perhaps recommendations made by ESA authorities are not enforced by water quality regulators, such that wastewater facilities discharge similar levels of pollutants regardless of designation.[1] Lastly, perhaps agency cooperation does improve water quality immediately surrounding wastewater facilities, but the aggregated effects are very small due to the greatly limited scope of water quality regulation of non-point source pollution.

Increased Recovery Efforts as a Potential Mechanism

Lastly, it is possible that critical habitat improves water quality by increasing available resources for water quality recovery plans. Under section 4 of the ESA, a recovery plan must be created for most listed species that details: a) site-specific actions taken to aid in species recovery, b) measurable criteria that would result in the recovery of a listed species, and c) estimates of time and cost required to complete recovery actions (Endangered Species Act of 1973, 2003). Examples of recovery plan actions include measures such as meeting with private landowners and local officials to discuss best management practices for preserving water quality, conducting additional research on species threats, and introducing species in

[1]For instance, Rosan (2000) highlights that water quality regulations in Oregon are deemed insufficiently enforced by some due to political pressures despite NMFS opinions that call for stronger water quality standards.

unoccupied locations. Furthermore, under section 5, the FWS or NMFS is authorized to purchase land to aid in species conservation. Land acquisition is an integral action in many species' recovery plans (U.S. FWS Kentucky Ecological Services Field Office, 2022) and is even supported by a FWS grant program that provides funding for acquisitions made by states or territories that contribute to endangered species conservation (U.S. FWS, 2023). Given that recovery plan measures are site-specific, it is likely that many efforts would be concentrated within critical habitat areas. Indeed, recovery strategy documents for listed species often allude to the cooperation between recovery plans under section 4 and critical habitat under section 7 of the ESA, indicating that recovery efforts often are focused within species' occupied habitat ranges (U.S. FWS Kentucky Ecological Services Field Office, 2022). Thus, water quality may improve more in critical habitat areas compared to non-designated areas if recovery actions that improve water quality are more likely to be conducted within critical habitat instead of outside these boundaries.

To illustrate the behavioral response to land acquisition, one of many recovery plan actions, let us again consider a landowner who chooses between developing and keeping undisturbed a forest land parcel within critical habitat boundaries. Assume the cost-benefit calculation made by the landowner is the same as discussed in the first section of this chapter; however, now that their property falls within critical habitat, it is valued higher by ESA authorities. The landowner now has a third choice: convert the parcel to cropland, leave it as forest land, or sell their property to the government, who wishes to acquire the critical habitat parcel for the purpose of species conservation. If the ESA regulatory agencies value the land parcel more than the net present value of benefits from converting or leaving undisturbed the land parcel, the landowner will sell the property to regulators. In turn, since their primary objective is to protect listed species, ESA regulators will likely avoid development activities that lower water quality, solving the market failure of water pollution posed by the private landowner choice.

LITERATURE REVIEW

Effects of ESA Critical Habitat Designation

Many economists have begun to address the uncertainty surrounding the costs and benefits of ESA provisions by exploring a variety of potential outcomes of designation and listing. The primary relationship of interest to this study is the effect of critical habitat designation on land development. Studies have examined the relationship with numerous outcome variables and in several settings, but the literature has not reached a consensus on the overall effects. A number of papers have examined land development directly preceding designation, and many indicate that landowners preemptively develop to avoid future regulation or designation altogether. As mentioned previously, Lueck and Michael (2003) and Byl (2019) find that timber plots closer to Red-cockaded woodpeckers, a species listed under the ESA, are more likely to be harvested younger, presumably to avoid future critical habitat designation and increased regulatory scrutiny. Additionally, List et al. (2006) find that land development accelerates by one year on average directly preceding critical habitat designation to finish development projects before regulations tighten.[1] Papers that examine land development outcomes after designation uncover mixed results. For instance, Zabel and Paterson (2006) examine the response in housing permits after critical habitat designations using a unit fixed-effects estimator. They find that housing permits decline by almost a quarter in the short run and more than a third in the long run, indicating that critical habitat signals higher costs to developers. Alternatively, Nelson et al. (2017) employ a matching estimator and accompanying OLS regressions (with control polygons derived from matching methods) to examine the effect of critical habitat designation on land cover change using pooled cross-sections from the contiguous U.S. in 1992, 2001, and 2011. The authors

[1] Additional work by Innes et al. (1998); Newburn, Brozovic, and Mezzatesta (2011); Zhang and Flick (2001) highlights the potential for critical habitat to create perverse incentives, similar to the results found by Lueck and Michael (2003) and List et al. (2006).

find no average impact on land cover change, suggesting that critical habitat designation may not alter the opportunity cost of land development compared to non-designated areas.

Other analyses have evaluated the impact of critical habitat designations on land values. In general, the literature forms a consensus that critical habitat designations decrease land values inside critical habitat boundaries, albeit to varying degrees. The literature on land values provides insight into the behavioral responses of landowners and developers on land parcels within critical habitat boundaries. These housing market effects may provide insight into responses in land development (and subsequent impacts on water quality); for instance, lower land values on vacant land may signal less desire to develop in critical habitat, less future development, and better water quality. Auffhammer et al. (2020) use a two-way fixed effects [TWFE] design to estimate the impact of critical habitat designation on California vacant lot transactions using data from 1993-2008. The authors conclude that ESA designations reduce vacant land values by roughly 50% during the sample period. Melstrom (2021) also employs a TWFE hedonic model to examine the impacts of species listing on agricultural land values using county-level data from four censuses between 2002-2017. Again, the ESA regulations appear to decrease land values; however, the effect is more modest at an estimated 4% in value, and this effect is concentrated in non-irrigated counties. While partial equilibrium settings tend to illustrate decreases in land values, Quigley and Swoboda (2007) point out the importance of general equilibrium analysis in this setting. Their study finds that critical habitat does decrease land values of designated parcels, but the rents that are redistributed to landowners outside of critical habitat boundaries outweigh these losses. This redistribution decreases consumer welfare due to higher housing prices. On an even broader scale than the housing market, another paper by Duffy-Deno (1997) finds no evidence of an effect of counties' listed species density on county employment and population. Since many papers identify negative impacts of critical habitat on housing prices, this suggests that landowners and developers do respond to designation. More research is needed to connect these behavioral responses to activities that impact water quality, such as land development.

Literature on the effectiveness of the ESA is also relevant since this paper seeks to provide context on the effectiveness of critical habitat. Some non-economics research finds that the ESA is effective in its aim to support species recovery (Taylor, Suckling, & Rachlinski, 2005), but economics work has found that some provisions are more effective than others. Funding has emerged as a key driver in species recovery as Ferraro, McIntosh, and Ospina (2007) find that listing efforts are successful only when accompanied by adequate funding. The degree of species preservation driven by critical habitat, however, is largely unknown. Similarly, many impacts of critical habitat designation are still ambiguous. As the ESA is an environmental policy, it is particularly surprising that environmental outcomes are generally uncharted territory for economists, especially as many potential benefits of critical habitat designation are likely to be environmental. A balanced analysis should consider more of these potential outcomes to effectively evaluate the costs and benefits of designation. Additional investigation into the economic benefits of biodiversity protection, ecosystem services, and environmental quality improvements should be addressed to fill the gap in the literature as it pertains to ESA provisions such as critical habitat.

Surface Water Quality and Land Development

The relationship between land use and water quality is well-documented in the natural sciences literature. Studies generally find negative correlations between anthropogenic land use activities, such as agriculture and urban development, and water quality. Many studies identify increases in nutrient pollution of nitrogen and phosphorus linked to agricultural land use.[2] Others find that agricultural and urban land cover correlate negatively with water

[2]See Bhattarai, Srivastava, Marzen, Hite, and Hatch (2008); Rothenberger, Burkholder, and Brownie (2009)

quality measures such as dissolved oxygen, biochemical oxygen demand, and pathogenic bacteria (Chen & Lu, 2014; Tong & Chen, 2002).

While the studies highlighted above are informative, their correlational nature does not disentangle land use practices from all other variables that could impact water quality such as population growth or GDP. Some economists have begun to fill this gap utilizing econometric techniques used to infer causality. To begin, Atasoy, Palmquist, and Phaneuf (2006) use detailed spatial data from land parcels in a North Carolina county to measure the effect of residential development on water quality. Using an econometric model that accounts for spatial autocorrelation between upstream and downstream observations, the authors document increases in total nitrogen and phosphorus caused by increases in total residential development and in the rate of residential development. This study provides several upgrades over correlational approaches: it incorporates spatial correlation among water quality readings, it avoids aggregation that obscures heterogeneity in land use and regulation in larger land parcels, and it employs panel data to estimate the temporal impact of land use changes on water quality. The authors note that these improvements also increase the relevancy of the results for informing policy governing land use and water quality. Abildtrup, Garcia, and Stenger (2013) incorporate spatial econometric models and IV estimation to analyze the impact of forest land use on the cost of drinking water. The authors show that increases in forest land use result in lower drinking water supply cost by protecting water from pollutants such as herbicides and pesticides that are used less in forests, illustrating the potential for natural land use to improve water quality in some circumstances.

Some studies have used econometric techniques to analyze federal water policy—namely the CWA. Keiser and Shapiro (2019a) examine trends in water quality outcomes and conduct an analysis of CWA grants on water pollution using detailed water monitoring data and triple-differences techniques. While the results, which indicate decreases in water pollution caused partially by CWA expenditures, have limited application to land use policy, the econometric techniques and spatial scale of the analysis provides an example of how federal land use policy could be examined at a national level using large water pollution datasets.

In general, there is a paucity of causal documentation of the links between land use and water quality in economics research. This gap is even larger for studies on the effects of land use policies. Still, other studies illustrate the importance of land use and land use policy when it comes to regulating water pollution. This paper aims to fill this gap and the gap in the ESA literature by evaluating the effects of the critical habitat designation, primarily a land use policy, on water quality.

DATA

This study merges data on water quality measurements, critical habitat boundaries, and weather variables to create an unbalanced, nation-wide panel from 1970-2018. The unique identifier of each observation is at the water monitoring station-by-hour level. Each source of the data is detailed below.

Water Quality Data

Water quality data is obtained from the Water Quality Portal [WQP], an online repository that centralizes access to three federal sources of water quality data: the USGS National Water Information System [NWIS], the EPA Water Quality Exchange [WQX] Data Warehouse, and the USDA ARS Sustaining The Earth's Watersheds - Agricultural Research Database System [STEWARDS]. For some stations, data are available dating back to the early 20th century. The WQP is the largest publicly available source of water quality monitoring data for the United States to the best of my knowledge, making it the best candidate for administrative data covering the outcomes of this study. Data is downloaded in each state in the contiguous U.S. for eleven water quality outcomes: dissolved oxygen, pH, temperature, biochemical oxygen demand, total phosphorus, total nitrogen, total suspended solids, fecal coliform, salinity, specific conductance, and dissolved nitrate.[1] Outcomes were chosen based on other commonly explored water quality measures in the economics literature as well as measures mentioned in federal register documents of final critical habitat rulings.

Since water quality is not defined by a single measure, further explanation of these outcomes provides context for the aspects of water quality that I study in this paper. Dissolved oxygen is often used as an omnibus measure of water quality in economics research.

[1] The approach developed by Koenig, Platt, and Padilla (2023) was instrumental in downloading WQP data.

Aquatic life generally requires adequately high dissolved oxygen levels (Keiser & Shapiro, 2019a), it is a proximate stressor on aquatic life,[2] and it can be altered by many other water quality factors such as temperature or nutrients (U.S. EPA, 2017). pH measures the acidity or alkalinity of a solution. Both highly acidic and highly basic water can produce harmful effects to species such as tissue damage or altered reproductive capacity (U.S. EPA, 2017). Additionally, various pollutants can be either acidic or alkaline in nature—mine wastes and carbon dioxide are often acidic, while asphalt and limestone from roads are often basic. Other pollutants, such as industrial discharges or effluents, can be either acidic or basic depending on the compounds present in the discharge. Biochemical oxygen demand measures the consumption of oxygen by microbial substances during decomposition of organic material. It reflects the amount of dissolved oxygen removed by aerobic bacteria and generally increases with additional pollutants (U.S. EPA, 2017). Phosphorus and nitrogen (including nitrate) are used to determine the nutrient condition in freshwater. While some nutrient concentration is essential for plant growth, excess may cause increases in algal blooms that impact species through subsequent changes to other stressors such as lowered dissolved oxygen. Agricultural activities, industrial and wastewater effluents, and runoff often contribute to higher nutrient concentrations (U.S. EPA, 2017). Total suspended solids represent particles removed from water by filtration. Increases in suspended solids can alter fish and invertebrate assemblages or harm submerged vegetation (U.S. EPA, 2017). Fecal coliform are pathogenic bacteria (such as *E. coli*) that are commonly associated with human and animal waste (Keiser & Shapiro, 2019a). Salinity and specific conductance measure the ionic strength of water. In freshwater systems, increases in salinity or conductance may adversely impact dissolved oxygen capacity or increase erosion, but its direct harm to organisms varies based on species (U.S. EPA, 2017).

[2]Proximate stressors are causal agents that directly evoke biological responses in aquatic life. Other stressors can indirectly harm aquatic life through their impact on proximate stressors (U.S. EPA, 2017).

Each observation in the WQP contains information on the geographic location and time that a sample occurred, the organization that submitted the record, and a variety of fields describing the sampling context such as if the measurement pertains to dissolved, suspended, or total particles in the sample; the depth of the water column at which the sample was taken; and any quality flags associated with the measurement. I harmonize units by aggregating observations of the same outcome with convertible units and identical sample fractions (i.e total fraction vs. suspended fraction) and by filtering out observations that use non-comparable units or methods to identify water quality outcomes.[3] I use a simple mean to average measurements within the same hour of day to get a single observation per site for any hour in the panel with non-missing data. Additionally, I restrict the data to surface water observations.[4]

Table 5.1 lists summary statistics and units for the final dataset, which includes almost 13 million unique observations and 30 million combined water quality measurements. Median water quality is generally satisfactory across stations according to water quality standards used by Keiser and Shapiro (2019a). Dissolved oxygen saturation is above the 36% and 17% deficit cutoffs for “fishable” and “swimmable” waters, respectively; and fecal coliform are below both 1000 and 200 MPN/100 ml cutoffs. Median stations meet standards for “fishable” waters in total suspended solids (< 50 mg/l) and biochemical oxygen demand (< 2.4 mg/l), but not for “swimmable” waters (TSS $<$ 10 mg/l and BOD $<$ 1.5 mg/l). Most water quality outcomes, excluding dissolved oxygen, pH, and temperature are right-skewed with a mean greater than each outcome’s respective median. Right-skewed water quality outcomes generally have a multitude of observations at 0. Dissolved oxygen is slightly left-skewed, while pH and temperature are approximately normally distributed. Over 91 thousand water stations appear in the data, which are reported by 618 different organizations. 87% of the

[3]These filters drop approximately less than 1% of the sample.

[4]Appendix B details comprehensive data downloading and cleaning steps.

nation's 8-digit HUC watersheds and all states in the contiguous U.S. are represented in the data. Additionally, the dataset provides coverage for all years in the panel. Figure 5.1 shows the locations of all stations that appear in the panel. In general, monitoring is denser in the U.S. east, southeast, northern-midwest, and pacific northwest states, with comparably fewer stations reporting data in the western and southwestern states. Monitoring for each outcome varies temporally such that some watersheds record data in most years of the panel, while other watersheds have less coverage across years. Figure 5.2 provides an example of the temporal coverage of pH data over time by 8-digit HUC.

Critical Habitat Boundaries

I obtain shapefiles for critical habitat areas from the U.S. FWS (2022) and the U.S. NOAA Fisheries (2022), the two agencies who administer the ESA. The agencies provide critical habitat shapefiles for species that are currently listed with active critical habitat, so species that have been delisted are not included in the analysis. Of all delisted species, only twelve have critical habitat designations in the contiguous U.S. that are not found in the data. Conversely, 375 designations in the contiguous U.S. are present in the data. Omitting these twelve designations may cause model estimates to be attenuated towards zero if some stations used as controls are, in reality, treated. However, given the proportion of included designations to excluded designations, I expect attenuation bias to be only a minor concern.

Each critical habitat shapefile lists the publishing date of the final critical habitat rule, the species for which the designation pertains, and some species information such as taxon[5]. While most designations become effective roughly one month after a final rule is published, the publishing date listed in each species shapefile is used as the designation start date since most shapefiles do not contain the "effective" start date. Shapefiles are then spatially merged to water station data to determine the day each water station received a

[5]A taxon denotes the larger group within which a species belongs (i.e. fish, clam, mammal, etc.)

critical habitat designation, if ever. In some cases, stations are designated more than once by separate species' critical habitat; consequently, some observations for these stations may correspond to multiple event time indicators in the analysis. Table 5.1 indicates that just over 3% (426,531 observations) of the sample is designated by critical habitat. Almost 5% of stations become designated at some point during the sample period. Figure 5.3 shows all active critical habitat designations colored by the year the designation was published.

Additional Data

I obtain weather data from Wolfram Schlenker, who adapts the PRISM weather dataset to create a balanced panel with daily minimum and maximum temperatures and precipitation for the U.S. at a 2.5 mile resolution.[6] Data on each weather variable is matched to water quality observations using an inverse distance-weighted average of the three closest weather grid centroids to each monitoring location on the day the measurement is recorded. Data are kept in their original units: degrees Celsius for the temperature variables and cm/day for precipitation. Minimum and maximum temperatures are included separately in all models.

Additionally, shapefiles for HUC boundaries are obtained from the USGS (n.d.) Watershed Boundary Dataset, which maps HUC watersheds for the entire U.S. at a 1:24,000 scale. These boundaries are used primarily to create figures, as the water quality data already lists 8-digit HUC codes for each station. Thus, merging HUC shapefiles with water quality data does not require spatial querying; instead, matches are performed simply using HUC codes present in both datasets.

I obtain county-level urban-rural classifications data from the U.S. CDC (2017), which classifies U.S. counties on a scale from 1 (large central metropolitan area) to 6 (non-core rural county). The data was last updated in 2013. While counties may change classification over time, this static measure serves as a reasonable proxy.

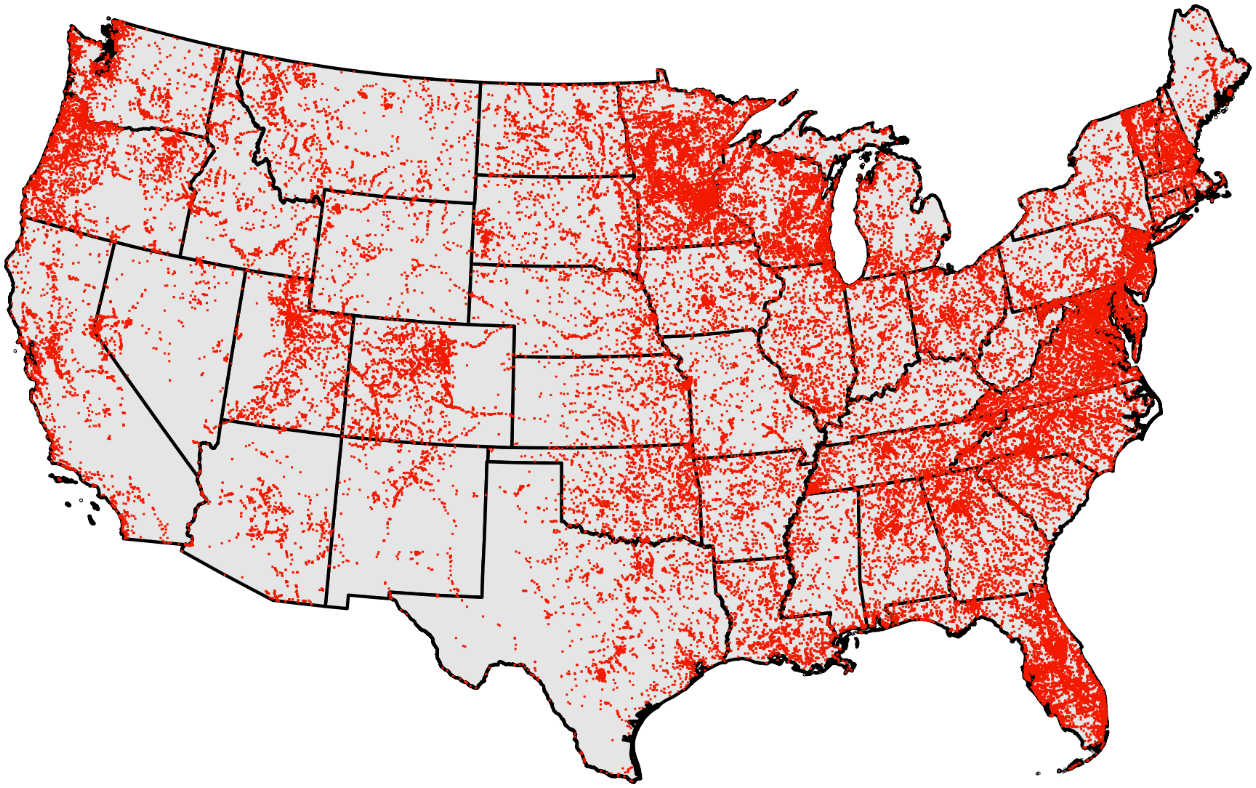
[6] Data is available at columbia.edu/%7ews2162/

Table 5.1: Summary statistics of dataset variables

Variable	Units	N	Unique Values	Mean	SD	Min	Median	Max
DO	% DO sat. -100	3,562,405	1,102,367	-17	24.1	-100	-12.8	42.5
pH	standard units	4,198,549	64,798	7.5	0.8	4.7	7.6	9
Temp	degrees Celsius	11,450,588	499,258	11.3	8.1	-5	10.7	29.7
BOD	mg/l	508,915	3,564	2.7	3.1	0	1.9	20
Phosphorus	mg/l	2,162,889	18,154	0.2	0.4	0	0.1	2.5
Nitrogen	mg/l	763,373	32,073	2	2.5	0	1.1	15
TSS	mg/l	1,781,203	20,085	42.3	105.7	0	11	784
Fecal coliform	MPN per 100 ml	1,145,059	4,681	1,059	4,009	0	100	32,000
Salinity	parts per thousand	508,915	39,825	2.8	6.6	0	0.2	32.5
Specific cond.	uS/cm	3,944,092	191,081	1,067	3,523	0	348	29,000
Dis. nitrate	mg/l	381,686	21,306	3.8	7.2	0	1.2	43.4
CH Designation		12,722,876	2	0.03	0.2	0	0	1
Org. ID		12,722,876	618					
Monitor ID		12,722,876	91,605					
State		12,722,876	49					
County		12,722,876	1,771					
HUC8		12,722,876	1,976					
Year		12,722,876	49	2003	11.1	1970	2005	2018
Month		12,722,876	12	6.7	3.1	1	7	12
Day		12,722,876	31	15.6	8.7	1	15	31
Hour		12,722,876	24	10	6.2	0	10	23
Hour missing		12,722,876	2	0.1	0.3	0	0	1
Max air temp	Deg. C	12,722,876	7,026,946	1.6	3.2	-15.6	1.7	25.5
Min air temp	Deg. C	12,722,876	7,026,946	6.1	3.5	-10.3	6.7	38
Precipitation	cm/day	12,722,876	6,549,297	0.9	2	0	0.1	123.7

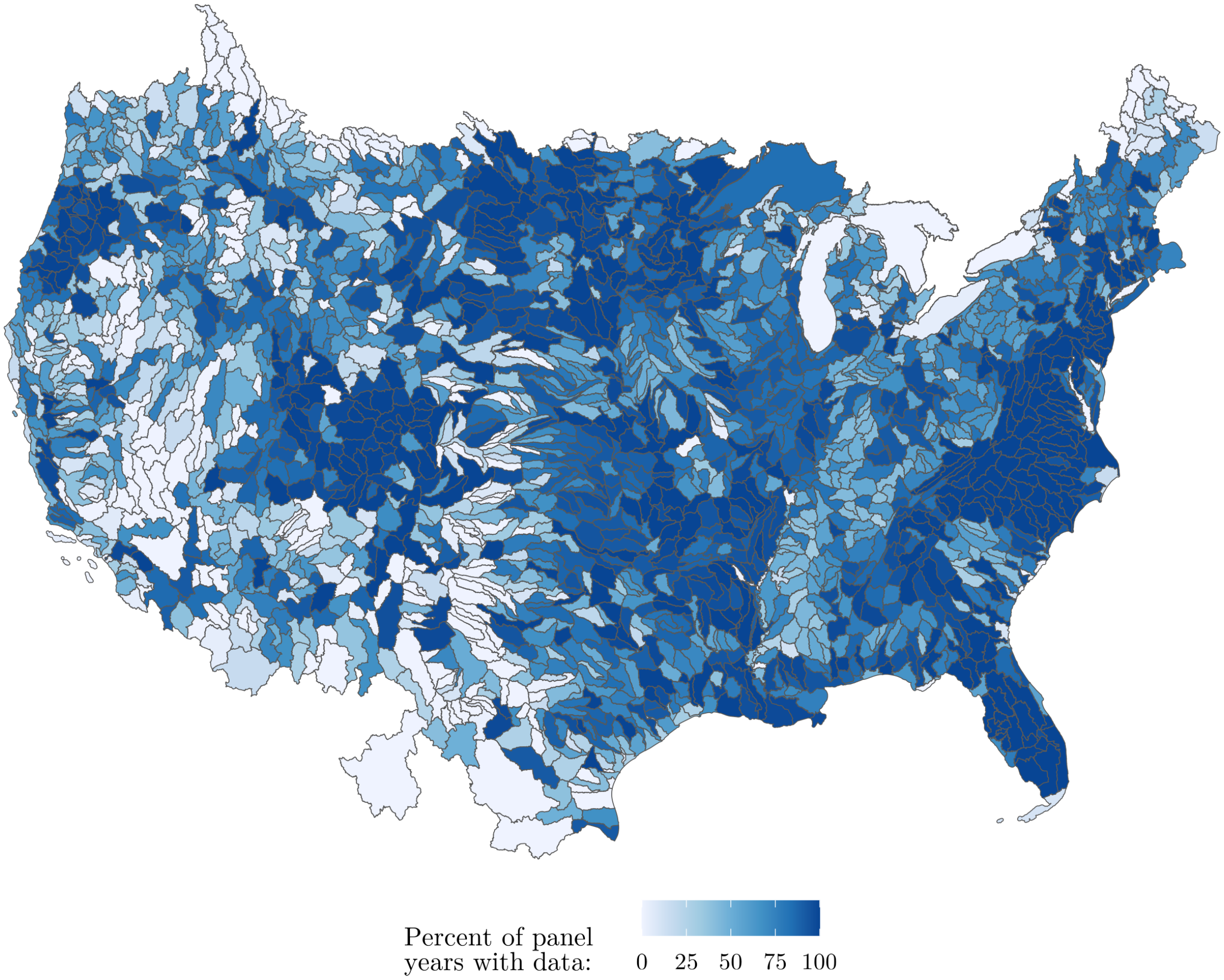
Notes: The table above includes descriptive statistics for variables in the final dataset used in the analysis. Outcome variables are displayed first, followed by the treatment variable, followed by geographic, time, and weather covariates. “DO” represents dissolved oxygen. “Temp” represents water temperature. “BOD” represents biochemical oxygen demand. “TSS” represents total suspended solids. “Specific cond.” represents specific conductance. “Dis. nitrate” represents dissolved nitrate. “Org. ID” represents sampling organization ID.

Figure 5.1: Water quality monitoring stations in the dataset



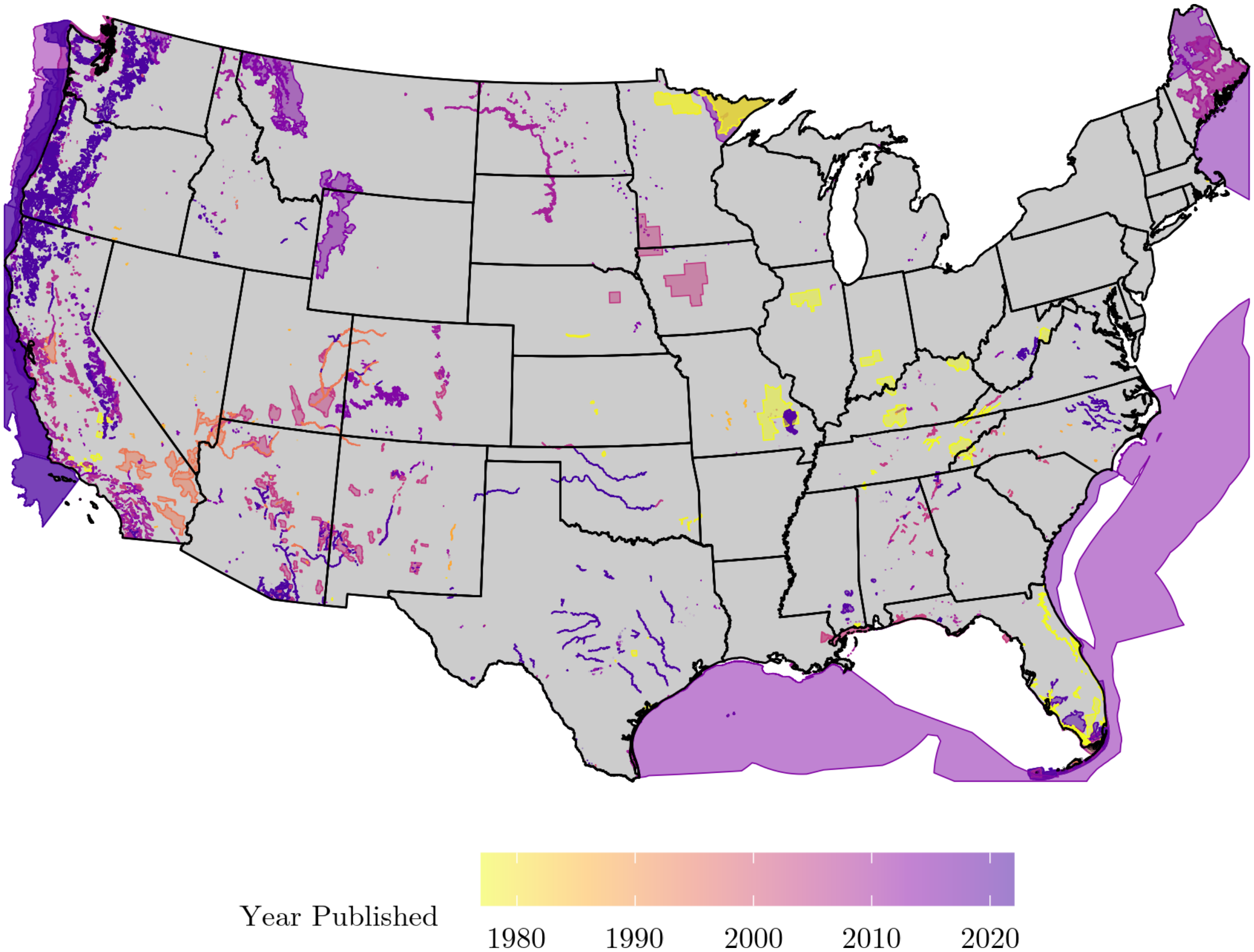
Notes: The figure above shows water monitoring stations that appear in the dataset. Each red dot represents a station that reports at least one observation during the sample period. Water station networks are denser in the Pacific Northwest, Rocky Mountain south, upper Midwest, New England, and the Gulf and Atlantic coasts.

Figure 5.2: pH coverage over time by 8-digit HUC watershed



Notes: The figure above shows coverage of pH data across 8-digit HUC watersheds in the contiguous U.S. from 1970-2018. pH data is recorded in more years for watersheds in the lower pacific northwest, lower mountain west, upper Midwest, and eastern Atlantic coast.

Figure 5.3: Active critical habitat designations in the contiguous U.S., colored by the year the designation was published



Notes: The figure above shows current critical habitat designations in the contiguous U.S. that are used in this analysis. Designations are colored by publishing year to indicate the length they have been in place. Designations of lengths of rivers often appear as lines, while other designations appear as polygons. Designations are spread throughout the contiguous U.S., but many designations are concentrated in the southwestern states, western coastal states, upper Midwest, and southern states.

METHODOLOGY

Identifying the causal effect of critical habitat designations on water quality outcomes requires leveraging plausibly exogenous policy variation with a data panel that records water quality readings at treated and control sites before and after designations occur. In this setting, multiple comparisons are helpful in identifying the relationship of interest: a) changes over time in water quality outcomes at the same site with varying treatment status and b) cross-sectional differences in designated and undesignated stations with otherwise similar characteristics. The ideal experiment would be to select a large number of sites that are identical in factors impacting water quality and to randomly assign some of these sites to receive a critical habitat designation while leaving others undesignated. Responses in water quality outcomes at treated and control stations after designations occur could then be compared to identify the average causal effect of the critical habitat policy on water quality outcomes. Random assignment of the policy would alleviate concerns over omitted variables bias or selection bias that would otherwise present threats to identification.

Of course, such an experiment does not exist in this case. To identify a causal effect, this study implements quasi-experimental econometric methods using high-dimensional fixed effects and an event study framework to estimate the effect of critical habitat policy on water quality outcomes. Fixed effects are employed to remove variation from unobserved factors that could correlate with both the designation and the water quality outcomes, resulting in cases of omitted variables bias. For instance, I use station fixed effects to net out differences in baseline water quality levels in each location. Removing these baseline differences avoids comparing inherently more polluted sites to cleaner sites without accounting for fixed characteristics that may correlate with both critical habitat designation and water quality. These methods, in combination with the event study framework, rely on plausibly exogenous variation in the timing of critical habitat designations conditional on model covariates and fixed effects to identify the causal effect of interest.

The baseline event study specification used in this paper is shown in equation 6.1:

$$Y_{isymdh} = \sum_{e=-10}^{10} \tau_e * D_i(y+e) + \theta_i + \delta_{sy} + \gamma_m + \phi_h + \boldsymbol{\beta X_{isymd}} + \varepsilon_{isymdh} \tag{6.1}$$

where Y_{isymdh} is a water quality outcome of monitoring station i in state s during year y, month m, day d, and hour h. The summation: $\sum_{e=-10}^{10} \tau_e * D_i(y+e)$ is a set of event year indicators, where τ_e is the coefficient of interest for event year e. $D_i(y+e)$ is an indicator variable that is equal to 1 if station i is designated as critical habitat in year $y+e$ and 0 otherwise. Event years -10 and 10 are binned, with all data 10 years before or after a designation used to calibrate each respective event year parameter. θ_i is a vector of water monitoring station fixed effects that control for time-invariant characteristics of a station i, such as baseline levels of pollution or type of surface water. δ_{sy} is a vector of state-by-year fixed effects that control for state-specific characteristics of a state s that vary across calendar years y, such as changes in state regulatory policies, anomalous weather events or natural disasters, and state population changes. δ_{sy} also removes unobserved place-invariant time characteristics such as national trends in pollution or federal water quality regulations. γ_m is a set of month fixed effects that removes seasonal variation in water quality outcomes. ϕ_h is a set of hour-of-day fixed effects that controls for daily variation in water quality outcomes. Both γ_m and ϕ_h are intended to remove variation due to natural fluctuations in water quality; for instance, dissolved oxygen levels are generally higher in the summer and during mornings (Keiser & Shapiro, 2019a). $\boldsymbol{X_{isymd}}$ is a covariate matrix that includes controls for daily minimum temperature, daily maximum temperature, and precipitation at the water monitoring station level as well as a missing hour indicator variable since some observations do not include information on hour-of-day.[1] Weather variables also provide

[1] In robustness checks, a variety of covariate combinations are used including hour-of-day and day-of-year cubic polynomials in place of hour-of-day and monthly fixed effects similar to Keiser and Shapiro (2019a).

additional precision by removing variation in water quality associated with fluctuations in temperature or precipitation on the day a sample was taken. ε_{isymdh} is the idiosyncratic error term that includes all other factors that impact water quality. Similar to Keiser and Shapiro (2019a), all models use standard errors that are clustered at the 8-digit hydrologic unit code [HUC] watershed level.[2] As noted by Keiser and Shapiro (2019a), "a watershed is defined by the USGS as an area of land in which all water within it drains to one point." Therefore, clustering standard errors at the watershed level accounts for error correlation among water quality measurements from stations within the same 8-digit HUC.

Additionally, I examine water quality outcomes using a pooled difference-in-differences [DiD] model shown in equation 6.2:

$$Y_{isymdh} = \tau * CHD_{isymd} + \theta_i + \delta_{sy} + \gamma_m + \phi_h + \boldsymbol{\beta X_{isymd}} + \varepsilon_{isymdh} \tag{6.2}$$

where CHD_{isymd} is the DiD interaction variable of interest equal to 1 if station i in state s is located within a critical habitat designation during year y, month m, and day d. τ is the coefficient of interest that measures the mean response in water quality outcomes for treated stations in the post-period compared to non-treated stations caused by critical habitat designation. All other model variables are identical to those defined in the baseline specification shown in equation 6.1.

The event-study and pooled DiD frameworks I implement are subject to standard model assumptions. Specifically: a) treatment and control stations would have followed parallel water quality trends in the period after designation and in the absence of critical habitat; b) there are no anticipatory behaviors that impact water quality before designation; and

Covariates are also interacted with state fixed effects to allow covariate parameters to vary geographically. In all cases, the results do not change significantly from the baseline specification in equation 6.1.

[2]The HUC classification is an organizational code by which the "United States is divided and sub-divided into successively smaller hydrologic units" according to the USGS (2022). Classifications were historically made at the 2-, 4-, 6-, and 8-digit levels, although some watersheds now include finer 10-, 12-, or 14- digit classifications.

c) there are no spillover effects of critical habitat designation on water quality outcomes of control stations. Violations of these assumptions would impact validity of the calibrated treatment effects. Event study figures assist in evaluating the first two assumptions by examining trends in water quality before designation. Furthermore, specifications that utilize alternate comparison groups of stations surrounding individual designations are reported in the robustness checks section. These additional comparisons present results that can be compared to the baseline specification to evaluate the robustness of baseline results. In a separate model extension, I remove stations most likely to receive spillover effects from designations—those stations that are never-treated but are located in a watershed with critical habitat—to test the no spillovers assumption. These results are also reported in robustness checks, and they find similar effects to the baseline specification.

In addition to the standard event study assumptions, the baseline specification above implicitly assumes that, conditional on model covariates and fixed effects, variation in the timing of the policy is as good as random. This assumption is violated if there are omitted variables that correlate with both the designation and water quality outcomes and that vary at finer geographic or time scales. For instance, it would not violate my identification assumption if designations occur in areas with generally greater concern for environmental quality, as these characteristics are absorbed by station fixed effects. My identification assumption would be violated if, for example, local efforts to improve water quality occur simultaneously in critical habitat areas around the time designations are published. However, most water quality regulation occurs at the national and state levels, which provides additional confidence that state-by-year fixed effects control for many changes in regulation over time (U.S. EPA, 2023).

RESULTS

Main Results

Figure 7.1 presents results from the baseline model shown in equation 6.1 for six primary outcomes of interest: dissolved oxygen saturation, pH, specific conductivity, total suspended solids, total nitrogen, and total phosphorus. These outcomes are highlighted in this main figure since they are repeatedly mentioned as outcomes of interest in final critical habitat rulings for many aquatic species such as the Fluted Kidneyshell, Laurel Dace, and Waccamaw Silverslide (U.S. FWS, 1987, 2012, 2013). For all tables and figures, coefficients are adjusted to normalize interpretation such that a negative coefficient value indicates (generally) a decrease in water quality, while a positive coefficient value indicates (generally) an increase in water quality.[1] The exception to this rule is for pH, which is left unaltered such that a negative coefficient indicates higher acidity and a positive coefficient indicates higher alkalinity, both of which may be undesirable if they result in pH outside of the range that supports aquatic ecosystems.

In addition to the outcomes reported in figure 7.1, other outcomes analyzed include fecal coliform, biochemical oxygen demand, salinity, dissolved nitrates, and temperature. These results are reported in appendix figure A.1 because the outcomes are either mentioned sparsely in the federal register documents that were examined for this study or they correlate highly with other outcomes that have more coverage (i.e salinity correlates highly with specific conductance, which has over six times as many observations; and dissolved nitrate correlates highly with total nitrogen, which has twice as many observations). Temperature is included along with this group of outcomes despite being repeatedly mentioned in federal register documents because it is unclear by which mechanism the critical habitat policy may

[1] Specifically, the following changes are made before regressions are run: specific conductance, total suspended solids, total nitrogen, and total phosphorus are all multiplied by -1 such that a positive coefficient value would indicate an increase in water quality but a decrease in the level each outcome.

meaningfully alter water temperatures in the long run, and I find no empirical evidence of any effect.

The main specification results indicate that most water quality outcomes were not substantially impacted by critical habitat policies after designations were implemented. The main exception is pH, which experiences statistically significant decreases in multiple periods after designation at the 5% significance level. In post-designation years 6 and 8, pH decreases by 0.054 and 0.088 units, respectively, relative to the year before designation *ceteris paribus*.[2] These results indicate that designated surface waters become slightly more acidic up to 8 years after critical habitat designations occur. I also run a model with the absolute deviation of pH from 7 as the outcome and find no qualitative change in these results. Appendix figure A.3 shows these results.[3]

The significance of these pH indicators, however, must be considered within the context of multiple hypothesis testing of a number of different outcomes and event-time indicators. The probability of type I error is proportional to the amount of hypothesis tests being run.[4] Figure 7.1 shows estimates for 20 coefficients and 6 different outcomes. Assuming the model assumptions hold, estimations are internally valid, and there are no true population effects on water quality, I would still expect at least 6 coefficients to be statistically significant at a 5% significance level. Indeed, 6 coefficients from all outcomes in the baseline specification are statistically significant. Thus, robustness checks will evaluate if pH results withstand additional testing to provide context for whether the pH results represent a true effect or an artifact of statistical chance.

Alternatively, the baseline results indicate that dissolved oxygen saturation, specific

[2]Results for 8 years after are also significant at the 1% level

[3]Coefficients in this figure are opposite in sign to preserve interpretation of positive effects as "good" for water quality and negative coefficients as "bad" for water quality.

[4]In my case, type I error would be finding a statistically significant change in water quality due to critical habitat designation from my sample even though there is no true average effect on water quality in the population.

conductance, and total suspended solids are non-significant in all post-designation event time indicators at the 5% significance level, suggesting it is unlikely that designation causes large changes in any of these outcomes. These results indicate there is no evidence that the amount of dissolved oxygen, suspended solids, and dissolved solids in surface waters are significantly impacted by critical habitat designations at the national level. Total nitrogen does not significantly change in any post-designation years, while total phosphorus levels experience significant increases in post-designation years 3, 7, and 10 only at the 10% significance level. In general, nutrient amounts appear largely unaffected, if perhaps slightly negatively impacted, by designation status.

Relative to each outcome's respective mean and standard deviation, the calibrated parameters in the baseline event study model represent precise null-effects. For instance, all dissolved oxygen coefficients range from -1.214% saturation to 1.455% saturation with an average standard error of 0.68% saturation. For the entire panel, the average dissolved oxygen saturation (subtracted by 100%) is -16.8% with a standard deviation of 24%. Ultimately, the calibrated event study coefficients represent a tight window around zero. This pattern is repeated for specific conductance, total suspended solids, total nitrogen, and total phosphorus, which similarly form narrow bands around zero when compared with respective means and standard deviations in the raw data.

To provide additional context for the event time regressions, pooled DiD estimates are calculated for each of the six outcomes. Table 7.1 shows the results from this regression. For all outcomes except specific conductance, there is no evidence that critical habitat designation significantly changes water quality in the post-period for treated stations after the inclusion of fixed effects and after holding model covariates constant. Although two pH coefficients were slightly negative in the event study models, these pooled results find no average declines in pH in the post-period. This is somewhat consistent with the non-significance of event time indicators in the few years following designation and the latter event time indicators. While the DiD results do not contradict the negative pH event time indicators, they do suggest that there is little evidence that any negative effects persist throughout the entire sample post-period for the average treated station. For specific conductance, there is a predicted decrease of 89 microseimens per centimeter in the post-period for treated stations caused by critical habitat designations. However, this result is marginally significant at the 10% significance level, and the calibrated effect is a small proportion of the sample mean, meaning any potential effect is likely small in size.

The baseline specification and pooled DiD results above are dependent upon the assumption that water quality outcomes in designated and undesignated water stations would have followed parallel trends in the post-period in the absence of designation. The pre-event trends generally lend credibility to this assumption, as no pretrend coefficients are significant at a 5% significance level for pH, total nitrogen, or total phosphorus. While pre-event indicators 8 and 9 are significant for dissolved oxygen, and while specific conductance and total suspended solids each have one significant pre-event indicator, there do not appear to be any consistent trends among the indicators for these outcomes. Additionally, at least one event indicator is expected to be significant at a 5% level by chance for each outcome since there are a total of 20 event indicator variables. Thus, there do not appear to be any notable signs of differing pre-event trends for designated and undesignated stations for the examined outcomes.

Robustness Checks

Patterns in Monitoring Station Sampling

Not all stations are sampled in each day, month, or even year of the panel. As mentioned by Keiser and Shapiro (2019a), there are no obvious explanations for the variation in the density of water quality measurement across time and space. Still, exploring patterns in sampling networks provides valuable insight into the drawbacks of water monitoring data used in this analysis. There are two potential issues that present reasonable concerns in my setting.[5] First, if sites are sampled at points in time when water quality is expected to be worse or better than average (i.e. after storms or other events that decrease water quality), regression estimates could be biased downwards since the treatment effect would also incorporate information from selective sampling within stations. Second, if sampling of monitoring sites is a function of critical habitat designation itself (i.e. sites are sampled less after designation compared to before), there may be extra noise added to treatment effects. Additionally, any correlation between sampling frequency and critical habitat may raise concerns over underlying sampling patterns that could cause bias, although I cannot directly test for any of these hypothetical concerns. First, to examine correlation in the raw data between sampling frequency and the event years of interest for treated stations, I plot the median number of yearly observations per treated station by event year. Figure 7.2 shows the results of this exercise for dissolved oxygen and pH observations. The figure shows that event years before designation are sampled slightly more than event years after designation for the median station. Still, the difference is not large, as the average event year (excluding binned event years -10 and 10) appears to be sampled 6-7 times for the median station for both dissolved oxygen and pH.

Still, monitor and state-by-year fixed effects may account for baseline levels or regional yearly trends in water quality sampling. Thus, to examine patterns in the identifying variation used in my analysis, I run two pairs of regressions to explore a) if areas that receive more samples in a year exhibit higher or lower levels of water quality outcomes and b) if critical habitat areas are systematically sampled more or less than non-designated areas, even after the inclusion of model fixed effects. Table 7.2 shows the results of these regressions for stations that report data for dissolved oxygen and pH. The first pair of models regress number of observations for a station in a given year on the mean water quality outcome

[5]Grainger and Schreiber (2019) discuss these issues in the use air pollution monitoring data.

measurement for all stations (columns 1 and 5) and only treated stations (columns 2 and 6). While this regression has the drawback that stations without any samples in a year cannot be included since the water quality level of these observations is unobserved, it does evaluate whether stations that do get sampled exhibit correlation between frequency of sampling and water quality levels. The second pair of models regresses the number of observations for a station in a given year on the pooled interaction DiD variable indicating if a station is within critical habitat designation boundaries in the period following designation (columns 3 and 7). I also run this regression with an alternate dependent variable: an indicator variable for whether a station is sampled at all in a given year (columns 4 and 8). For these models, I create a balanced panel at the water station-by-year level where each station in the data is given an observation for each year in the panel. Again, all models include monitoring station and year-by-state fixed effects to examine patterns in the identifying variation leveraged in the baseline model.

Figure 7.2, column 1 suggests there is not significant evidence to conclude that dissolved oxygen levels are predicted to be higher or lower based on frequency of sampling for the average station. For pH (column 5), there is a significant trend at a 5% significance level, indicating that a 0.1 unit increase in pH is correlated with a station having .03 less observations in a given year. However, this effect is very small as the average station is sampled around 8 times per year, conditional on being sampled at all. Even for stations that are sampled at the 99th percentile of number of observations per year, this correlation only predicts pH values would be 0.0056 units lower than a station sampled the average number of times per year. In general, the results provide additional confidence that dissolved oxygen measurements do not appear to be sampled more or less based on dissolved oxygen levels. While pH values may exhibit slight downward trends based on monitoring frequency, these trends represent very small changes in pH values. pH coefficients could be biased downwards, but the bias would likely be very small.

The second pair of regressions indicate that both dissolved oxygen and pH exhibit slight patterns in sampling behavior based on whether stations are designated or not. Being a treated station post-designation leads to a predicted 0.2 less yearly observations and a 2.6% smaller probability of being sampled at all for dissolved oxygen even after inclusion of model fixed effects. Similarly, being a treated station post-designation leads to a predicted 0.24 less yearly observations and a predicted 2.1% smaller probability of being sampled at all for pH after including model fixed effects. While there does appear to be slight under-sampling of designated stations, it is not easy to determine the impact of this effect. There are several important considerations. First, the effects are small since the average station records data in around 16.7% of years in the panel. Second, the decrease in sampling of designated stations likely indicates that there is additional noise in regression coefficients due to there being less data for treated observations than for untreated observations. While only hypothetical, monitoring patterns in critical habitat areas that contribute to this under sampling could also bias treatment effects. Alternatively, it could be that critical habitat sites are more highly sampled leading up to designation to determine the environmental state of the ecosystem, then sampling occurs less frequently after critical habitat is designated—a situation which is uncorrelated with the water quality outcome and would not cause bias.

While it is not clear if underlying sampling frequency in critical habitat would bias coefficients, examining correlation between sampling frequency and water quality in only treated stations provides some additional context. If monitors are sampled at points in time that are systematically cleaner or dirtier than average within critical habitat boundaries, this may indicate that bias is more likely. Columns 2 and 6 of figure 7.2 estimate the effects of the regression of water quality level on number of observations per year on a sub-sample of only treated stations. I find that there is no significant correlation within critical habitat boundaries for pH, and dissolved oxygen displays a marginally significant effect, indicating that an increase in dissolved oxygen deficit of 1 percentage point correlates with 0.008 less yearly observations compared to pre-designation years. This result is very small given the mean dissolved oxygen deficit of roughly -17 and that stations, conditional on being sampled in a given year, report roughly eight samples per year.

I complete one final exercise to alleviate concerns that unbalanced sampling influences the results of the baseline specification. First, I keep only treated stations with at least one observation in each event year such that the treated sample portion forms a balanced panel of observations in event time.[6] Next, I regress the number of yearly observations per station on the set of event-time indicators, including station and year-by-state fixed effects. This regression examines correlation between sampling frequency and the independent variables of interest using identifying variation in the model. Figure 7.3 shows results from this exercise. The regression indicates that sampling frequency is uncorrelated with event time indicators after controlling for model fixed effects using the balanced treated stations. Lastly, I re-run the baseline specification keeping only balanced treated stations to examine changes compared to the results shown in figure 7.1. I show the results of this exercise in appendix figure A.2. There are no significant qualitative differences between the results. Dissolved oxygen displays slightly more significant pre-trends, but pH is still significantly negative in event year 8. Additionally, pH decreases in the year of designation, and all other post-event coefficients remain negative.

There are several key takeaways from these exercises. Any bias to regression coefficients would likely be small and negative. While monitors in critical habitat are sampled slightly less after designation than other monitors, there does not appear to be a correlation between the frequency of sampling and the level of pH for these treated stations. This provides some confidence that monitors would not be sampled strategically less (or more) in polluted places within designation boundaries. Any endogeneity of monitor sampling frequency would likely

[6] I also align calendar and event time so that I can aggregate the data to the station-by-year level for this analysis. Doing so additionally simplifies the analysis to ease in interpretation.

add some noise to regression estimates for both dissolved oxygen and pH. Rerunning the results using only treated stations that are balanced in event years yields similar results to the baseline specification, adding confidence that sampling patterns are not a key driver of the results. Ultimately, the small significant effects found highlight the additional need for representative water quality data that fully cover critical habitat areas.

Alternate Specifications

To assess the event study assumptions' validity in the baseline specification and the robustness of the results from this model, two alternate specifications are explored. The first alternate specification uses a sample that excludes never-treated stations in watersheds that have a critical habitat designation at some point in time. Excluding these never-treated stations addresses the assumption of no spillovers in the baseline specification, as they are most likely to receive spillover effects from nearby designations in the same watershed. If spillovers are present, any effects of designation would be understated because some control stations would be partially treated, and one would expect the coefficients under this model extension to be greater in magnitude.

The second alternate specification uses "designation group" fixed effects in place of state fixed effects. A "designation group" is constructed by selecting a listed species' critical habitat area and identifying all watersheds that intersect with designation or that are adjacent to a "designated" watershed. All stations within these selected watersheds are coded as belonging to the species' designation group. Since many watersheds and water stations belong to more than one "designation group," some observations are repeated to create a stacked dataset that uses the same observation for several comparison groups.[7] This alternate specification provides an alternative control group for cross-sectional comparisons

[7] Additionally, some water stations belong to a designation group and eventually receive a critical habitat designation, but not for the species that defines that particular designation group. These observations are excluded.

between designated and undesignated stations. While the baseline specification compares watersheds within the same state that have different treatment status at a point in time, this specification instead compares treated and control water stations that are geographically close. This alternate treatment group would absorb omitted variables that vary within the same year as designations in the watersheds surrounding a critical habitat area as opposed to omitted variables that vary at the state level at the same time as designation. This would provide a more valid comparison if, for example, designations status brings more awareness to water quality in areas immediately surrounding critical habitat and leads to improvement efforts around the time designations are announced. The drawback of using this specification is that it relies on less data for comparison groups, lowering statistical power of the coefficients. Figure 7.4 provides a visual example of the "designation group" alternate specification using a single geographical area.

Table 7.3 displays results from the baseline and alternate specifications for dissolved oxygen—as it is commonly explored as a comprehensive water quality measure in the economics literature (Keiser & Shapiro, 2019a)—and pH, as the baseline results find significant decreases in year 8 after designation. The results for dissolved oxygen are largely unchanged in the specification that removes never-treated stations in designated watersheds, as all event time indicators in the post-period remain statistically indistinguishable from zero. When compared to the baseline specification for pH, this model also finds qualitatively similar results. All statistically significant event time indicators in the baseline model remain statistically significant at the same significance levels, and the effects of designation appear statistically significant at a 10% significance level even 9 years after the designations occur. These results indicate that pH decreases by as much as 0.092 units 8 years after critical habitat designation relative to the year before designation when controlling for model covariates at a 1% significance level. These results provide additional confidence that spillovers are not a cause of bias in the baseline results since the results are qualitatively similar across outcomes when stations close to designations and within the same watersheds are removed from the estimation.

The second alternate specification, which utilizes "designation group" fixed effects in place of state fixed effects, leverages slightly different cross-sectional variation by comparing stations that are in neighboring watersheds instead of stations within the same state boundaries. For dissolved oxygen saturation, the model results change slightly as one post-designation event indicator is statistically significant at a 10% level, and three become significant at a 5% level. This specification finds potentially long-term negative effects of critical habitat designation on dissolved oxygen deficit, in contrast with the baseline specification which illustrated no effect of designation. Over ten years after designation, the designation groups model indicates that dissolved oxygen saturation decreases by 2.1% relative to the year before designation after accounting for model covariates, illustrating a slight decline in water quality. While the results are somewhat surprising since it would seem unlikely that designation would cause long-term *negative* impacts on dissolved oxygen, they could be biased downwards based on the negative correlation between sampling frequency and dissolved oxygen levels in treated stations discussed in the previous section. In any case, the dissolved oxygen results go from a null effect to a slightly negative result using the alternate comparison group.

Results from the "designation group" alternate specification for other main outcomes are shown in appendix table A.2. Total suspended solids and total phosphorus exhibit statistically significant and positive coefficients in some post-event years; however, the results for total phosphorus also suffer from potential violations of the parallel trends assumption due to significance of event year indicators immediately preceding designation. Again, results for these outcomes are not consistent across model specifications.

Alternatively, the designation groups model for pH again illustrates decreases after designation occurs; however, the results are not as definitive. While the effects of the policy are not statistically significant 6 years after designation, they remain significant at a 1% level 8 years after designation and are almost exclusively negative in sign in the post-period. Relative to the year before designation, critical habitat policies result in a 0.065 unit decrease in pH 8 years after the policy, in contrast with a 0.088 unit decrease found using the baseline model. The alternate comparison group highlights results that are qualitatively similar, but less significant and lower in magnitude for pH compared to the baseline model.

Heterogeneity Results

To further examine the mechanisms behind the pH results found in the baseline specification, various subgroups were analyzed to determine where the effects are concentrated. If certain subgroups drive the pH effects, this could help disentangle a notable pH decrease from a mildly significant, but spurious decrease after designation. Doing so also allows for speculation on the mechanisms behind the result. Consequently, several groups were analyzed to evaluate the heterogeneous effects of designation. First, the baseline model specification was estimated separately for urban and rural counties due to the differences in land use among these areas. If the water quality effects are concentrated in rural counties, for instance, this may indicate that the drivers of water quality effects from designation are related to activities that occur more often in rural areas—i.e. agriculture, forestry, etc. Alternatively, if the effect is stronger in urban areas, other mechanisms such as the policy's effect on construction or urban development could be the primary drivers. Next, effects were explored for aquatic and fish species designations to examine if water quality effects were concentrated in designations meant to support aquatic life. If so, this may indicate that critical habitat designations primarily impact water quality when designated species are directly impacted by water quality, whereas designations for species whose recovery is less directly impacted by water quality would result in little to no effect on water contaminants.

Figure 7.5 compares pH event study regressions separately for the heterogeneity groups described above. The decline in pH after designation occurs is concentrated in urban counties, with rural counties demonstrating no significant change in pH after designation. For urban counties, there is a significant decline in pH after designation in post-event years 3, 6, 7, and 8 at the 5% level. Results indicate that pH declines by as much as 0.13 units 8 years after designation compared to the year before designation, *ceteris paribus*. The magnitude of this effect is roughly one and a half times as large as the result for all stations, indicating that urban areas drive some of the results found in the baseline specification.

Decreases in pH after designation also appear to be concentrated among aquatic, and particularly fish species. Aquatic species experience significant declines in pH in post-designation years 6 and 8, while fish species show significant declines in post-designation years 0, 2, 3, 6, and 8. In the eighth year after designation, pH decreases by 0.11 units for fish species and by 0.095 units for aquatic species compared to the year before designation, *ceteris paribus*. These results indicate that aquatic species' designations, in particular fish designations, contribute heavily to the pH declines found in the baseline specification.

To supplement the event study results for heterogeneity groups, I run pooled DiD regressions for each of these four groups using interaction coefficients by subgroup. The results are shown in appendix table A.3. I conduct joint significance tests on the pooled DiD coefficient and the DiD coefficient interacted with subgroup indicator variables and report the p-values in this table as well. This joint test of significance indicates that treated stations with fish designations experience a statistically significant decrease in pH of -0.058 standard units in the post-period compared to untreated stations after controlling for model covariates. Joint significant tests for urban stations, rural stations, and aquatic designations are not statistically significant. Based on this result, I also run similar pooled DiD regressions for fish designations for each of other five main water quality outcomes. Results shown in appendix table A.4 indicate that treated stations with fish designations experience a statistically significant increase in total phosphorus of 0.035 mg/l (a slight decrease in water quality) and a statistically significant decrease in total suspended solids of 5.56 mg/l (a slight increase in water quality) in the post-period compared to untreated stations at 5% and 10% significance levels, respectively. Results from these tables indicate that fish designations drive water quality effects in pH and some other outcomes, but these effects remain quite small. Still, these results suggest fish designations may produce stronger average water quality effects than all designations pooled together.

Since the results are concentrated among fish species (and somewhat in urban counties), evaluating drivers of low and high pH could indicate possible mechanisms by which critical habitat designations decrease pH. The policy could either contribute to activities that lower pH or discourage activities that increase pH, either of which may lead to net decreases after the designations occur. According to the U.S. EPA (2017), contributors to low pH include wastes from mines, energy plants such as power plants and coal piles, industrial and landfill wastes, runoff from dairy farming, and wetland draining; while contributors to high pH include alkaline soils, road construction waste from sources such as asphalt and limestone gravel roads, agricultural lime, manufacturing of cement and soap, and industrial discharge and landfills. There is no evidence in the literature for critical habitat designations increasing mining activities or contributing to greater industrial waste, both of which could potentially lower pH in urban areas. While existing studies have found relatively little impact on land use changes following critical habitat designation, there may be reason to suspect that the designations could result in lower pH by reducing road construction activities that produce alkaline discharge. First, these activities would likely occur more in urban areas, where the pH effects appear to be concentrated, and construction projects would incur additional regulatory scrutiny in areas where they are more likely to occur. Second, federal register documents for numerous species mention road construction as a key potential contributor to species degradation through water quality impacts.[8] The critical habitat rule

[8]U.S. FWS (1987, 2007, 2012, 2013)

for the Waccamaw Silverslide, for instance, specifically mentions the road expansion of a section of US Highway 74 as a development project which may adversely impact pH in Lake Waccamaw, the critical habitat site of the fish species. While the importance of suppressed road construction as a driver of declining pH in critical habitat areas after designation remains speculative, further investigation is necessary to confirm this as a main factor driving the results found in this study.

Table 7.1: Pooled difference-in-differences results for main outcomes

	Dissolved oxygen	pH	Specific conductance	Total suspended solids	Total nitrogen	Total phosphorus
	(1)	(2)	(3)	(4)	(5)	(6)
CH designation	-0.543	-0.013	88.922*	2.372	-0.068	-0.010
	(0.633)	(0.015)	(51.815)	(2.849)	(0.080)	(0.014)
Min. temp	-1.231***	-0.017***	14.749***	-3.809***	-0.024***	-0.005***
	(0.070)	(0.001)	(3.007)	(0.305)	(0.007)	(0.001)
Max. temp	1.473***	0.025***	-19.213***	0.465**	0.038***	0.001
	(0.076)	(0.002)	(4.865)	(0.233)	(0.006)	(0.001)
Precip.	-0.272***	-0.012***	15.323***	-7.126***	-0.005	-0.010***
	(0.018)	(0.001)	(1.980)	(0.448)	(0.005)	(0.001)
N	3,620,836	4,182,501	3,996,720	1,834,381	777,015	2,116,390
Sample mean	-17	7.5	1,067	42.3	2	0.2
Year-by-state FE	Yes	Yes	Yes	Yes	Yes	Yes
Monitor FE	Yes	Yes	Yes	Yes	Yes	Yes
Hour & month FE	Yes	Yes	Yes	Yes	Yes	Yes
SE Cluster	HUC8	HUC8	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table above shows results for the pooled difference-in-differences regression of critical habitat designation on the main water quality outcomes: dissolved oxygen, pH, specific conductance, total suspended solids, total nitrogen, and total phosphorus. Coefficients for specific conductance, total suspended solids, total nitrogen, and total phosphorus are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

Table 7.2: Regressions exploring sampling patterns of water quality data using stations that report data for dissolved oxygen and pH

	Dissolved Oxygen				pH			
	Yearly number of observations			Any data indicator	Yearly number of observations			Any data indicator
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean yearly outcome	-0.004 (0.002)	-0.008* (0.004)			-0.336*** (0.117)	-0.255 (0.196)		
CH designation			-0.209* (0.114)	-0.026*** (0.009)			-0.248** (0.115)	-0.021** (0.009)
Stations in sample	All	Only treated	All	All	All	Only treated	All	All
N	456,638	28,984	2,899,722	2,899,722	524,778	33,738	3,137,078	3,137,078
Monitor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-by-state FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SE cluster	HUC8	HUC8	HUC8	HUC8	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table above shows results for four regressions run to explore patterns in the sampling of two outcomes each: dissolved oxygen and pH measurements. Regressions 1, 2, 3, 5, 6, and 7 use mean yearly water quality measurement as the dependent variable. Regressions 4 and 8 use an indicator for if the water quality measurement is recorded at the monitoring location at all in a given year (1 if yes, 0 otherwise) as the dependent variable. Regressions 1, 2, 5, and 6 include number of water quality measurements per year as the independent variable, while regressions 3, 4, 7, and 8 include an indicator for critical habitat designation (1 if designated, 0 otherwise) as the independent variable. Regressions 2 and 6 only include stations that are treated at some point during the sample period. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

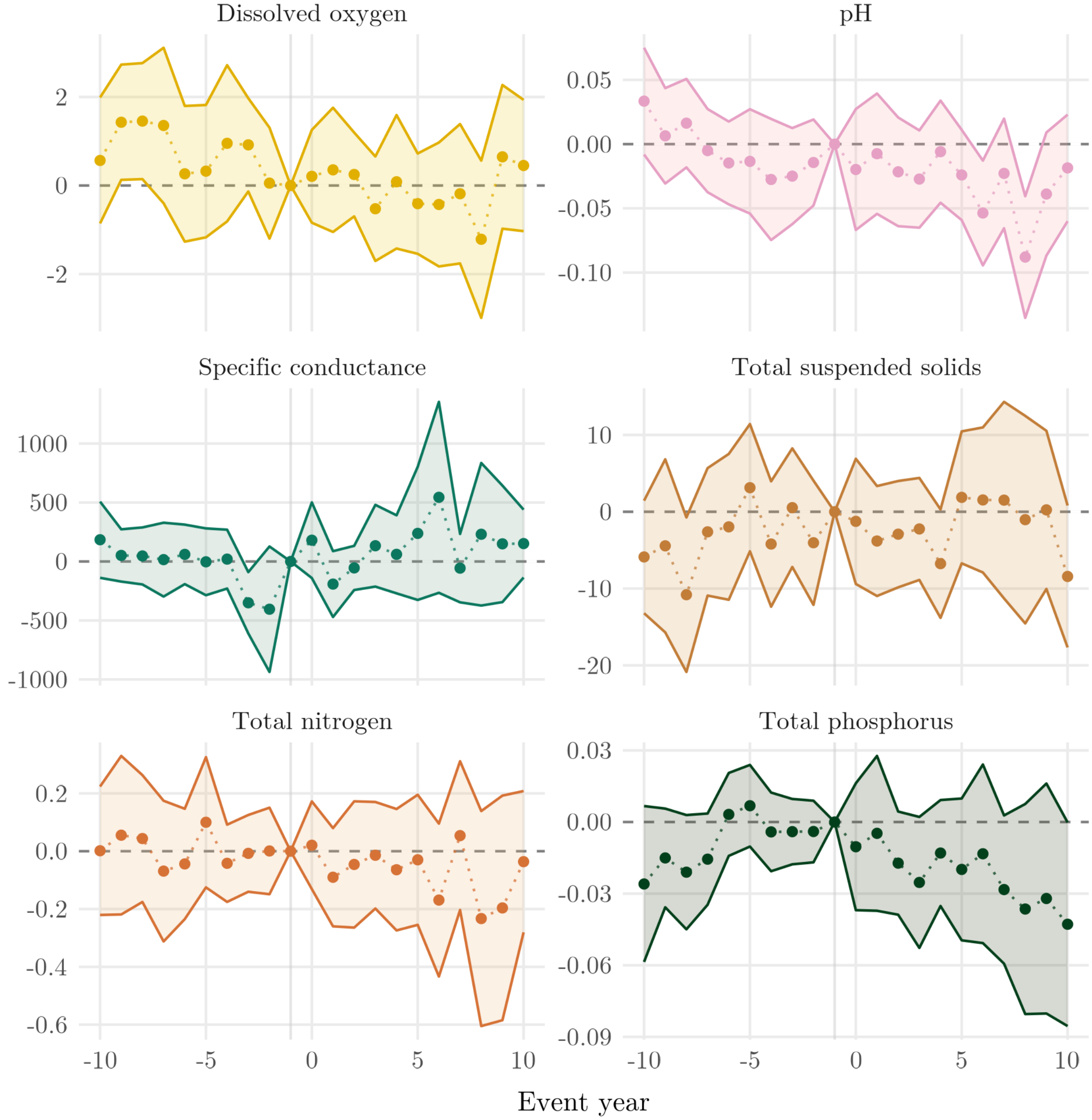
Table 7.3: Dissolved oxygen and pH results comparing baseline and alternate specifications

	Dissolved Oxygen			pH		
	Baseline model results	Extension 1: Removing spillover stations	Extension 2: Designation groups	Baseline model results	Extension 1: Removing spillover stations	Extension 2: Designation groups
	(1)	(2)	(3)	(4)	(5)	(6)
Event Year -5	0.324	0.165	-0.613	-0.013	-0.017	-0.011
	(0.762)	(0.685)	(0.907)	(0.021)	(0.022)	(0.020)
Event Year -4	0.954	0.968	0.177	-0.028	-0.033	-0.014
	(0.900)	(0.796)	(1.068)	(0.024)	(0.026)	(0.022)
Event Year -3	0.919*	1.066*	0.093	-0.025	-0.033	-0.008
	(0.537)	(0.572)	(0.700)	(0.019)	(0.021)	(0.018)
Event Year -2	0.054	0.033	-1.259*	-0.014	-0.019	-0.007
	(0.639)	(0.687)	(0.699)	(0.017)	(0.018)	(0.018)
Event Year 0	0.208	0.437	-0.180	-0.020	-0.030	-0.017
	(0.535)	(0.543)	(0.710)	(0.024)	(0.025)	(0.026)
Event Year 1	0.354	0.607	-0.805	-0.007	-0.015	-0.011
	(0.715)	(0.751)	(0.939)	(0.024)	(0.026)	(0.023)
Event Year 2	0.250	0.549	-0.801	-0.022	-0.024	-0.034
	(0.482)	(0.563)	(0.672)	(0.022)	(0.023)	(0.022)
Event Year 3	-0.524	-0.383	-1.503**	-0.027	-0.033	-0.022
	(0.602)	(0.682)	(0.613)	(0.019)	(0.021)	(0.020)
Event Year 4	0.085	0.269	-0.653	-0.006	-0.012	0.002
	(0.769)	(0.810)	(0.764)	(0.020)	(0.020)	(0.019)
Event Year 5	-0.409	-0.344	-1.199*	-0.024	-0.026	-0.020
	(0.578)	(0.557)	(0.696)	(0.018)	(0.018)	(0.020)
Event Year 6	-0.428	-0.305	-1.017	-0.054**	-0.059**	-0.035
	(0.714)	(0.679)	(0.917)	(0.021)	(0.023)	(0.024)
Event Year 7	-0.186	0.133	-1.048	-0.023	-0.027	-0.017
	(0.803)	(0.760)	(1.046)	(0.022)	(0.022)	(0.023)
Event Year 8	-1.214	-0.804	-2.557**	-0.088***	-0.092***	-0.065***
	(0.906)	(0.894)	(1.239)	(0.024)	(0.025)	(0.024)
Event Year 9	0.648	0.955	-0.948	-0.039	-0.045*	-0.011
	(0.828)	(0.742)	(1.294)	(0.024)	(0.025)	(0.022)
Event Year 10	0.453	0.848	-2.108**	-0.019	-0.028	-0.016
	(0.756)	(0.765)	(1.071)	(0.021)	(0.025)	(0.020)
N	3,620,836	2,671,750	11,393,451	4,182,501	3,104,359	12,660,140
Sample mean	-17	-17	-17	7.5	7.5	7.5
Year-by-state FE	Yes	Yes	No	Yes	Yes	Yes
Year-by-des. FE	No	No	Yes	No	No	Yes
Monitor FE	Yes	Yes	Yes	Yes	Yes	Yes
Hour & month FE	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
SE cluster	HUC8	HUC8	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

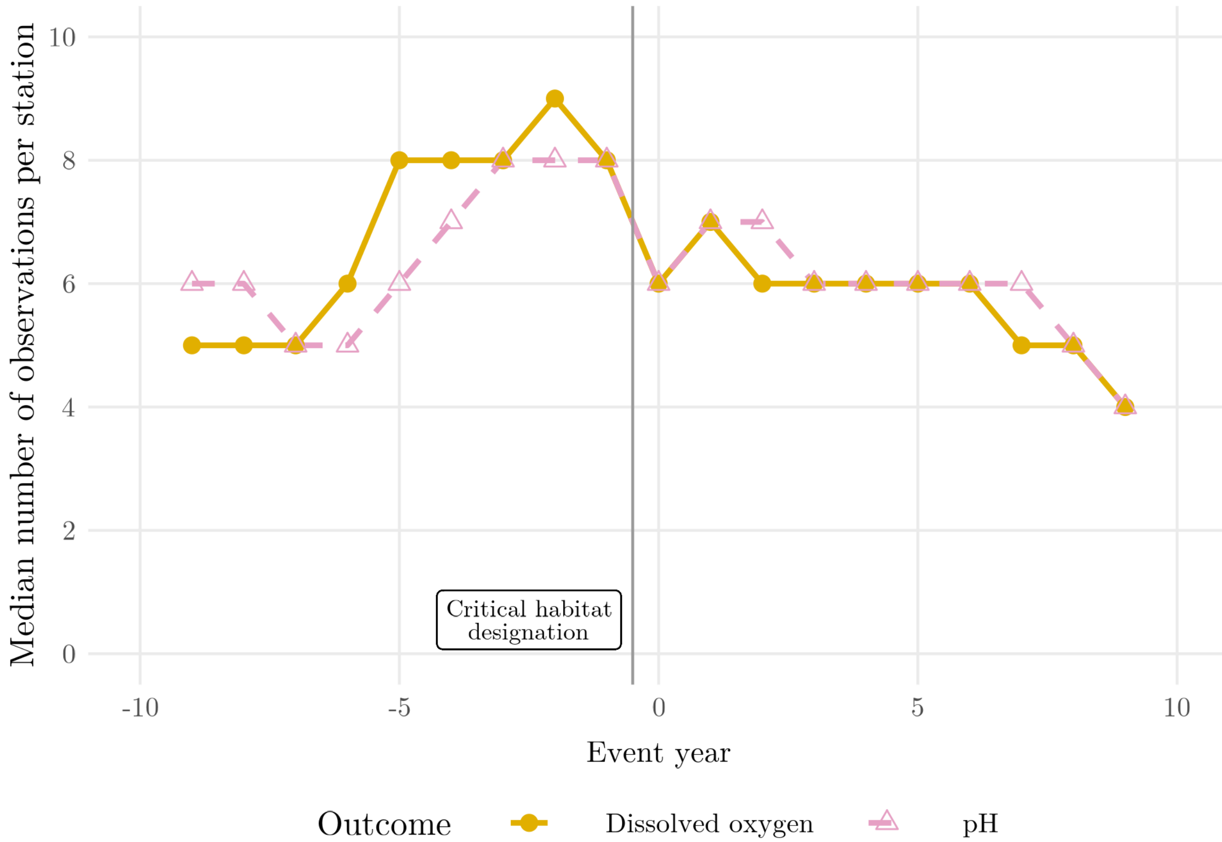
Notes: The table above shows results for the baseline model specification compared with two alternate specifications. Alternate specification 1 is the baseline specification using a restricted sample that removes never-treated water stations in watersheds with critical habitat designation. Alternate specification 2 is the "designation groups" model that replaces year-by-state fixed effects with year-by-designation group fixed effects. Although event year indicators for event years -10 through -6 are also included in these models, they are omitted from this table for brevity. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

Figure 7.1: Event time figures of baseline specification results for main outcomes



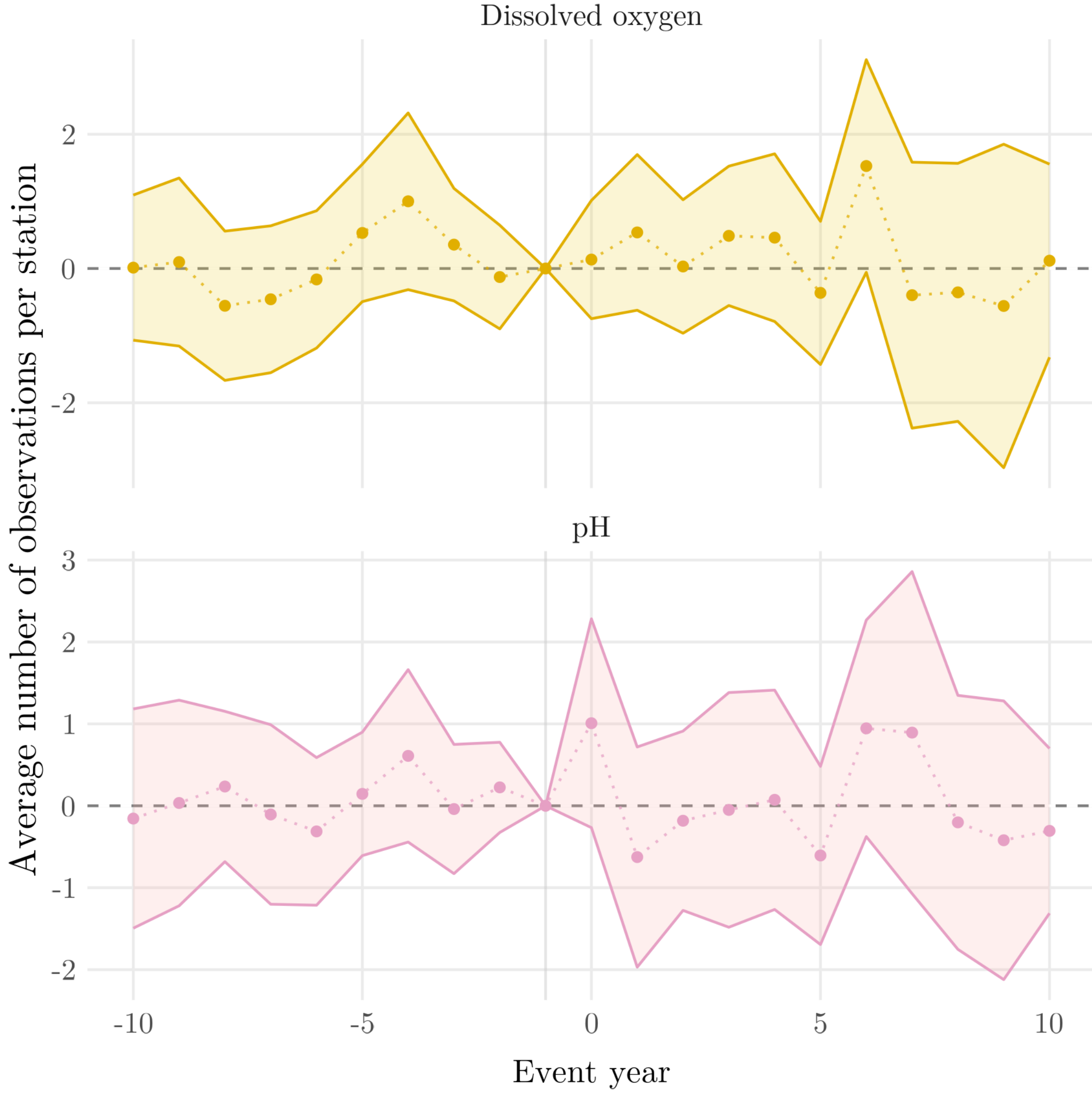
Notes: The figure above shows event time figures for the results of the baseline model specification for the main outcomes: dissolved oxygen, pH, specific conductance, total suspended solids, total nitrogen, and total phosphorus. Dissolved oxygen is recorded in percent saturation deficit; pH is in standard units; specific conductance is in microsiemens per centimeter; total suspended solids, total nitrogen, and total phosphorus are in milligrams per liter. Coefficients for specific conductance, total suspended solids, total nitrogen, and total phosphorus are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. Standard error bands around each event year coefficient are clustered at the 8-digit HUC level. Vertical bars at event year -1 indicate that this coefficient is left out of the analysis.

Figure 7.2: Median number of dissolved oxygen and pH observations per treated station by event year



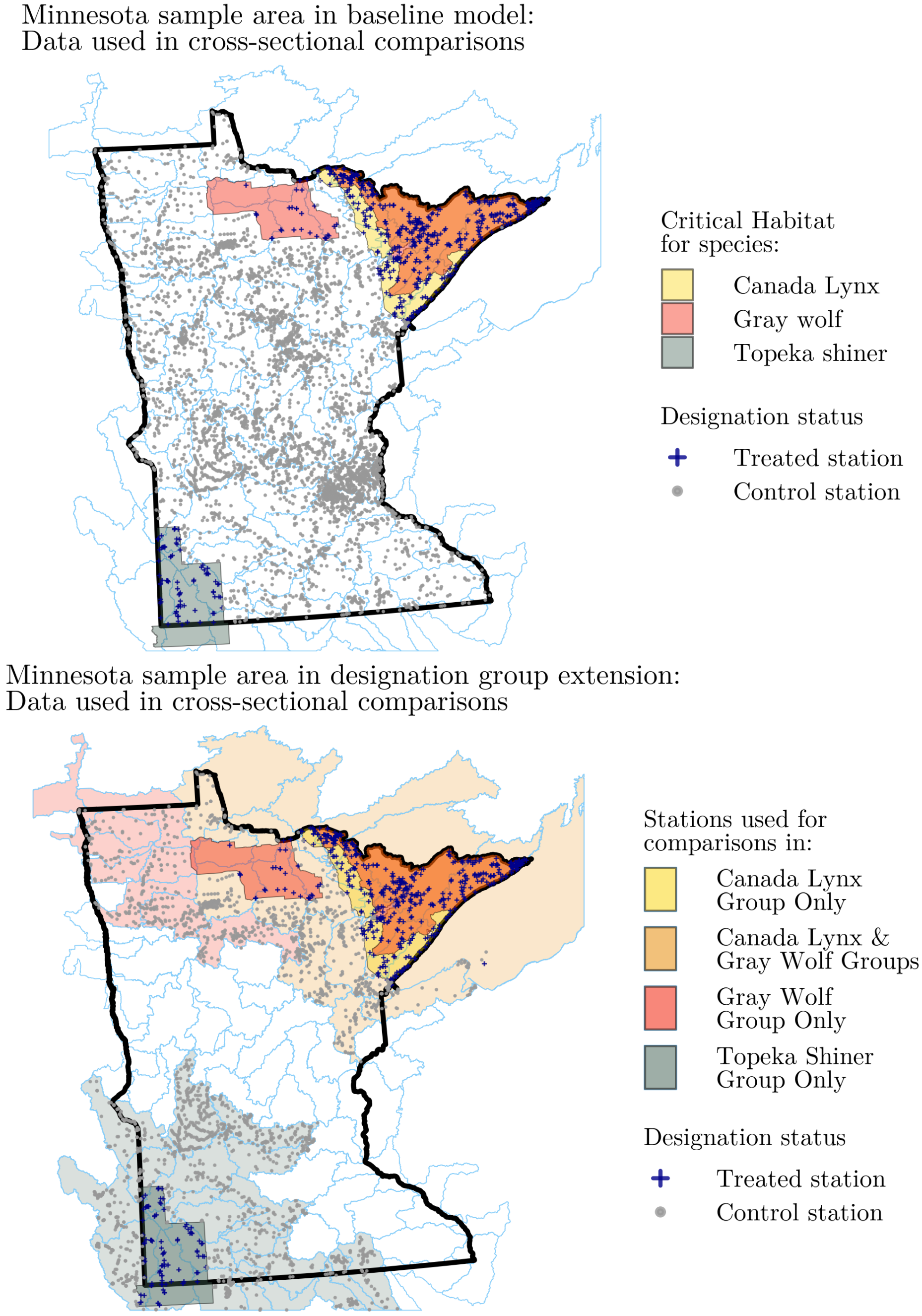
Notes: The figure above shows the median number of dissolved oxygen and pH observations for treated stations by event year. Event years -10 and 10 are omitted since they are binned—all years before or after ten years from designation are coded as belonging to each respective event year. Consequently, these event years contain a comparatively higher number of observations than all other event years. For both outcomes, there are generally higher amounts of median observations in the raw data before designation than after designation.

Figure 7.3: Event time regression of yearly number of dissolved oxygen and pH observations per station on event time indicators using balanced panel of only treated stations with at least one observation in each event year



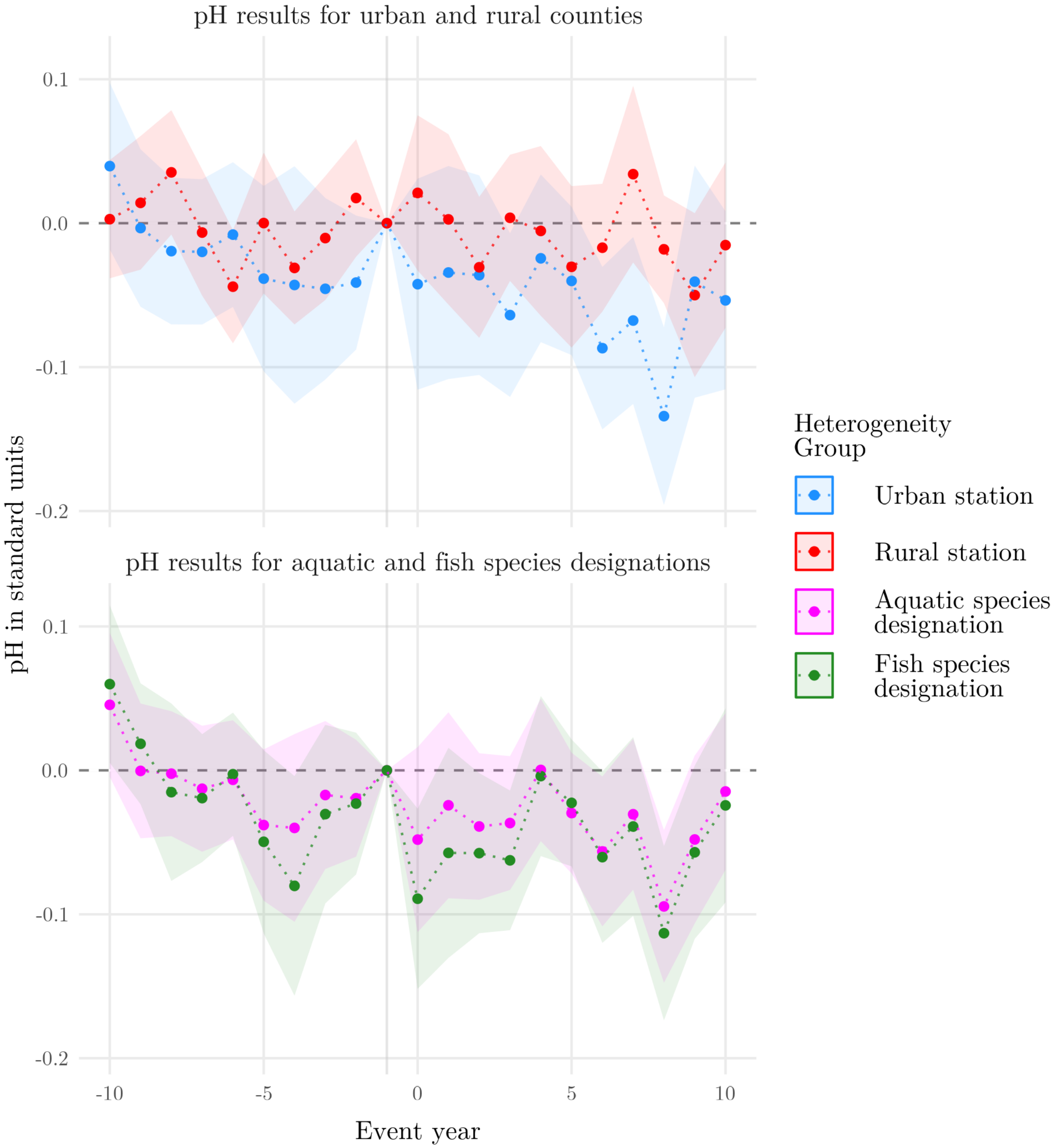
Notes: The figure above shows results from a regression of number of yearly observations per station on a set of event year indicators. The regressions are run on a sample with all untreated stations and only treated stations with at least one observation in each event year. All event time indicators for both outcomes are statistically indistinguishable from zero at a 95% significance level, indicating no correlation between the sampling frequency of treated stations and critical habitat designation after controlling for model fixed effects. All regressions include monitor station and state-by-year fixed effects. Standard error bands around each event year coefficient are clustered at the 8-digit HUC level. Vertical bars at event year -1 indicate that this coefficient is left out of the analysis.

Figure 7.4: Example of comparisons used in baseline specification versus alternate designation group specification



Notes: The figure above shows cross-sectional comparisons made between treated and control stations in the sample area of Minnesota under two specifications. Year-by-state fixed effects compare treated stations to control stations in the same state, as shown in the top image. Alternatively, year-by-designation group fixed effects compare treated and control stations from watersheds surrounding a single species designation, as shown in the bottom image.

Figure 7.5: Event time figures of pH heterogeneity results using baseline specification on urban and rural station and aquatic and fish critical habitat designation subgroups



Notes: The figure above shows event time figures for the results of the baseline model specification on pH for four subgroups. The top image compares effects for urban and rural water stations. The bottom image compares results for aquatic and fish species critical habitat designations. All standard error bands around event year coefficients are clustered at the 8-digit HUC level.

CONCLUSION

This study finds no evidence that ESA critical habitat designations in the contiguous U.S. impact most water quality outcomes on average. However, I find mixed results that indicate potential decreases in pH by as much as 0.09 units 8 years after designations occur. While speculative, one mechanism for this decrease could be the designation's effect on certain types of development such as road construction, which can produce alkaline pollutants. In some circumstances, final critical habitat rulings mention road construction as a threat to species preservation, and some mention specific road construction projects that could adversely affect species. Still, uncovering the specific mechanism behind the effect of critical habitat on pH requires additional study of the effect of designation on mechanisms such as land development, regulatory efforts, and recovery plan success. Additionally, pooled DiD results find no significant change in pH from designation in post-event years, signalling that mechanisms may not result in pH decreases throughout the entire post-period.

Ultimately, the results highlight the limited impacts of ESA critical habitat on water quality outcomes. Given the implied importance of water quality to achieving ESA goals, as indicated by its repeated mention in federal documents on ESA rulings and even in the Act itself, this result is somewhat surprising. However, when considering similar papers on the impact of the critical habitat, the result aligns closely with other findings about the designation's effects. Specifically, Nelson et al. (2017) find that critical habitat does not significantly impact land cover change in their setting, leading them to question if the designation is as effective as other measures to improve species preservation. While water quality improvement is not the main goal of the ESA, this study finds similarly insignificant impacts of the designation on water quality. This adds to a body of research indicating that critical habitat designations may not be the most effective policy measures for achieving outcomes related to species recovery and ecosystem preservation. For instance, other studies illustrate that funding plays a more important role in the recovery process (Ferraro et al., 2007). Other potential policy designs, such as payments or regulatory assurances to private landowners in exchange for conservation, may be promising based on a number of theoretical studies detailed by Langpap et al. (2018). In general, this study reaches a similar conclusion to other studies that find negligible critical habitat impacts, bringing into question if this provision of the ESA should be reformed to focus on more effective drivers of species recovery and, in this case, water quality improvement.

There are several caveats to this study that should inform the conclusions that are made above. First, this study focuses on critical habitat and may not highlight the water quality impacts of other ESA provisions such as listing. Second, although this study uses the largest centralized and publicly available water quality database in its analysis, the sporadic nature of water quality sampling may lead to under-powered estimates and/or a small downward bias in coefficient estimates. If additional data exists in state or other agency databases, merging it with WQP data could provide a more balanced panel and greater precision. Third, critical habitat boundaries for species that have been de-listed are not included in this study since they are not available with current critical habitat boundaries. The exclusion of these designations could lead to attenuated effects if some control units are, in reality, treated; however, the number of missing designations only represent roughly 3% of the amount of designations that are included in the analysis, so any attenuation is likely quite modest. Lastly, extended results indicate that additional water quality outcomes may experience significant changes resulting from fish habitat designations. Thus, the relative insignificance of water quality responses to critical habitat designation found in this study represent average effects for all species' designations. This analysis may uncover different results if conducted using case studies of individual designations or designation groups based on similar species. Additional analysis must be conducted to uncover if designations for fish species result in consistent and more generalized impacts on water quality.

There are several avenues for follow-up research. In terms of highlighting the water quality effects of the ESA, future work could explore the impacts of other provisions, such as funding for species recovery, on water quality outcomes. These studies would identify if other mechanisms lead to greater water quality impacts. Additional work may also investigate water quality responses specifically for fish critical habitat designations. For additional future work on critical habitat, there are many other outcomes and incentives that remain poorly understood. For instance, Langpap et al. (2018) point out that literature is inconclusive about the costs of the consultation process induced by critical habitat. Important questions also remain about the incentives of agents such as environmental consultants who likely benefit when designations are implemented. Determining whether these incentives influence decisions over listing and habitat declarations could be a fruitful area for future study. Also, other environmental outcomes such as water use remain unconsidered and could be addressed in follow-up studies on the impacts of critical habitat.

Ultimately, my results find no evidence that water quality is a significant and previously unconsidered cost or benefit of critical habitat on average. This conclusion informs the larger question of whether or not the ESA represents an efficient trade-off of preservation of natural resources over development by suggesting that water quality benefits of the designation are not substantial. This result also highlights a growing trend in research on critical habitat that finds insignificant effects besides decreases in land values. Given intense criticism from advocates of private property rights and resounding support from environmental groups, the reality of critical habitat's effects may be less consequential than either side of the debate claims. While this claim should be qualified by the fact that not all costs and benefits of critical habitat have been explored, and effects may vary widely based on designation, my study indicates that there is no consistent evidence that the average critical habitat designation meaningfully alters water quality.

REFERENCES

Abildtrup, J., Garcia, S., & Stenger, A. (2013, Aug). The effect of forest land use on the cost of drinking water supply: A spatial econometric analysis. *Ecological Economics*, *92*, 126–136. doi: 10.1016/j.ecolecon.2013.01.004

Atasoy, M., Palmquist, R. B., & Phaneuf, D. J. (2006, Jun). Estimating the effects of urban residential development on water quality using microdata. *Journal of Environmental Management*, *79*(4), 399–408. doi: 10.1016/j.jenvman.2005.07.012

Auffhammer, M., Duru, M., Rubin, E., & Sunding, D. L. (2020, Mar). The economic impact of critical-habitat designation: Evidence from vacant-land transactions. *Land Economics*, *96*(2), 188–206. doi: 10.3368/le.96.2.188

Babbitt v. Sweet Home Chapter of Communities for a Great Oregon, 515 U.S. 687. (1995). Retrieved from `https://supreme.justia.com/cases/federal/us/515/687/`

Bhattarai, G., Srivastava, P., Marzen, L., Hite, D., & Hatch, U. (2008, Jul). Assessment of economic and water quality impacts of land use change using a simple bioeconomic model. *Environmental Management*, *42*(1), 122–131. doi: 10.1007/s00267-008-9111-9

Brown, G. M., & Shogren, J. F. (1998, Aug). Economics of the Endangered Species Act. *Journal of Economic Perspectives*, *12*(3), 3–20. doi: 10.1257/jep.12.3.3

Byl, J. P. (2019, Mar). Perverse incentives and safe harbors in the Endangered Species Act: Evidence from timber harvests near woodpeckers. *Ecological Economics*, *157*, 100–108. doi: 10.1016/j.ecolecon.2018.11.008

Carlson, R. R., Foo, S. A., Burns, J. H. R., & Asner, G. P. (2022, Dec). Untapped policy avenues to protect coral reef ecosystems. *Proceedings of the National Academy of Sciences*, *119*(49), e2117562119. doi: 10.1073/pnas.2117562119

Chen, J., & Lu, J. (2014, Aug). Effects of land use, topography and socio-economic factors on river water quality in a mountainous watershed with intensive agricultural production in east china. *PLoS ONE*, *9*(8), e102714. doi: 10.1371/journal.pone.0102714

Duffy-Deno, K. T. (1997, May). Economic effect of endangered species preservation in the non-metropolitan west. *Growth and Change*, *28*, 26.

Endangered Species Act of 1973. (2003). As Amended through the 108th Congress, 16 U.S.C. §§ 1531-1544.

Ferraro, P. J., McIntosh, C., & Ospina, M. (2007, Nov). The effectiveness of the US Endangered Species Act: An econometric analysis using matching methods. *Journal of Environmental Economics and Management*, *54*(3), 245–261. doi: 10.1016/j.jeem.2007.01.002

Friedman, L. (2022, Jun). Can a law protecting endangered animals stop new oil drilling? *The New York Times*. Retrieved from https://www.nytimes.com/2022/06/15/climate/lawsuit-oil-gas-drilling-endangered-species.html

Gallup. (2022). *Environment.* Retrieved from https://news.gallup.com/poll/1615/environment.aspx

Grainger, C., & Schreiber, A. (2019, May). Discrimination in ambient air pollution monitoring? *AEA Papers and Proceedings*, *109*, 277–282. doi: 10.1257/pandp.20191063

Innes, R., Polasky, S., & Tschirhart, J. (1998, Aug). Takings, compensation and endangered species protection on private lands. *Journal of Economic Perspectives*, *12*(3), 35–52. doi: 10.1257/jep.12.3.35

Keiser, D. A., & Shapiro, J. S. (2019a, Feb). Consequences of the Clean Water Act and the demand for water quality. *The Quarterly Journal of Economics*, *134*(1), 349–396. doi: 10.1093/qje/qjy019

Keiser, D. A., & Shapiro, J. S. (2019b, Nov). US water pollution regulation over the past half century: Burning waters to crystal springs? *Journal of Economic Perspectives*, *33*(4), 51–75. doi: 10.1257/jep.33.4.51

Koenig, L., Platt, L., & Padilla, J. (2023). *Large data pulls from water quality portal - a pipeline-based approach.* Retrieved from https://waterdata.usgs.gov/blog/wqp-large-pull-targets/

Krutilla, J. V. (1967, Sep). Conservation reconsidered. *American Economic Review*, *57*(4), 777.

Langpap, C., Kerkvliet, J., & Shogren, J. F. (2018, Jan). The economics of the U.S. Endangered Species Act: A review of recent developments. *Review of Environmental Economics and Policy*, *12*(1), 69–91. doi: 10.1093/reep/rex026

List, J. A., Margolis, M., & Osgood, D. E. (2006). *Is the Endangered Species Act endangering species?* (Working Paper No. 12777). Retrieved from https://www.nber.org/papers/w12777 doi: 10.3386/w12777

Lueck, D., & Michael, J. A. (2003, Apr). Preemptive habitat destruction under the Endangered Species Act. *The Journal of Law and Economics*, *46*(1), 27–60. doi: 10.1086/344670

Melstrom, R. T. (2021, Jan). The effect of land use restrictions protecting endangered species on agricultural land values. *American Journal of Agricultural Economics*, *103*(1), 162–184. doi: 10.1111/ajae.12127

Nelson, E. J., Withey, J. C., Pennington, D., & Lawler, J. J. (2017, Feb). Identifying the impacts of critical habitat designation on land cover change. *Resource and Energy Economics*, *47*, 89–125. doi: 10.1016/j.reseneeco.2016.12.002

Newburn, D. A., Brozovic, N., & Mezzatesta, M. (2011, Jul). Agricultural water security and instream flows for endangered species. *American Journal of Agricultural Economics*, *93*(4), 1212–1228. doi: 10.1093/ajae/aar039

Plantinga, A. J., Helvoigt, T. L., & Walker, K. (2014, Feb). Critical habitat for threatened and endangered species: How should the economic costs be evaluated? *Journal of Environmental Management*, *134*, 127–135. doi: 10.1016/j.jenvman.2013.10.025

Quigley, J. M., & Swoboda, A. M. (2007, Mar). The urban impacts of the endangered species act: A general equilibrium analysis. *Journal of Urban Economics*, *61*(2), 299–318. doi: 10.1016/j.jue.2006.08.004

Radford, D. (2022, Aug). Lawsuit challenges marina, housing development. *Bonner County Daily Bee*. Retrieved from `https://bonnercountydailybee.com/news/2022/aug/26/lawsuit-challenges-marina-housing-development/`

Rosan, E. (2000). EPA's approach to endangered species protection in state Clean Water Act programs. *Environmental Law*, *30*(2). Retrieved from `https://www.jstor.org/stable/43266768`

Rothenberger, M. B., Burkholder, J. M., & Brownie, C. (2009, Sep). Long-term effects of changing land use practices on surface water quality in a coastal river and lagoonal estuary. *Environmental Management*, *44*(3), 505–523. doi: 10.1007/s00267-009-9330-8

Taylor, M. F. J., Suckling, K. F., & Rachlinski, J. J. (2005). The effectiveness of the Endangered Species Act: A quantitative analysis. *BioScience*, *55*(4), 360. doi: 10.1641/0006-3568(2005)055[0360:TEOTES]2.0.CO;2

Tong, S. T., & Chen, W. (2002, Dec). Modeling the relationship between land use and surface water quality. *Journal of Environmental Management*, *66*(4), 377–393. doi: 10.1006/jema.2002.0593

U.S. CDC. (2017). *NCHS urban–rural classification scheme for counties.* Retrieved from `https://www.cdc.gov/nchs/data_access/urban_rural.htm#Data_Files_and_Documentation` (retrieved from the Centers for Disease Control and Prevention (CDC))

U.S. EPA. (2001). *Policy guidance: Water quality standards and the Endangered Species Act.* Retrieved from `https://www.epa.gov/sites/default/files/2018-10/documents/wqs-endangered-species-act-factsheet.pdf`

U.S. EPA. (2017). *Causal analysis/diagnosis decision information system (CADDIS).* Retrieved from https://www.epa.gov/caddis

U.S. EPA. (2023, Jan). *Regulatory and guidance information by topic: Water.* Retrieved from https://www.epa.gov/regulatory-information-topic/regulatory-and-guidance-information-topic-water

U.S. FWS. (n.d.). *History of the Endangered Species Act: Principal amendments.* Retrieved from https://www.fws.gov/page/endangered-species-act-amendments

U.S. FWS. (1987, Apr). Endangered and threatened wildlife and plants; determination of threatened status and critical habitat for the Waccamaw Silverside. *Fed. Reg.*, *52*(67).

U.S. FWS. (2007, Nov). Endangered and threatened wildlife and plants; designation of critical habitat for five endangered and two threatened mussels in four northeast Gulf of Mexico drainages; final rule. *Fed. Reg.*, *72*(220).

U.S. FWS. (2012, Oct). Endangered and threatened wildlife and plants; designation of critical habitat for the cumberland Darter, Rush Darter, Yellowcheek Darter, Chucky Madtom, and Laurel Dace. *Fed. Reg.*, *77*(200).

U.S. FWS. (2013, Sep). Endangered and threatened wildlife and plants; designation of critical habitat for the Fluted Kidneyshell and Slabside Pearlymussel. *Fed. Reg.*, *78*.

U.S. FWS. (2017). *Critical habitat.* Retrieved from https://www.fws.gov/project/critical-habitat

U.S. FWS. (2022, Dec). *USFWS threatened & endangered species active critical habitat report.* Retrieved from https://ecos.fws.gov/ecp/report/critical-habitat

U.S. FWS. (2023). *Recovery land acquisition grants.* Retrieved from https://www.fws.gov/service/recovery-land-acquisition-grants

U.S. FWS Kentucky Ecological Services Field Office. (2022). *Draft recovery implementation strategy for the Fluted Kidneyshell (ptychobranchus subtentus).* Retrieved from https://ecos.fws.gov/docs/recovery_plan/FlutedKidneyshell_RIS_1.pdf

U.S. NOAA Fisheries. (2022, Nov). *Critical habitat.* Retrieved from https://www.fisheries.noaa.gov/national/endangered-species-conservation/critical-habitat

USGS. (n.d.). *Access National Hydrography products.* Retrieved from https://www.usgs.gov/national-hydrography/access-national-hydrography-products

USGS. (2022). *Hydrologic unit maps.* Retrieved from https://water.usgs.gov/GIS/

huc.html

Wernik, A. (2022, Sep). *Biden challenges California's offshore drilling ban; reversal could affect Santa Barbara Channel.* Retrieved from https://www.kcbx.org/environment-and-energy/2022-09-12/biden-challenges-californias-offshore-drilling-ban-reversal-could-affect-santa-barbara-channel

Wood, J., & Watkins, T. (2020, Oct). Critical habitat's unique private land problem. *The Center for Growth and Opportunity at Utah State University*, *Working Paper 2020.018*. Retrieved from https://www.thecgo.org/wp-content/uploads/2020/10/Critical-Habitats-Unique-Private-Land-Problem.pdf

Zabel, J. E., & Paterson, R. W. (2006, Feb). The effects of critical habitat designation on housing supply: An analysis of California housing construction activity. *Journal of Regional Science*, *46*(1), 67–95. doi: 10.1111/j.0022-4146.2006.00433.x

Zhang, D., & Flick, W. A. (2001, Aug). Sticks, carrots, and reforestation investment. *Land Economics*, *77*(3), 443–456. doi: 10.2307/3147135

APPENDICES

APPENDIX A

SUPPLEMENTARY TABLES AND FIGURES

Table A.1: Pooled difference-in-differences regression results for additional outcomes

	Temperature	Biochemical oxygen demand	Fecal coliform	Salinity	Dissolved nitrate
	(1)	(2)	(3)	(4)	(5)
CH designation	0.133	-0.127	-147.741*	0.319**	-0.079
	(0.139)	(0.102)	(87.984)	(0.161)	(0.199)
Min. temp	0.860***	-0.039***	-58.953***	0.047**	0.011
	(0.063)	(0.009)	(11.033)	(0.020)	(0.017)
Max. temp	0.372***	-0.004	23.439*	-0.072***	0.061***
	(0.029)	(0.015)	(14.078)	(0.019)	(0.019)
Precip.	-0.047***	-0.058***	-332.848***	0.028***	0.029**
	(0.006)	(0.007)	(25.249)	(0.008)	(0.014)
N	11,401,966	508,466	1,189,239	566,659	416,464
Sample mean	11.3	2.7	1,059	6.6	3.8
Year-by-state FE	Yes	Yes	Yes	Yes	Yes
Monitor FE	Yes	Yes	Yes	Yes	Yes
Hour & month FE	Yes	Yes	Yes	Yes	Yes
SE Cluster	HUC8	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table above shows results for the pooled difference-in-differences regression of critical habitat designation on five additional water quality outcomes: temperature, biochemical oxygen demand, fecal coliform, salinity, and dissolved nitrate. Coefficients for biochemical oxygen demand, fecal coliform, salinity, and dissolved nitrate are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

Table A.2: Results for alternate specification using designation group fixed effects for additional main outcomes

	Specific conductance	Total suspended solids	Total nitrogen	Total phosphorus
	(1)	(2)	(3)	(4)
Event Year -5	7.700	-5.037	-0.125	0.010
	(134.068)	(5.197)	(0.139)	(0.010)
Event Year -4	-15.733	7.921*	0.051	0.016
	(118.293)	(4.796)	(0.093)	(0.013)
Event Year -3	-316.108**	-0.004	0.058	0.023**
	(145.764)	(4.780)	(0.091)	(0.010)
Event Year -2	-491.272	4.831	-0.008	0.021**
	(343.482)	(5.486)	(0.079)	(0.009)
Event Year 0	98.778	4.795	-0.010	0.017
	(181.483)	(5.828)	(0.083)	(0.013)
Event Year 1	-315.829*	3.793	0.071	0.014
	(178.068)	(4.688)	(0.089)	(0.018)
Event Year 2	-119.874	7.798	0.008	0.023*
	(99.829)	(4.828)	(0.104)	(0.013)
Event Year 3	111.136	8.681*	0.003	0.034**
	(192.181)	(4.950)	(0.104)	(0.013)
Event Year 4	96.797	10.879**	0.016	0.021*
	(141.451)	(4.364)	(0.112)	(0.011)
Event Year 5	310.832	1.662	0.082	0.028
	(290.574)	(5.836)	(0.136)	(0.017)
Event Year 6	785.640	5.742	0.186	0.025
	(504.338)	(7.173)	(0.172)	(0.023)
Event Year 7	-43.372	-1.892	0.075	0.043**
	(187.448)	(9.861)	(0.185)	(0.018)
Event Year 8	352.348	6.159	0.492	0.070*
	(355.932)	(15.265)	(0.406)	(0.041)
Event Year 9	187.201	6.007	0.410	0.060
	(313.869)	(9.883)	(0.420)	(0.037)
Event Year 10	193.773	22.100***	0.090	0.047*
	(154.838)	(8.106)	(0.123)	(0.026)
N	12,695,878	4,749,503	1,941,509	6,411,590
Sample mean	1,067	42.3	2	0.2
Year-by-des. FE	Yes	Yes	Yes	Yes
Monitor FE	Yes	Yes	Yes	Yes
Hour & month FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
SE Cluster	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table above shows results for the alternate model specification using year-by-designation group fixed effects. Although event year indicators for event years -10 through -6 are also included in these models, they are omitted from this table for brevity. Coefficients for specific conductance, total suspended solids, total nitrogen, and total phosphorus are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

Table A.3: Pooled difference-in-differences results for pH by heterogeneity subgroup

	(1)	(2)	(3)	(4)	(5)
CH designation	-0.007	-0.013	-0.001	0.035*	0.068**
	(0.013)	(0.021)	(0.016)	(0.020)	(0.028)
CH designation*urban	-0.014				-0.040
	(0.027)				(0.027)
CH designation*rural		-0.003			-0.036
		(0.025)			(0.027)
Aquatic CH designation			-0.017		
			(0.025)		
Fish CH designation				-0.093***	-0.095***
				(0.032)	(0.032)
χ^2 test of joint sig. p-value	0.67	0.57	0.65	0.01	0.04
N	4,191,784	4,191,784	4,191,784	4,191,784	4,191,784
Sample mean	7.5	7.5	7.5	7.5	7.5
Year-by-state FE	Yes	Yes	Yes	Yes	Yes
Monitor FE	Yes	Yes	Yes	Yes	Yes
Hour & month FE	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes
SE Cluster	HUC8	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table above shows results for the pooled difference-in-differences regression of critical habitat designation on pH with interaction terms for four subgroups: urban stations, rural stations, aquatic designations, and fish designations. P-values from χ^2 tests of joint significance on the pooled DiD coefficients and interacted coefficients are reported along with coefficient estimates and standard errors. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

Table A.4: Pooled difference-in-differences results for main outcomes in fish designations

	Dissolved oxygen	pH	Specific conductance	Total suspended solids	Total nitrogen	Total phosphorus
	(1)	(2)	(3)	(4)	(5)	(6)
CH designation	0.106	0.035*	-46.842	-2.079	0.013	0.015
	(0.537)	(0.020)	(195.755)	(2.310)	(0.087)	(0.020)
Fish CH designation	-1.179*	-0.093***	251.907	7.636**	-0.160	-0.050**
	(0.704)	(0.032)	(318.351)	(3.397)	(0.127)	(0.021)
χ^2 test of joint sig. p-value	0.24	0.01	0.12	0.07	0.40	0.02
N	3,629,843	4,191,784	3,996,720	1,839,299	779,089	2,122,269
Sample mean	-17	7.5	1,067	42.3	2	0.2
Year-by-state FE	Yes	Yes	Yes	Yes	Yes	Yes
Monitor FE	Yes	Yes	Yes	Yes	Yes	Yes
Hour & month FE	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
SE Cluster	HUC8	HUC8	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table above shows results for the pooled difference-in-differences regression of critical habitat designation on pH with interaction terms for one subgroup: fish designations. Regressions are conducted for six main water quality outcomes: dissolved oxygen, pH, specific conductance, total suspended solids, total nitrogen, and total phosphorus. Coefficients for specific conductance, total suspended solids, total nitrogen, and total phosphorus are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. P-values from χ^2 tests of joint significance on the pooled DiD coefficients and interacted coefficients are reported along with coefficient estimates and standard errors. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

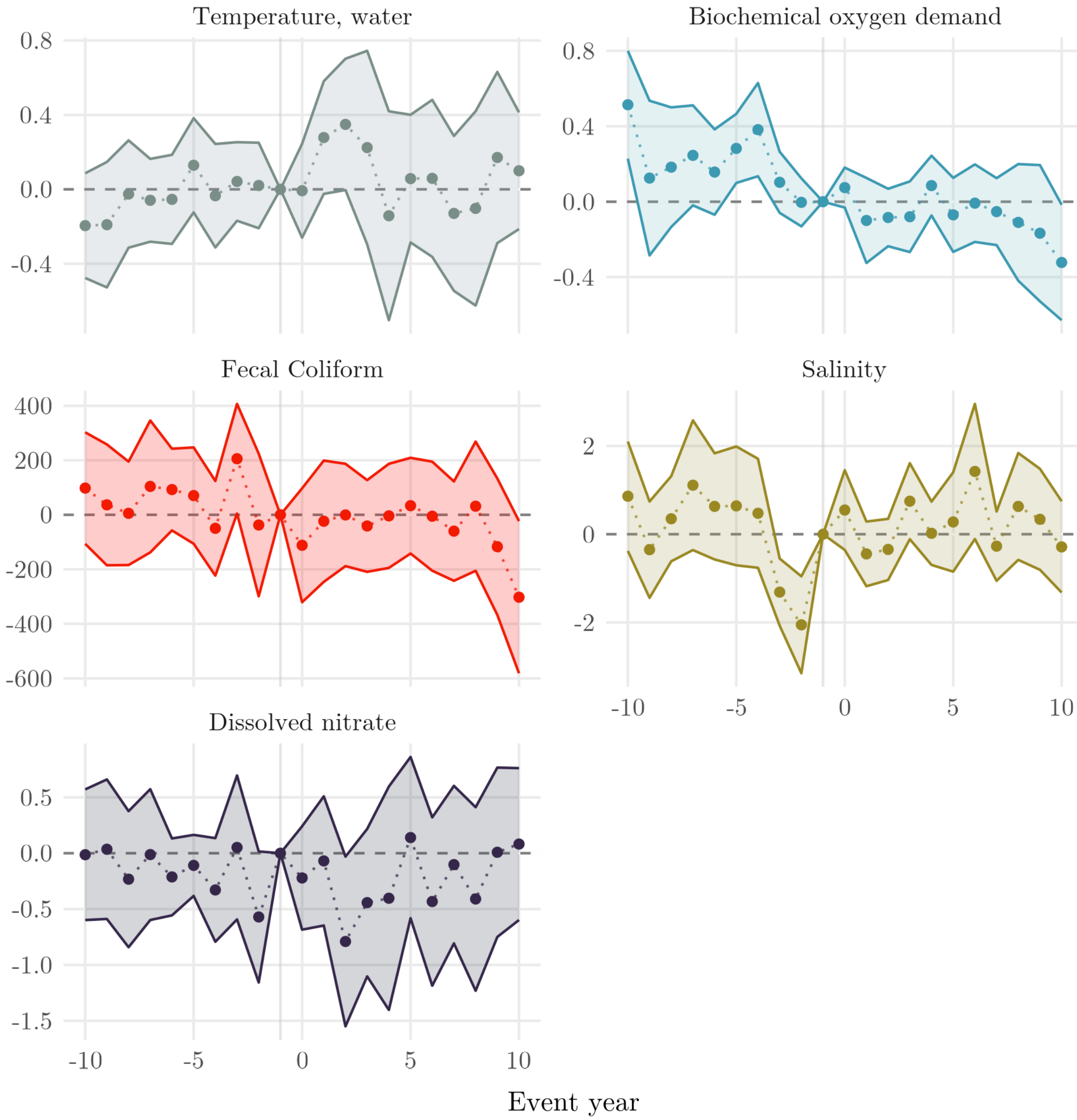
Table A.5: Pooled difference-in-differences results for main outcomes using year-by-designation group fixed effects

	Dissolved oxygen	pH	Specific conductance	Total suspended solids	Total nitrogen	Total phosphorus
	(1)	(2)	(3)	(4)	(5)	(6)
CH designation	-1.265**	-0.027*	57.949	3.148	0.128	0.023
	(0.609)	(0.016)	(52.148)	(5.771)	(0.119)	(0.016)
Min. temp	-1.576***	-0.026***	14.990***	3.715***	0.009	0.011***
	(0.214)	(0.004)	(4.334)	(0.518)	(0.013)	(0.003)
Max. temp	1.666***	0.033***	-27.995**	-0.426	-0.042***	-0.007**
	(0.206)	(0.005)	(11.392)	(0.425)	(0.008)	(0.003)
Precip.	-0.211***	-0.010***	16.683***	4.676***	0.001	0.005***
	(0.032)	(0.001)	(3.471)	(0.966)	(0.006)	(0.001)
N	11,393,451	12,660,140	12,695,878	4,749,503	1,941,509	6,411,590
Sample mean	-17	7.5	1,067	42.3	2	0.2
Year-by-state FE	Yes	Yes	Yes	Yes	Yes	Yes
Monitor FE	Yes	Yes	Yes	Yes	Yes	Yes
Hour & month FE	Yes	Yes	Yes	Yes	Yes	Yes
SE Cluster	HUC8	HUC8	HUC8	HUC8	HUC8	HUC8

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

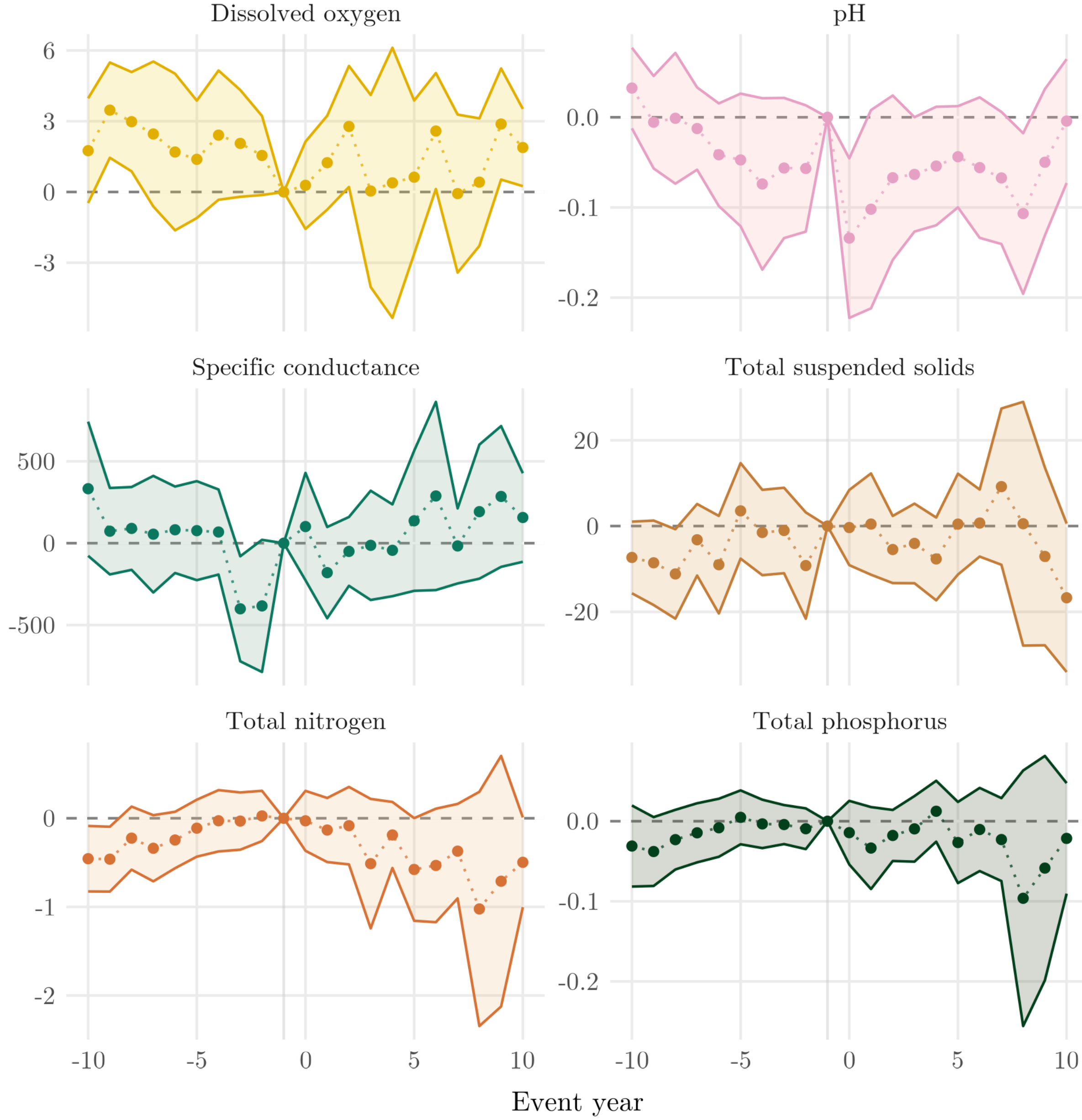
Notes: The table above shows results for the pooled difference-in-differences regression of critical habitat designation on the main water quality outcomes: dissolved oxygen, pH, specific conductance, total suspended solids, total nitrogen, and total phosphorus. The regression substitutes year-by-state fixed effects for year-by-designation group fixed effects. Coefficients for specific conductance, total suspended solids, total nitrogen, and total phosphorus are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. Standard errors clustered at the 8-digit HUC level are shown in parentheses.

Figure A.1: Event time figures of baseline specification results for additional outcomes



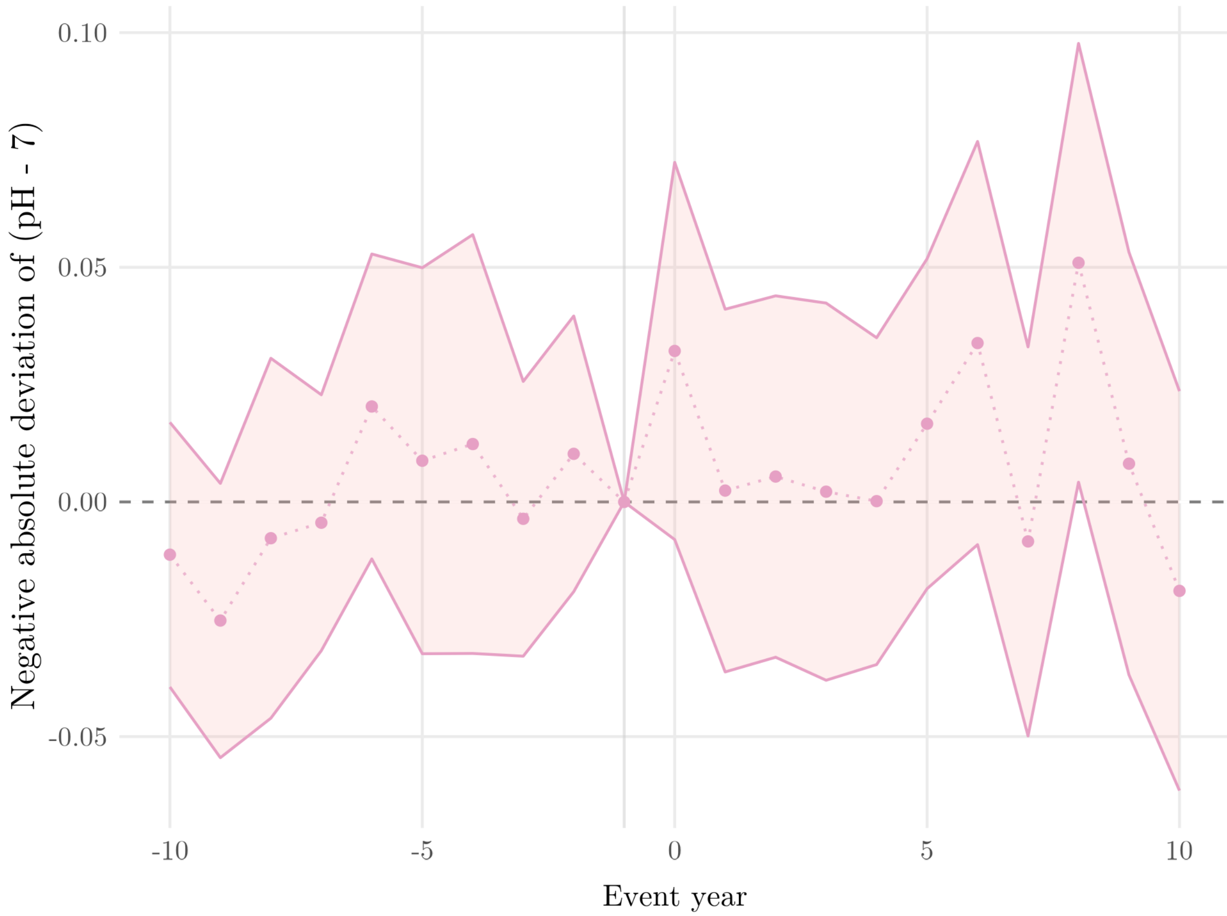
Notes: The figure above shows event time figures for the results of the baseline model specification for additional outcomes: temperature, biochemical oxygen demand, fecal coliform, salinity, and dissolved nitrate. Water temperature is recorded in degrees Celsius; fecal coliform is recorded in MPN per 100 milliliters. Salinity is recorded in parts per thousand. Biochemical oxygen demand and dissolved nitrate are recorded in milligrams per liter. Coefficients for biochemical oxygen demand, fecal coliform, salinity, and dissolved nitrate are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. Standard error bands around each event year coefficient are clustered at the 8-digit HUC level. Vertical bars at event year -1 indicate that this coefficient is left out of the analysis.

Figure A.2: Event time figures of baseline specification results for main outcomes using balanced panel of only treated stations with at least one observation in each event year



Notes: The figure above shows event time figures for the results of the baseline model specification using a sample with all untreated stations and only treated stations with at least one observation in each event year. The regression is a robustness check for the baseline specification results shown in figure 7.1. Results are qualitatively similar for most outcomes. Coefficients for specific conductance, total suspended solids, total nitrogen, and total phosphorus are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, decreases in absolute concentration of each respective outcome. Standard error bands around each event year coefficient are clustered at the 8-digit HUC level. Vertical bars at event year -1 indicate that this coefficient is left out of the analysis.

Figure A.3: Event time figure of baseline specification results for absolute deviation of pH from 7 (neutral)



Notes: The figure above shows event time figures for the results of the baseline model specification on pH using the absolute deviation from 7. The coefficients are multiplied by negative one such that positive coefficients represent improvements in water quality but, consequently, pH levels closer to 7 (neutral). This regression explores the possibility that pH could change heterogeneously across sites, resulting in net-positive or net-negative impacts on water quality that would be masked without examining the absolute deviation of pH from 7. However, results are consistent with the baseline specification and indicate that the baseline model does not mask heterogeneous (but net-positive or -negative) pH effects. Standard error bands around each event year coefficient are clustered at the 8-digit HUC level. Vertical bars at event year -1 indicate that this coefficient is left out of the analysis.

APPENDIX B

DATA APPENDIX

This appendix details complete steps used to download, clean, aggregate, and merge data into panels used for analysis in the results section.

Downloading and Cleaning Water Quality Data

To begin, data are downloaded using the pipeline-based approach developed by Koenig et al. (2023). This approach breaks up data queries in to smaller, manageable requests to avoid back-end issues with large data pulls. The code is adapted to accept a shapefile defining the geographic area of interest for a data query. Other query parameters include the time period of interest and the water quality measurements of interest, defined using characteristic names in the WQP database. Shapefiles for all 48 contiguous U.S. states and the District of Columbia are used to download data by state for the years 1970-2022, even though data beyond 2018 are filtered out of the final analysis due to restrictions in weather data availability. Data is downloaded for 11 outcomes: dissolved oxygen, pH, temperature, biochemical oxygen demand, total phosphorus, total nitrogen, total suspended solids, fecal coliform, salinity, specific conductance, and dissolved nitrate. Since the WQP requires characteristic names that match to water quality outcomes, and each outcome generally has multiple potential characteristic names that match, the following characteristic names were requested from the database after exploring names available in the WQP: "Oxygen," "Dissolved oxygen," "Dissolved oxygen (DO)," "Dissolved oxygen saturation," "pH," "PH," "Temperature, water," "Temperature, water, deg F," "Biochemical oxygen demand, standard conditions," "Phosphorus," "Total Particulate Phosphorus," "Total Phosphorus, mixed forms," "Nitrogen," "Nitrogen, mixed forms (NH3), (NH4), organic, (NO2) and (NO3)," "Total Nitrogen, mixed forms (NH3), (NH4), organic, (NO2) and (NO3)," "Total Nitrogen, mixed forms," "Total suspended solids," "Fecal Coliform," "Condition class, chemical (Conductivity/Salinity)," "Salinity," "Conductance," "Conductivity," "Specific conductance," "Specific Conductance, Calculated/Measured Ratio," "Specific conductivity," "Inorganic nitrogen (nitrate and nitrite)," "Inorganic nitrogen (nitrate and nitrite and ammoni," "Inorganic nitrogen (nitrate and nitrite) as N," "Inorganic nitrogen (NO2, NO3, & NH3)," "Nitrate," "Nitrate as N," "Nitrate-N," "Nitrate + Nitrite," "Nitrate-Nitrite," "Nitrate-nitrogen," and "Nitrate-Nitrogen."

Data is then cleaned by state for all outcomes. Several steps are applied across all outcomes. First, data is filtered to surface water observations by limiting location type to one of the following: "Stream," "Lake," "Impoundment," "River," "Canal," "Floodwater," "Reservoir," "Floodwater," "Pond-Anchialine," or "Pond-Stormwater" and media subdivision name to "Surface Water." Monitoring events are limited to "Routine sample," "Not applicable," or a missing value to remove instances of sampling that occur after an anomolous event. Observations with quality flags are removed using WQP accompanying documentation to identify codes associated with quality issues. All observations with codes excluding: "J," "RNAF," "VVRR," "VVRR2," "RV," and "RP" are filtered out.[1] Observations near dams are filtered out using regular expression searches for monitoring location names containing permutations of the word "dam." Observations with missing site latitude, longitude, 8-digit HUC code, or sample measurement value are removed. Data from the same location and hour of the day is averaged to get a single datapoint for each hour. Except for temperature, data are filtered to keep observations greater than or equal to 0. Lastly, data for each outcome are topcoded to the 99th percentile observation value, similar to Keiser and Shapiro (2019a). pH data is also bottomcoded to the 1st percentile observation value due to the approximately normal distribution of this variable.

Apart from these common data cleaning steps, extensive efforts are made to ensure data for each outcome are combined for sensible, comparable observations only. There are

[1] "J" indicates: "estimated: the analyte was positively identified and the associated numerical value is the approximate concentration of the analyte in the sample." "RNAF" indicates: "result not affected by noted QC issue." "VVRR" indicates: "value verified by rerun." "VVRR2" indicates: "value verified by rerun, 2nd method." "RV" indicates: "Results are within precision limits and are analytically equal." "RP" indicates: "value reported is preferred."

several common steps taken for each outcome. First, measurements taken can represent various fractions of a total water sample based on how the water sample was filtered. If the measurement is based on an unfiltered sample, the fraction is "Total," if the measurement is based on a filtered sample portion, the fraction is "Dissolved," and if a measurement is based on the portion of a sample filtered out, the fraction is "Suspended." Except for nitrates, which were only reserved for "dissolved" sample fractions, observations are only kept if a) the sample fraction for the measurement is recorded as "Total" or b) at least one sample fraction for both "Dissolved" and "Suspended" are recorded in the same hour and are added to get a "Total" sample fraction observation. Next, parameter codes in the USGS portion of the database help to filter observations with comparable units (and sample fractions) of the same water quality outcome. For each outcome, a list of acceptable parameter codes is developed using accompanying WQP documentation, and each outcome is filtered such that parameter codes from USGS observations matched one of the acceptable parameter codes.

For each outcome, observations are again filtered to keep data with units that can be easily converted to a single standard unit. Observations are only converted if there is no measurement of the same outcome in the same hour using the standard unit instead of the non-standard unit. For each outcome, an observation's units fall into three distinct categories: a) the unit is the same as the standard unit for the outcome, b) the unit can easily be converted using a standard conversion formula with no additional necessary information to complete the conversion (i.e. flow rate), or c) the unit cannot easily be converted since additional information is required or the observation's unit appears to be erroneous. As an example, the standard unit for measuring total nitrogen is in milligrams per liter (mg/l). Most total nitrogen observations are measured in mg/l, so they fall into the first category described above. Some observations are recorded in units that are not mg/l, but because they represent weight-per-volumetric unit measurements, they are easily converted to mg/l (i.e. 1000ug/l = 1mg/l or 1g/l = 1000mg/l). Additionally, parts-per-million is a commonly used unit that is identical to mg/l, so no conversion is necessary. These types of units belong to the second category described above. Lastly, all other observations represent the third category described above because they cannot easily be converted to mg/l (i.e. observations recorded in "lb/day" require knowing the flow rate of the surface water source when the measurement was recorded) or they are clearly erroneous (i.e observations recorded in "ug" are dropped because the volumetric aspect of the unit is missing). Dissolved oxygen is commonly recorded in both mg/l and percent saturation, which represents the concentration of dissolved oxygen over the theoretical maximal concentration at the same temperature in a water sample.[2] Thus, 100% saturation at a water temperature of 40 degrees Celcius represents a higher concentration in mg/l than 100% saturation at a water temperature of 50 degrees Celcius. All weight-by-volumetric units are converted to percent dissolved oxygen saturation to, similar to Keiser and Shapiro (2019a), using the same formula: $DO_{percent} = \frac{DOmg/l}{468/(31.5+Temp_{Celcius})}$. Percent saturation is then converted to difference in percent saturation from 100 by subtracting 100 from each observation: $DO_{differencex} = DO_{percent} - 100$. For pH, observations for the unit "mole/l" were converted to standard units using the formula: $pH_{standard} = -log(pH_{mole/l})$. For temperature, observations in Fahrenheit were converted to Celsius using the formula: $Deg_C = \frac{Deg_F-32}{9/5}$.

Merging Critical Habitat Shapefiles and Weather Variables

Shapefiles from critical habitat designations are combined from the FWS and NMFS to form a single dataset with all active critical habitat designations. All designations are buffered on the border by 100m to avoid measurement error caused by slight deviations in latitude/longitude measurements between water quality measurement stations and critical habitat boundaries. Longitude/latitude coordinates of stations and critical habitat shapefiles

[2]Dissolved oxygen saturation is not bounded at 100%, as sudden temperature fluctuations and photosynthetic processes of aquatic life can produce amounts of dissolved oxygen that exceed theoretical equilibriums.

are converted to the US National Atlas Equal Area coordinate reference system projection before merging (ESPG code: 2163). To merge, a dataset with each station's coordinates is spatially joined with critical habitat boundaries by year to determine if a station is ever designated and, if so, on which day. Each time a station intersects a critical habitat boundary, the designation species and taxon are recorded in the final dataset. After recording the designated water quality stations and the designation start dates, event year indicators are created by determining the number of event years each observation was recorded away from the station's designation date. Some stations that are designated multiple times include observations that belong to multiple event years.

Next, weather data is merged by converting grid numbers to longitude/latitude coordinates using a formula provided by Wolfram Schlenker. Then, each water monitoring station's coordinates are compared against each weather station's coordinates in the nearby area to find the three closest weather stations to each water monitoring station. Since Schlenker's data is split up by year and fips code, datasets from multiple fips codes are loaded in at once for some water stations to find the closest three weather stations. Next, the inverse-weighted averages of minimum and maximum temperature and precipitation are calculated for water station in the data to create a single data-point for all three variables for each unique station and day combination in the dataset.

Panels Created for Alternate Specifications and Heterogeneity Results

To create stacked datasets for alternate specifications using designation group fixed effects, each unique species designation that appears in the data is intersected with 8-digit HUC watershed boundaries to determine which watersheds cross each designation. Then, spatial queries are used to determine the watersheds that were adjacent to designated HUCs. Any observations in an 8-digit HUC that is not designated or adjacent to a designation is filtered out. Furthermore, for stations appearing in HUCs that belong to multiple designation groups due to close proximity of critical habitat designations, each observation is duplicated such that control stations can be used in multiple groups' estimations. Lastly, stations in a designation group that are treated only by a separate species were removed. For example, some control stations for the Canada Lynx are designated as critical habitat for the Gray wolf only. These stations are removed from the Canada Lynx designation group.

To examine heterogeneous effects of the designation, final panels are merged with data on urban/rural classifications for U.S. counties in 2013. Since water quality data already includes information on U.S. counties, matching the data is done using county names with no spatial queries required. Urban rural classification data ranks counties into one of six hierarchical categories: large central metro (1), large fringe metro (2), medium metro (3), small metro (4), micropolitan (5), and non-core (6). A county is labelled as "rural" if it is a non-core county, while "urban" counties are those defined by any of the metropolitan rankings (1-4). Additional robustness checks broaden the definition of "rural" to include micropolitan counties, but this does not alter the results. In regressions, indicator variables for "rural" and "urban" counties are interacted with event year variables to analyze heterogeneous effects.

Creating variables for aquatic and fish species designations requires further data manipulation since a fixed indicator variable cannot be created for each water station. This is because stations may receive multiple designations, and there may be some designations representing aquatic or fish species along with others representing non-aquatic or non-fish species. On the contrary, urban-rural classifications are treated as fixed characteristics of stations over the panel. Thus, in the creation of event time variables for aquatic and fish species, two separate classes of event time variables are created: event time indicators for all designations, and event time indicators for aquatic/fish designations. Then, all event time indicators are explicitly included in heterogeneity regressions.