# The differential benefits of market-based water pollution control policy [*]

Zach Raff [†]  Andrew Meyer [‡]  Arthur R. Wardle [§]

January 30, 2025

## Abstract

In this paper, we examine the water quality benefits of Wisconsin's phosphorus rule, which created the most stringent water quality standards for phosphorus in the US. We highlight the differential benefits realized downstream from point sources that comply with the rule through offset trading vs. those that comply via treatment technology upgrade. To our knowledge, ours is the first study to empirically estimate the water quality benefits of water quality markets compared to those of traditional (command-and-control) regulation, over a large spatial scale. We find that Wisconsin's phosphorus rule decreases surface waterbody concentrations of total phosphorus downstream of regulated point sources by 26%. The reductions are larger for facilities with technology upgrades (29%) than those that participate in offset trading (21%), but offset trading represents the cost-effective option. Upgrading facilities pay roughly \$34,000/year in capital costs to produce the same water quality improvements as offset trades that cost \$5,900/year.

**JEL Classification:** H23, K32, L21, Q58
**Keywords:** nonpoint source pollution, point source, total phosphorus, surface water quality, water pollution abatement, water pollution offset trading

---

[*] *Acknowledgements*: We thank Andrew Rosenberg, Ben Gramig; seminar participants at PERC, the University of Kansas; and participants at the AERE summer conference and SEA and MEA meetings, for their helpful comments. We also thank Colin Welch for the research assistance. Raff is grateful for generous funding provided by the Property and Environment Research Center's Lone Mountain Fellowship (summer 2021), which helped support this project. This research was supported in part by the US Department of Agriculture, Economic Research Service. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or US government determination or policy. All remaining errors are ours.

[†] USDA Economic Research Service, Washington, DC; zachary.raff@usda.gov

[‡] Marquette University, Milwaukee, WI; andrew.g.meyer@marquette.edu

[§] UC Berkeley, Berkeley, CA; arw@berkeley.edu

# 1 Introduction

Water pollution control policy in the US relies primarily on the regulation of point source wastewater discharges. These regulations, part of the Clean Water Act's (CWA) National Pollutant Discharge Elimination System (NPDES), have contributed to improvements in surface water quality in the US (Keiser et al. 2019). However, the marginal costs of technological abatement at point sources likely exceed the marginal benefits. Alternatively, nonpoint source water pollution is effectively ignored by the CWA and most state water pollution control policies. As a result, nonpoint source pollution has become the largest source of surface waterbody impairments in the US (Olmstead 2010).

From a regulatory perspective, controlling water pollution from nonpoint sources provides greater marginal benefits at lower marginal costs than at point sources. However, regulating nonpoint sources remains difficult. Instead, the CWA grants state agencies the authority to establish water quality markets, where point and nonpoint sources trade offset credits. In theory, water quality markets are more cost-effective and produce the same or better water quality improvements than traditional command-and-control regulation. EPA and state agencies have therefore encouraged their use for years (USEPA 2003, 2019).

Despite strong support, there are few examples of water quality markets that have brokered enough trades to allow for observational analysis. However, there exists comparatively many trades in Wisconsin because of recent regulation. In 2010, Wisconsin revised its phosphorus water quality rules to adopt the most stringent standards in the country, known as Wisconsin's "phosphorus rule" (Meyer and Raff 2024).[1] Under the rule, hundreds of point sources were subject to stringent water quality-based effluent limits (WQBEL), which required most point sources to reduce their total phosphorus (TP) discharges by over 90% compared to previous technology-based limits. To meet the new limits, point sources must upgrade to tertiary, or filtration, treatment technology. Alternatively, point sources can comply with the rule at oftentimes cheaper cost by participating in water quality markets,

---

[1] Meyer and Raff (2024) discuss the motivation for the rule's implementation and the political economy behind it. Briefly, Wisconsin's economy relies heavily on outdoor, water-based recreation (Skidmore et al. 2023). But much of the state's surface waterbodies are impaired by high TP concentrations (see https://dnr.wisconsin.gov/topic/SurfaceWater/ConditionLists.html), which decreases the state's recreation and tourism dollars. As a result, Wisconsin's state legislature passed the phosphorus rule to combat surface waterbody impairments.

which many pursued.

In this paper, we examine the water quality benefits of Wisconsin's phosphorus rule. We focus on the differential benefits realized from point sources that comply with the rule through offset trading vs. those that upgrade their wastewater treatment technology to comply. To our knowledge, ours is the first study to examine the differential benefits realized through water quality markets vs. the traditional technological regulation (command-and-control) over a large spatial scale, i.e., more than a single watershed. In addition to supporting many trades compared to other programs, Wisconsin's program also allows for causal interpretation. Wisconsin's implementation of the rule rolled out across point sources, e.g., wastewater treatment plants, according to exogenous NPDES permit reissuance dates, providing a natural experiment that allows us to identify the effects of this water pollution regulation on surface water quality outcomes. We also leverage the US stream and river network and the upstream-downstream relationship between different surface waterbody segments to better identify suitable treatment and control groups.

Some background on water pollution offset trading and Wisconsin's phosphorus rule elucidates why the benefits of the rule may be heterogeneous. Economists have long endorsed market-based trading programs as cost-effective pollution control instruments, both in theory and in practice. And policymakers laud as successes among historical air pollution control policies well-known examples, such as the US Acid Rain Program for trading sulfur dioxide emissions (Goulder 2013; Schmalensee and Stavins 2013). The analogous water pollution control policies, which governmental agencies, environmental organizations, and agricultural associations promote (e.g., USEPA 2019), typically allow for point sources to trade water pollution (i.e., discharge) offsets with agricultural or other nonpoint sources.[2] The intuition behind these policies is clear; point sources avoid the high marginal costs of technological pollution abatement by paying nonpoint sources to implement less expensive practices

---

[2] Although often conflated, there are two distinct types of pollution trading markets. First are cap-and-trade markets (like the Acid Rain Program), where there exists an authoritative emission or discharge "cap" for the entire market. Participants in these programs then trade pollution allowances—that collectively equal the cap—with other market participants. Second, offset markets consist of regulated entities purchasing pollution "offsets", typically from non-regulated entities. Total emissions or discharges in these markets are not capped, so the individual permits of regulated entities determine the overall stringency of the market. Trading programs for water pollution control generally take the form of offset markets. For this reason, we refer to trading programs for water quality as "water pollution offset trading".

that reduce pollution elsewhere.[3] In theory, then, water pollution offset trading is more cost-effective than command-and-control technological standards. However, and despite the existence of several water pollution offset trading programs in the US, few trades occur in practice.[4]

In addition to representing a useful setting to identify the effects of water pollution regulation on surface water quality, Wisconsin also provides a good setting to study water pollution offset trading. Wisconsin's Department of Natural Resources (WDNR) administers a program that comparatively facilitates many trades because of the stringency of the state's phosphorus rule. The regulation resulted in point sources experiencing large changes in their permitted TP discharge limits, often from 1 mg/L to 0.075 mg/L. As a result of the sizable increase in costs to comply with their discharge limits via treatment technology upgrade, many point sources now trade phosphorus credits with other, usually nonpoint, sources.[5]

Several studies catalog water pollution offset trading programs throughout the US, but do not focus on their efficacy (e.g., Woodward and Kaiser 2002; Morgan and Wolverton 2008; Shortle et al. 2021). There exist some ex ante studies that model agricultural and environmental systems to predict the effects of these programs.[6] However, few papers conduct ex post program evaluations of water pollution offset trading programs, because the programs are relatively new and transaction costs in these markets are high (Newburn and Woodward 2012). As a result, many programs do not have enough existing trades to conduct empirical

---

[3] The practices often consist of agricultural "best management practices", such as the use of cover crops or riparian buffers, and are cheaper per unit of phosphorus abated than the installation of wastewater treatment technology at the point source.

[4] The inherent right to pollute given to nonpoint sources, the localized and uncertain nature of water pollution, and high transaction costs all contribute to this lack of trades (Fisher-Vanden and Olmstead 2013; Raff 2022).

[5] The CWA, which state environmental protection agencies primarily administer, allows for states to establish trading markets for individual point sources to use to comply with their discharge permits. Wisconsin is one such state that administers a water pollution trading program.

[6] For example, Fleming et al. (2020) use an integrated assessment model to predict the effects of Maryland's proposed water pollution offset trading program. The authors focus on the behavioral responses of farmers, finding that the trading program would likely decrease the effectiveness of existing conservation subsidy programs. Rabotyagov et al. (2014) compare trading programs with command-and-control and performance standards. The authors simulate the policies in an agricultural watershed in Iowa and find trading programs would produce cost-effective outcomes for nitrogen abatement, although the simulations suggest that trading programs may not attain the regulated abatement requirement. Ghosh et al. (2011) simulate the effects of baseline requirements in the Conestoga watershed in Pennsylvania and find that such requirements can discourage trades.

analyses. Ex post evaluations are generally limited to case studies of small-scale (i.e., single watershed) programs describing the geographic pattern of trades (Saby et al. 2021), the best management practices (BMP) that nonpoint sources use to generate credits (Newburn and Woodward 2012), bidding behavior (Newburn and Woodward 2012), market response to baseline stringency (Ribaudo and Savage 2014), social context and farmer participation (Breetz et al. 2005), and transaction costs (Fang et al. 2005; Stephenson and DeBoe 2016). In general, there exists a sizable gap in the literature on the ex post performance of water pollution offset trading programs relative to the alternatives.

Wisconsins's phosphorus offset market is also an illustrative case in understanding the potential of offset markets for environmental problems more generally. Despite their theoretical benefits in reducing the total cost of pollution abatement, numerous practical challenges can disrupt the real-world efficacy of pollution offset schemes (Wardle 2023), including establishing additionality (Mason and Plantinga 2013), producing realistic expectations of how much abatement offset practices will achieve, and monitoring for compliance. Though these issues have each received theoretical treatment in the literature, ex post evaluation of offset markets for many pollutants is often difficult due to the regional or global mixing of emissions/discharges, making empirical work on offset efficacy rare (see, e.g., Probst et al. (2023) for a review of carbon offsets' empirical assessment). The Wisconsin phosphorus offset market offers a near best case scenario, with explicit and reasonably defined additionality requirements, trade-by-trade administrative review, and geographically restricted effects. The Wisconsin phosphorus offset market is therefore a promising test case for offset markets generally.

We contribute to the literature in two key ways. First, we identify the effects of Wisconsin's phosphorus rule on surface water concentrations of TP. Prior work examines the costs of the phosphorus rule and its effect on sewer utility bills (Meyer and Raff 2024). Other, related work, examines ambient water quality (Chakraborti 2016; Raff and Meyer 2022) and water pollution control policy (Earnhart and Harrington 2014; Cohen and Keiser 2017; Keiser and Shapiro 2019), in general. Ours is the first study to provide an ex post program evaluation of the water quality benefits of a sizable regulatory program, using plausibly exogenous variation in compliance timing to estimate causal effects. Second, Wisconsin's phosphorus rule allows point sources to comply with the rule by either upgrading their wastewater treatment technology to expensive tertiary filtration systems or engaging in water pollution offset trading. We extend our analysis to estimate the differential water quality benefits of Wisconsin's water pollution offset trading program relative to a technological compliance option. Does Wisconsin's water pollution offset trading program empirically improve water quality outcomes, i.e., do trading programs produce benefits commensurate with those of command-and-control regulation? We use the regulatory setting in Wisconsin to provide the first causal estimates of the efficacy of a water pollution offset trading program compared to a technological abatement option, over a large spatial scale. And importantly, we assess the cost-effectiveness of each option by estimating the compliance cost \$/unit of water quality improvements generated.

We gather data from several sources to develop our contributions. First, we gather water quality data from the Water Quality Portal(WQP). Second, we gather from WDNR compliance option reports. We collect from these reports—which WDNR requires for all point sources subject to the phosphorus rule—the possible compliance options (e.g., offset trading, technological upgrade), their costs, and the path to compliance for each source. Also from WDNR and these reports, we quantify the TP discharge reductions and trade ratios required as part of each facility's compliance plan.[7] Third, we use EPA's Enforcement and Compliance History Online (ECHO) database for each facility's permit issuance and reissuance dates. Importantly, the NPDES permit process results in temporal variation of each point source's permit reissuance. We therefore use this as a source of identifying variation, which is a common challenge when studying water pollution regulation (Keiser and Shapiro 2019). Finally, we use the National Hydrography Dataset (NHD) to identify water quality monitoring locations upstream and downstream from each point source or trading location.

To estimate the water quality benefits of the phosphorus rule, we first use an upstream-downstream difference-in-differences (DD) research design to compare changes in water quality outcomes near point sources subject to Wisconsin's phosphorus rule with changes near point sources not subject to the rule. Like recent work using water quality outcomes, (e.g.,

---

[7] Because facilities complete the reports prior to achieving compliance with the phosphorus rule, we lack data on the actual phosphorus discharge reductions post-implementation. We discuss this further in subsequent sections.

Keiser and Shapiro 2019; Taylor and Druckenmiller 2022), we leverage Wisconsin's stream and river network to estimate the differences between average TP concentrations upstream and downstream of each facility regulated by the rule. For all point sources subject to the rule, treatment in our setting is determined by exogenous temporal variation in each point source's NPDES permit reissuance date. Then, we examine the differential impacts on water quality by compliance option. Importantly, we restrict our sample to vary tight geographies so that we can identify impacts from offset trades that, overall, cover only a small part of the landscape. Our empirical approach requires that control waterbodies are upstream and in a different HUC12 subwatershed than the regulated facility. This setting ensures that waterbody segments upstream of the discharging facility that are "treated" via offset trading do not bias our estimates.

Applying the methods of the recent two-way fixed effects (TWFE) literature (e.g., Borusyak et al. 2021), we find that Wisconsin's phosphorus rule significantly decreases TP discharges from regulated point sources and improves downstream water quality. TP concentrations downstream of point sources regulated by the rule decrease by 26% after facilities are subject to the rule, compared to the counterfactual of upstream waterbodies (in a different HUC12) and point sources not yet regulated by the rule. We also find that water pollution offset trading throughout Wisconsin provides compliance cost savings of roughly \$6.4 million per year over treatment technology upgrades. Our analysis of compliance options suggests that technological upgrades produce larger improvements in downstream water quality than offset trades, suggesting that offset trading ratios are set too low. Compliance with the phosphorus rule via technological upgrade results in downstream TP concentration reductions of over 29%. For facilities that comply via water pollution offset trading, the downstream TP concentration decreases are on average 21%. Perhaps most important, we use our estimates to identify the cost-effective option. Upgrading facilities expend nearly \$34,000/year in capital equipment costs, on average, to produce a reduction in downstream TP concentrations of 0.01 mg/L. For this same downstream reduction via offset trades, it costs only \$5,900/year. These costs imply that the same water quality improvements from offset trades are possible at only 17% of the cost of technological upgrades. Our results show that stringent TP regulations in Wisconsin successfully improved water quality directly downstream of affected point sources. And that water pollution offset trading produces noticeable improvements in surface water quality in a cost-effective way.

The rest of this paper proceeds as follows. Section 2 describes Wisconsin's phosphorus rule. Section 3 describes water pollution offset trading in the US and Wisconsin's program. Section 4 presents a conceptual framework that guides our empirical analysis. Section 5 provides an analysis of how Wisconsin's phosphorus rule impacts surface water quality in the state. Section 6 investigates heterogeneity in the effects of the phosphorus rule on downstream TP concentrations by compliance option. Finally, Section 7 concludes.

## 2 Wisconsin's phosphorus rule

This section describes the necessary background and regulatory setting of our study. First, we describe the statutory framework of Wisconsin's phosphorus rule. Then, we provide information about point source compliance with the rule, including compliance options, requirements, and schedules.

Effective December 1, 2010, Wisconsin jointly implemented administrative code NR102 (Water Quality Standards for Wisconsin Surface Waters) and administrative code NR217 (Effluent Standards and Limitations for Phosphorus). NR102 established a set of water quality standards for surface waterbodies in the state. NR217 established point source discharge limits, schedules of compliance, and alternative compliance options such as water pollution offset trading. The administrative codes jointly comprise the "phosphorus rule", which is one the most stringent regulatory policies controlling TP water pollution in the country (Meyer and Raff 2024).[8] Affected point sources in Wisconsin experienced sizable changes in their TP discharge limits because of the phosphorus rule. Most affected point sources discharge to streams and rivers that have TP concentrations above the NR102-regulated WQBEL of 0.075 mg/L. As a result, they were required to meet TP discharge limits of 0.075 mg/L, which WDNR considers "stringent" limits and require tertiary treatment technology, or filtration systems, to attain. Previously, the same point sources faced technology-based effluent lim-

---

[8] Wisconsin's phosphorus rule also includes administrative code NR151, which manages nonpoint source runoff pollution. NR151 is not applicable in the present study.

its (TBEL) of 1 mg/L. Point sources can attain TBELs with only secondary, or biological treatment technology. As a result of the phosphorus rule, NPDES permit compliance costs for point sources increased considerably. Because WDNR issues NPDES permits in five-year cycles (by specific date), point sources are subject to the phosphorus rule at different times depending on their initial permit date. For nearly all facilities in our sample, the initial permit date was determined many years prior to implementation of the rule because there are few new point sources in the state. Our primary source of identifying variation is therefore the exogenous reissuance of NPDES permits after the promulgation of the 2010 rule. Given this timeline, most point sources in the state faced an initial compliance timeline of between seven and nine years.

Rather than mandate technological upgrades to attain the new TP discharge limits, administrative code NR217 provides two primary options for point sources to comply with the rule: 1) upgrade from secondary to tertiary treatment technology or 2) water pollution offset trading, where point sources offset their pollutant load by reducing phosphorus pollution elsewhere in the watershed.[9] We discuss the water pollution offset trading compliance options in greater depth in the subsequent section.

As part of NR217 and the administration of Wisconsin's NPDES program, WDNR gives point sources a flexible compliance schedule. First, WDNR requires facility-level reports three years following the first NPDES permit reissuance of each facility that faces a TP WQBEL after the promulgation of the 2010 laws. (For example, consider a NPDES permit holder whose active permit, at the time of the 2010 phosphorus rule, expired on April 1, 2012. As part of its new NPDES permit, WDNR would require that the facility develop a third-year report by April 1, 2015.) The third-year report, also called the preliminary compliance alternatives plan, outlines the potential options for the point source to comply with the phosphorus rule. These plans, which contracted environmental engineering consultants produce for each point source, contain a list of the facility's phosphorus rule compliance

---

[9] There are other, less frequently used compliance options. Such options include regionalization or discharging to a new surface waterbody. In rare cases, point sources may request a variance if all compliance options are too costly and would result in an economic hardship for an industry or community. EPA must grant approval for WDNR to authorize a variance. Meyer and Raff (2024) discuss variances in more detail, but we exclude facilities complying with the phosphorus rule via variances or other rarely used options from our analysis.

options and their associated costs. For example, a single compliance alternatives plan can include cost estimates for water pollution offset trading, treatment technology upgrades, plans to combine with another point source (i.e., regionalization), or other options that allow the point source to comply with the phosphorus rule. For compliance options that contain offset trades, WDNR factors in the uncertainty of nonpoint source pollution reduction through "trading ratios". Depending on several characteristics of the offset trade, e.g., practice type, proximity to point source, the third-year reports calculate a ratio (from a formula from WDNR) that outlines the TP load reductions from offsets necessary to meet the required TP load reductions from the point source (e.g., two pounds of offset reductions are necessary for every one pound of reductions from the point source; we discuss trading ratios in depth in subsequent sections). Many of the preliminary compliance alternative plans also contain a compliance option recommendation, based almost exclusively on the least cost option.

The final compliance alternatives plan, also known as the fourth-year report, is due four years after the first NPDES permit reissuance following the promulgation of the 2010 phosphorus rule. For point sources that comply with the phosphorus rule through water pollution offset trading, the fourth-year reports lay out the facility's chosen path for compliance. For these facilities, WDNR refers to the fourth-year report as the water pollution offset trading or adaptive management plan. The report contains the full, final plan for complying with Wisconsin's phosphorous rule through the water pollution offset trading option, which includes the necessary TP load reductions, trading partners, locations and types of trades, e.g., agricultural BMPs, and trading ratios. For facilities that comply via tertiary treatment technology upgrade, the fourth-year reports must include a final engineering design report and facility plan. WDNR must approve the final compliance alternatives plan, including its individual components. Finally, five years after the first NPDES permit reissuance following the promulgation of the 2010 phosphorus rule, WDNR reissues the next permit for the second term following rule promulgation. This second term permit contains the approved compliance components.

# 3 Water pollution offset trading

Our interests primarily lie in whether water pollution offset trading programs, which allow point sources to abate at lower marginal cost than through treatment technology upgrade, empirically deliver water quality benefits. In this section, we first discuss water pollution offset trading programs in the US, while summarizing the associated challenges of the programs documented in the literature to date. We then provide more information about Wisconsin's water pollution offset trading program in the context of the phosphorus rule.

## 3.1 Water pollution offset trading in the US

The CWA requires states to regulate point sources by issuing permits that correspond with effluent limits set by the NPDES program. These limits vary by pollutant and depend on available wastewater treatment technologies. According to the CWA, states must develop water quality standards and assess surface waterbodies within their borders. Many waterbodies do not meet ambient water quality standards for their designated use(s), largely due to nonpoint source pollution, i.e., agricultural and urban runoff. Section 303(d) of the CWA requires states to list as impaired specific waterbody segments that do not meet their designated use(s). States must then develop Total Maximum Daily Loads (TMDL) for waters on the 303(d) "impaired list". A TMDL determines the maximum amount of a pollutant that a waterbody can assimilate without violating water quality standards. TMDLs allocate the allowable load among point sources, nonpoint sources, and a margin of safety.

As noted by Fisher-Vanden and Olmstead (2013), "in almost all water quality trading programs established in the US, the regulatory driver has been the establishment (or anticipated establishment) of a [TMDL]." TMDLs must inventory point and nonpoint pollution sources. While TMDLs are the impetus for water pollution offset trading programs throughout much of the US, the CWA also allows individual states to establish trading programs that can help point sources achieve compliance with their NPDES permits. As a result, water pollution offset trading can occur in both impaired and unimpaired watersheds.

Although point sources face discharge limits through the NDPES permitting process, water pollution from agricultural sources is effectively exempt from CWA regulations. Therefore, regulators have realized that programs to reduce agricultural runoff are important for achieving any substantial improvements in water quality (Olmstead 2010; Fisher-Vanden and Olmstead 2013). Moreover, marginal abatement costs for nonpoint sources are typically low compared to those for point sources. Thus, the largest scope for water pollution offset trading is between point sources with high marginal abatement costs and the largely unregulated nonpoint sources.

In 2003, EPA finalized its water pollution offset trading policy (USEPA 2003), although it had been working on draft frameworks since the 1990s. EPA updated this policy in 2019, again promoting the use of water pollution offset markets (USEPA 2019). Fisher-Vanden and Olmstead (2013) highlight two important aspects of the 2003 EPA water pollution offset trading policy: 1) if a TMDL has been created, all trading must occur within the watershed or defined area of the TMDL and 2) the policy generally supports trading of nutrients and sediment, but trading other pollutants needs prior approval. Trading is intended to facilitate a source's effort to attain additional restrictions from the TMDL.

Although growing, water pollution offset trading programs are relatively new and limited in number. Fisher-Vanden and Olmstead (2013) describe 21 active and pilot programs; 18 of the 21 listed trading programs are in the US. Stephenson and Shabman (2017) and Shortle et al. (2021) list 26 active trading programs in the US, while Selman et al. (2009) identify nearly 60 active trading or offset programs using a wider definition of qualifying programs. Most of these trading programs have a rather small number of market participants. In summary, trading markets are less common and thinner than is optimal.

Fisher-Vanden and Olmstead (2013) identify two classes of factors limiting the success of water pollution offset trading programs. The first class of challenges relates to spatial issues inherent in water pollution. Unlike uniformly mixed air pollutants, damages from water pollution are often heterogeneous based on the discharging location and are more likely to result in pollution "hotspots", making water pollution a more localized problem than air pollution (Doyle et al. 2014). This problem has technical and theoretical solutions such as spatial trading ratios, where policymakers establish exchange rates to reflect varying damages across reductions in different locations (Rodríguez 2000; Fisher-Vanden and Olmstead 2013). A perhaps more fundamental spatial challenge is that water pollution offset trading must occur within a sub-watershed or area defined by a TMDL. This spatial requirement means that many watersheds are limited to few potential trading partners and the scope of efficiency gains is smaller than with a larger geographic area. The second class of challenges relates to the de facto exclusion of agricultural nonpoint sources from CWA regulations. This implicit right to pollute limits the extent of water quality improvements possible through trading since point sources have become relatively small contributors to the overall pollutant load (Fisher-Vanden and Olmstead 2013; Raff 2022).

Another related spatial challenge is modeling and monitoring pollution reductions from nonpoint sources; there is significant uncertainty in the effectiveness of abatement practices from these sources (Raff 2022). Difficulties in establishing baseline pollution levels for nonpoint sources raise concerns about the additionality of credited pollution reductions (Ribaudo and Savage 2014; Ribaudo and Nickerson 2009; Shortle et al. 2021). Most water pollution offset trading programs account for this uncertainty through trading ratios, which require more pollution reduction credits through offset trading than would be required if the point source decreased effluent discharges at the facility itself. As an example, consider a point source that is required to decrease its TP discharges by 200 pounds per year. For this point source, a trading ratio of 1.5:1 requires its reductions through offset trades be at least 300 pounds per year.

Trading ratios determine the relative prices of different compliance options for regulated polluters. When ratios are set too high, money is left on the table by making less expensive compliance options unfeasibly expensive. When they are too low, polluters may "comply" with standards without achieving the desired pollution reductions. In general, because trading ratios are set by regulators, there is reason to expect that the market will clear in the sense of ensuring that marginal compliance with any given compliance pathway provides the same pollution reduction for equal marginal cost. Trading ratios are common among many pollution markets, but the ratios themselves vary between markets and regulators rarely update them once in place.

Finally, multiple case studies document the institutional challenges of water pollution offset trading programs. Woodward (2003) examines the factors that impeded trades in the Lake Dillon Reservoir in Colorado. Jarvie and Solomon (1998) review similar difficulties in the earliest example of water pollution trading in Wisconsin's Fox River program. One barrier cited in these case studies is high transaction costs. DeBoe and Stephenson (2016) quantify transaction costs for a trading program in Virginia and find relatively low costs for land conversion projects but predict high costs for an expanded program that allows credits for agricultural BMPs.[10]

## 3.2 Wisconsin's water pollution offset trading program

Next, we discuss WDNR's water pollution offset trading program. Wisconsin's adoption of strict TP WQBELs presents point sources in the state with a better opportunity to participate in the water pollution offset trading program because the cost of traditional compliance options is often prohibitive (tertiary treatment technology upgrade) or the compliance options are ineffective at meeting a WQBEL (chemical treatment). Our examination of the Wisconsin program enables us to analyze observational data on the program that do not exist in other contexts, where high transaction costs and low traditional compliance costs make trading rare.

To participate in Wisconsin's program, point sources must submit as their fourth-year report a water pollution offset trading plan. There exist several requirements of this plan for WDNR to approve it and write offset trading into the point source's next NPDES permit. The plan must contain basic information about the offset trades such as the number of credits they will generate (i.e., TP discharge pound reductions), trading partners, trading locations, BMPs or other actions that will produce the pollutant reduction credits, appropriate trading ratios, and timing of the trades. To be approved, the trades must satisfy several specific requirements. Each trade must be within the same HUC12 subwatershed (or TMDL area) and upstream of the point source.[11] Trades must also be additional, meaning that the activity used to generate the credits is not already in use or part of a different conservation program. As a specific example in Wisconsin, much of the state is part of a TMDL. If conservation

---

[10] Raff (2022) provides a summary of these and other challenges and policy recommendations for water pollution offset trading programs.

[11] In exceptional cases, close downstream trades or trades outside of the facility's HUC12 may be allowed, with an added uncertainty factor that increases the required trading ratio. In Wisconsin's program these types of trades are very rare, and we omit them from our empirical analysis.

practices are part of the load reduction requirements of the TMDL, they cannot be used as part of water pollution offset trading. Like most programs, WDNR's program also uses site-specific trading ratios to account for the uncertainty of nonpoint source discharges. The minimum allowable trading ratio in Wisconsin is 1.2:1, but ratios can go as high as 5:1 depending on the type and location of trades.[12]

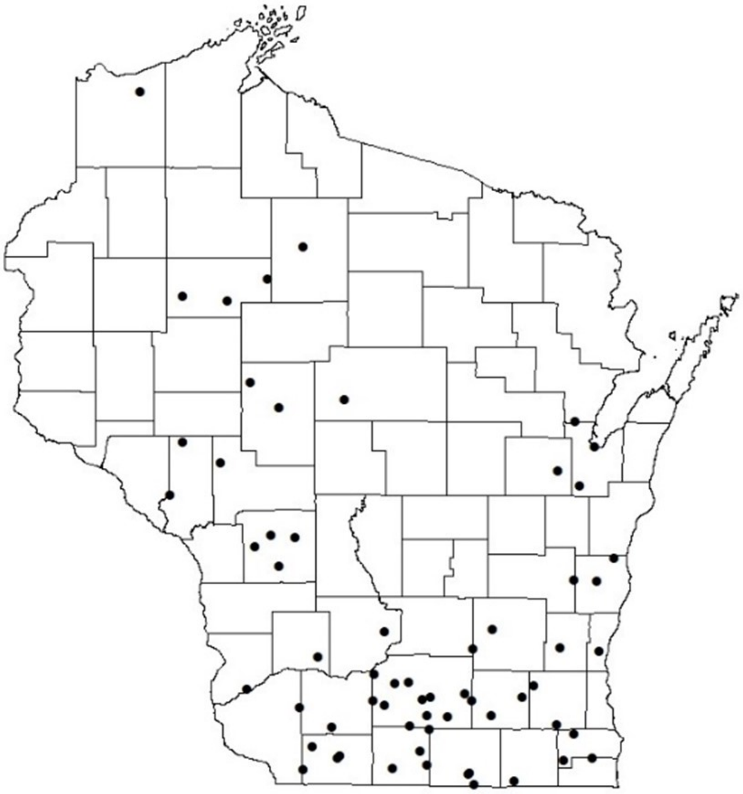


Figure 1: Location of point sources that participate in water pollution offset trading in Wisconsin

*Notes*: Locations identify the point source dischargers in Wisconsin that engage in water pollution offset trading through WDNR to comply with the phosphorus rule.

In Wisconsin, 63 water pollution offset trading agreements are part of WDNR's program. Figure 1 maps the locations of the point sources engaged in these agreements. To present the aggregate savings of Wisconsin's offset trading program (compared to the technological option), we gather from the facility-level third- and fourth-year reports descriptive informa-

---

[12] For example, trades that consist of more uncertain pollution reduction practices can be denied or given a high trading ratio.

tion on WDNR's water pollution offset trading program. (We analyze the program on a micro level in a subsequent section.) From the third-year (and some fourth-year) reports, we collect the compliance option cost data. We supplement these data with estimated costs from WDNR when compliance costs are not included in the facility-level reports. From the fourth-year reports, we collect information on the offset practices that nonpoint sources implement as part of the offset trading program. Finally, we receive directly from WDNR a list of the facilities whose final compliance option is a tertiary treatment technology upgrade.

For trading facilities, we calculate expected yearly capital equipment and trading costs (since the third- and fourth-year reports are completed before compliance options begin). We have data on total tertiary treatment technology costs for each facility, which are typically paid over a 20-year period (Meyer and Raff 2024). We therefore take the estimated capital costs and develop a yearly value, discounting by 5%. In aggregate, the facilities in Wisconsin that comply with the phosphorus rule via offset trading would have had to pay over \$7.8 million per year if they upgraded their treatment technology to comply. Alternatively, these same facilities collectively pay roughly \$1.4 million per year in offset trading costs.[13] Therefore, the program results in cost savings for point sources that comply with the phosphorus rule through water pollution offset trading (rather than through a treatment technology upgrade) of \$6.4 million annually throughout the state. These summaries are consistent with the theory of water pollution offset trading; the marginal costs of abatement at nonpoint sources are lower than those at point sources.

## 4 Conceptual framework

We next consider a brief conceptual framework that guides our empirical analysis.

Consider a network of $i \in \{1, \dots, I\}$ stream and river reaches (i.e., demarcated sections of the stream and river network) with ambient pollution measurements $a_t$. The pollutant concentrations measured at each reach can be expressed as the sum of new pollution discharges within the reach from point sources, nonpoint sources, and natural sources, as well

---

[13] It should be noted that the offset costs typically go to land users in the same watershed as the regulated point source, while upgrade costs go to the technology company.

as natural removals, collectively referred to as $d_t$, as well as flows from other water quality monitoring locations. Let $\mathbf{F}$ be a weighted and directed adjacency matrix (where element $i, j$ is equal to one when $j$ is the only stream reach flowing directly into $i$, and $i, j$ and $i, j'$ are both equal to 0.5 if they are equally sized and both flow into $i$, etc.). The vector $\mathbf{F}a_{it}$ therefore describes pollutant flows into each stream reach in the network. This gives the identity:

$$a_{it} = d_{it} + \mathbf{F}a_{it}. \tag{1}$$

To isolate the dependent variable on one side of this equation, it can be reorganized

$$(\mathbf{I} - \mathbf{F})a_{it} = d_{it}. \tag{2}$$

Therefore, knowledge of the adjacency matrix $\mathbf{F}$ and ambient pollution levels $a_{it}$ allows the study of discharges, even without complete discharge data. The left hand side of this equation is interpretable as subtracting each stream reach's ambient pollutant reading by its inflow reach's pollutant readings.

Consider further decomposing $d_{it}$ into $\tilde{d}_{it}$ ("baseline" pollution without a change in effluent policy), $c_{it}$, the mandated reduction in pollution to be met via non-offset changes at the point source, and $o_{it}$, mandated reductions to be met with offsets. Then,

$$(\mathbf{I} - \mathbf{F})a_{it} = \tilde{d}_{it} + \beta_1 c_{it} + \beta_2 o_{it}, \tag{3}$$

where $\beta_1 = \frac{1}{\gamma}\beta_2 = -1$ if compliance with the water quality program is perfect and $\gamma$ is an offset trading ratio. Econometricians lack information on $\tilde{d}_{it}$, but can decompose it further into fixed effects for stream reaches (or reach-seasons) and controls for precipitation, and, importantly, a composite error term interpretable as changes in water pollution loading not attributable to compliance with standard changes. Causal identification will come from independence assumptions with respect to this composite error. Testing the hypothesis $\hat{\beta}_1 = \frac{1}{\gamma}\hat{\beta}_2$ will reveal if the offset market is awarding the correct amount of credits per offset, and tests of the hypotheses $\hat{\beta}_1 = -1$ and $\hat{\beta}_2 = -\gamma$ will reveal whether regulated entities comply with the phosphorus rule.

# 5 The effect of Wisconsin's phosphorus rule on surface water quality

We next examine the water quality benefits of Wisconsin's phosphorus rule. To do so, we estimate the impact of the rule on average surface water concentrations of TP directly near affected facilities. In this section, we first discuss the data. Second, we describe our treatment definition and identification. Third, we present our empirical model specification. Fourth, we present the results, including an event study. Finally, we present the results of a randomized inference placebo test.

## 5.1 Data

We create our analysis dataset by merging together information from several sources. First, we gather NPDES permit data for all point sources in Wisconsin from EPA's ECHO database. To determine the phosphorus rule compliance timeline for each point source, we use the ECHO data to identify active permit dates, including permit expiration and new permit reissuance dates. We can identify from the permit dates and the statutory language of NR217 when each point source must comply with the phosphorus rule.

Next, we gather compliance option and water pollution offset trading information from the third- and fourth-year reports, respectively, that point sources regulated under the phosphorus rule submit to WDNR. These plans identify the treated group, i.e., the point sources that must comply with the phosphorus rule. In addition, the fourth-year reports identify the compliance option for each facility (e.g., water pollution offset trading, treatment technology upgrade).

For measures of surface water quality, we collect data from the WQP. Several recent studies in environmental and natural resource economics also use these data to study surface water quality and nonpoint source pollution (Keiser and Shapiro 2019; Paudel and Crago 2021; Raff and Meyer 2022; Skidmore et al. 2023; Meyer et al. 2024). The National Water Quality Council aggregates data from the US Geological Survey (USGS) National Water Information System (NWIS), EPA Storage and Retrieval (STORET), and USGS Biodata to provide in the WQP. These water quality sources contain measurements of the presence (i.e., concentration) of water quality indicators and the location and timing of the monitoring. Many measurements are taken as part of research studies at universities or government programs but they also contain samples that citizen scientists and volunteers collect. As our outcome of interest, we gather TP concentrations because the phosphorus rule regulates TP discharges. We restrict our collection of water quality readings to streams and rivers because very few point sources discharge to non-moving surface waterbodies, i.e., lakes, reservoirs,[14] and we wish to use the stream and river network of the NHD. The monitoring data contain some zero and non-detect data as well as very high readings. For the former, we transform the measurements to 1/2 of the smallest value in the sample, like previous studies (e.g., Keiser and Shapiro 2019).[15] For the latter, we winsorize readings at the 99% level.

It is possible that samplers are more likely to collect water quality information on certain days, such as those following precipitation events. We note that this concern is unlikely to bias our analysis. Raff and Meyer (2022) show that sampling timing is not endogenous to precipitation, using a similar analysis sample as ours. Nevertheless, we avoid overweighting more frequently monitored locations and smooth daily noise in the data by aggregating surface water quality readings to the monthly level. We separately average all TP readings along waterbody segments that are within a certain distance upstream and downstream of point sources (i.e., two observations per facility-month). The outcome in our empirical analysis is the average TP concentration upstream and downstream of a point source at the monthly level.

Finally, we gather information on Wisconsin's stream and river network (Figure 2), specifically the upstream and downstream relationship between different stream reaches, from the NHD. The NHD is a national geospatial surface water framework jointly developed by EPA and USGS. The NHD contains data on watershed and catchment boundaries, stream and river flow paths, and other information about the US hydrological network, such as the locations of USGS gage stations. The NHD provides information about the US streamflow

---

[14] This is likely because the discharge limits for these waterbodies are much more stringent than those for point sources discharging to streams and rivers.

[15] Results are robust to alternative substituted values and to an alternative random effects Tobit model that retains non-detect information.

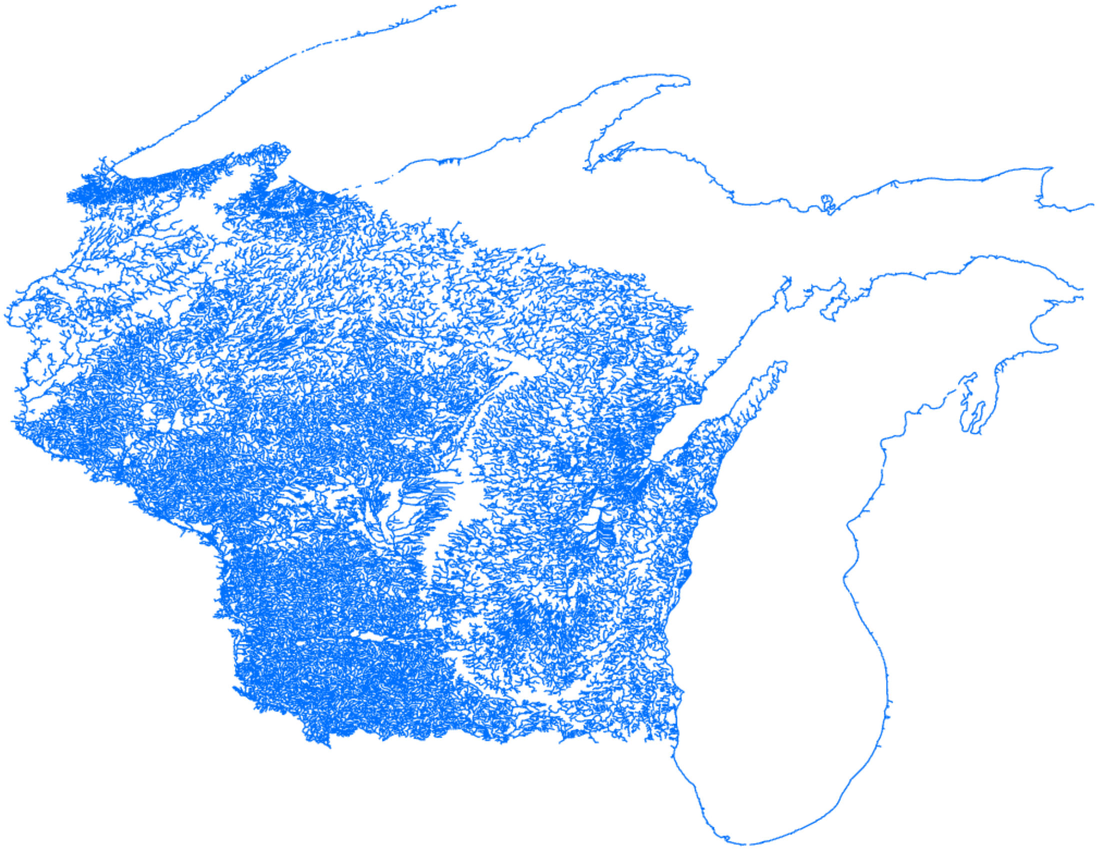


Figure 2: Stream and river network in Wisconsin

*Notes:* Blue lines represent the stream and river network in Wisconsin.

network at the stream reach level for all perennial streams and rivers in the conterminous US. As the primary unit of analysis important for our purposes, the NHD delineates stream and river segments at the "stream reach" level. A stream reach is a section of each waterbody, typically one to five km in length depending on if it is part of the mainstem or a tributary. The NHD provides information on, among other things, the type of waterbody, its streamflow direction, its outlet, the type of stream reach (e.g., mainstem vs. tributary), and flow. From this streamflow network, we can navigate upstream and downstream from point sources and water quality monitoring stream reaches.

To develop our analysis sample, we first match all water quality monitoring locations in Wisconsin from the WQP to their associated stream reach in the NHD. Concurrently, we match each point source in Wisconsin to the stream reach of its discharging waterbody. We use the interconnected nature of the NHD's stream and river network to aggregate average upstream and downstream TP concentrations within five km of each point source. Our unit of analysis, then, is the five km aggregated stream reach-month level. Importantly, we restrict our analysis sample to only those five km segments of streams and rivers; these segments serve as the treatment (downstream) and control (upstream) surface waterbodies.

Our choice of five km as the treatment and control bandwidth is important. Outside of technological upgrades at point sources, we are interested in the impacts of water pollution offset trades on surface water quality. These offset trades primarily take the form of agricultural BMPs. We do not know the precise location of the offset trades in Wisconsin's program. We do, however, know that they occur upstream of the point sources and within the same HUC12, so we can identify the general location of the offsets. Relevant here, prior work shows that agricultural BMPs and conservation program activity overall comprise a small portion of each watershed. For example, Hsieh and Gramig (2024) show that 3.2% of the average watershed's cropland in the Upper Mississippi River drainage plants cover crops. For offset trades, the percent coverage at the watershed level is even smaller. We therefore wish to restrict the analysis sample as much as possible to ensure that it contains meaningful coverage of the offset practice. Our restriction necessarily decreases the sample size and the number of water quality monitoring locations and samples that we examine. However, we ensure that the surface waterbodies affected by the offset trades represent a meaningfully proportion of surface waterbodies in the analysis sample. This empirical structure is superior to a HUC12 level analysis with a minimal proportion of affected surface waterbodies.

Finally, we combine NPDES permit reissuance dates with the statutory language of the NR102 and NR217 to determine the group of point sources that are regulated or yet to be regulated by the phosphorus rule; this group represents the treated group. For point sources treated by the phosphorus rule (based on the quality of the waterbody receiving their discharges and their NPDES permit reissuance date), we also merge the relevant information from their third- and fourth-year reports submitted to WDNR. As such, we build a monthly panel of aggregated stream reaches upstream and downstream of point sources in Wisconsin. Our period of analysis is from 2008 to 2022.

Table 1 displays summary statistics for the final analysis sample, which contains point

Table 1: Descriptive summaries, final analysis sample

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| *Panel A. Summaries* | | | | |
| TP concentration (mg/L) | 0.159 | 0.237 | 0.000007 | 3.660 |
| Total precipitation (cm) | 2.812 | 1.601 | 0.136 | 10.48 |
| Treatment | 0.0740 | 0.262 | 0 | 1 |
| Comply via water pollution offset trading | 0.346 | 0.476 | 0 | 1 |
| Trading ratio | 1.858 | 0.712 | 1.2 | 3.03 |
| Credits | 153.5 | 101.3 | 25 | 439.8 |
| *Panel B. Outcome summary by treatment status* | | | | |
| TP concentration\|Treatment=1 | 0.101 | 0.161 | 0.00841 | 1.669 |
| TP concentration\|Treatment=0 | 0.163 | 0.242 | 0.000007 | 3.660 |
| Observations | | 2,826 | | |

*Notes:* Summary statistics are at the stream reach-month level and represent observations in the final analysis sample. Treatment is a dummy representing stream reach-months downstream from the point source where the point source must comply with the phosphorus rule.

sources that are (eventually) subject to the phosphorus rule and surface water quality readings within five km upstream or downstream of these point sources. In Wisconsin, TP concentrations in surface waterbodies are comparatively high, which is the primary reason for the implementation of the phosphorus rule. Average TP concentrations in the final analysis sample are 0.159 mg/L, which is considered impaired under NR102. This is intuitive, because stream reaches in the final analysis sample are the waterbodies surrounding point sources subject to the rule, i.e., the TP concentrations in these waterbodies trigger the requirements of the phosphorus rule to take effect. Roughly 7.4% of average stream reach-month observations receive our definition of treatment (described in detail below). 65% of point sources in the final analysis sample comply with the phosphorus rule via a treatment technology upgrade, while 35% comply via water pollution offset trading. The mean trading ratio of 1.858 implies that for the average offset trading agreement, point sources receive one pound of TP discharge reduction credit for every 1.858 pounds of TP discharges to surface waterbodies that they eliminate via offset trading. So, from a mean credit-producing offset trade of 153.5 pounds, the point source would need to reduce upstream TP discharges by over 285 pounds.

Panel B summaries present preliminary evidence that the phosphorus rule decreases TP concentrations in Wisconsin surface waterbodies. For treated waterbodies, the average TP concentration is 0.101 mg/L, while the average TP concentration for control waterbodies is 0.163 mg/L.

## 5.2 Treatment definition and identification

We identify the effects of Wisconsin's phosphorus rule on surface waterbody concentrations of TP using an upstream-downsream DD research design, which is effectively a difference-in-difference-in-differences (DDD) design (Taylor and Druckenmiller 2022). Specifically, we leverage water quality monitoring data, Wisconsin's stream and river network, and the temporal variation in NPDES permit reissuance for affected point sources, and thus, phosphorus rule compliance timelines.

We define treatment in the following way. First, we consider the treated group ($Treated$) as the point sources in Wisconsin that are eventually regulated by the phosphorus rule. This group includes affected facilities that must comply with the rule after our analysis panel ends in 2022. Treated point sources discharge to surface waterbodies that do not meet the relevant water quality standard codified in NR102 (0.075 mg/L for our analysis sample). As a result, these point sources face a stringent WQBEL of 0.075 mg/L, which they can only attain (technologically) through an upgrade from secondary to tertiary treatment technology. Alternatively, point sources can meet the stringent TP discharge limits via water pollution offset trading or other means, such as a variance. Because point sources that receive a variance are inherently different than those that comply with the phosphorus rule in the standard ways (e.g., they are smaller and usually more financially leveraged), we eliminate from our final analysis sample all point sources that do not comply with the rule via treatment technology upgrade or water pollution offset trading.

Second, the "post" treatment period represents the period of our panel where the phosphorus rule regulates point sources in Wisconsin; we denote this period $Post$. Using an identification strategy like Meyer and Raff (2024), we consider the post period as the period five years or more after a point source's first permit reissuance after the promulgation of the phosphorus rule on December 1, 2010. This timing matches the schedule to compliance outlined in NR217, i.e., at this point facilities will have phosphorus rule discharge limits and compliance options written into their NPDES permit. WDNR reissues NPDES permits on five-year schedules by specific date, depending on original permit dates. Importantly, WDNR originally issued nearly all NPDES permits in the state before the legislature implemented the phosphorus rule, so the reissuance dates are plausibly exogenous for each point source. Our primary source of identifying variation is therefore temporal variation in NPDES permit reissuance.

Third, we leverage Wisconsin's stream and river network to better identify the effects of the phosphorus rule on surface water quality. Our identification strategy considers as treated water quality monitoring locations within five km (via the stream and river network) downstream of the point source, denoted *Downstream*. As controls, we consider water quality monitoring locations within five km upstream of that point source. This upstream-downstream definition is appropriate if we wish only to identify the water quality effects from facilities that comply with the phosphorus rule via treatment technology upgrade. For these facilities, we know the exact location of the abatement efforts: the geolocation of the point source. The upstream-downstream waterbody segments are therefore clear. But for facilities that comply via offset trading, we only know that the abatement efforts occur upstream of the point source and within the same HUC12. Identifying treatment waterbodies as downstream of the point source and control waterbodies as upstream of the point source would result in biased estimates. We therefore define control waterbody segments as those in the first five km of the HUC12 upstream of the point source. Figure 3 depicts the treatment and control waterbody segments in relation to each regulated point source. The downstream, or treated, waterbody segments are clear. These are the stream and river reaches that are downstream of all abatement activity that regulated facilities implement, be it through technological upgrade or offset trading. The waterbody segments directly upstream of each regulated facility, which we call the "treatment area", is where the abatement efforts take place for offset trading as a compliance option. We eliminate these segments from the analysis sample because it is impossible to tell if they are upstream or downstream of the practices that are part of offset trades. The third group of waterbody segments, the control group, are those where we are certain that no phosphorus rule abatement activity has taken place.

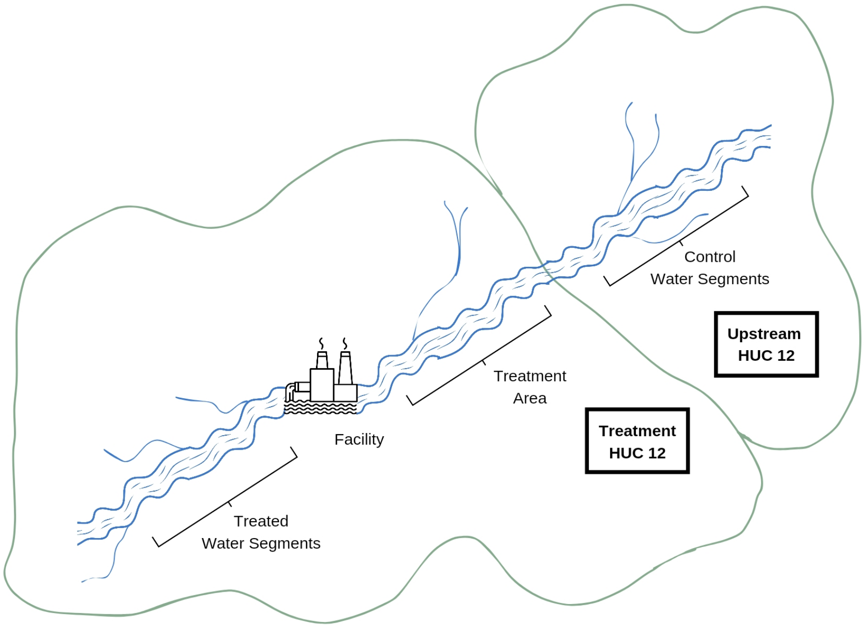


Figure 3: Identifying treatment and control surface waterbody segments

*Notes*: Figure outlines the types of surface waterbody segments as part of our research design. Treated waterbody segments are those downstream of both the point source and the areas where the abatement of offset trades occurs. These segments are within five km of the point source and are directly impacted by all phosphorus rule abatement efforts; they therefore represent the treated group. Treatment areas are waterbody segments that are directly upstream of regulated point sources and may be impacted by the abatement of offset trades. We remove surface waterbodies in the treatment area from our analysis sample. Control waterbody segments are those within the first five km of the HUC12 directly upstream of the point source. These waterbody segments serve as the controls in our analysis.

This within-estimator, which recent work uses (Keiser and Shapiro 2019; Taylor and Druckenmiller 2022), considers as a natural counterfactual upstream water quality that is not impacted by regulated point source abatement activity. Clearly, decreases in TP discharges from either the point source itself or other upstream abatement activity will impact surface water quality downstream of the point source. But the same decreases will have no impact on water quality directly upstream of the abatement location. We therefore leverage variation in treatment timing within a 10-km stretch of streams and rivers to identify our effects of interest, as seen in Figure 3.

The use of a within-estimator allows us to control for time-invariant factors that impact surface water quality near the point source, such as the flow of the waterbody or topological factors. If we fail to incorporate the upstream and downstream nature of average TP concentrations, then control observations (i.e., those not impacted by the phosphorus rule) come from any location in the state, which may have very different unobservable characteristics that impact water quality. We are able to use this identification strategy because of the location of point sources and the sufficient collection of both upstream and downstream water quality data in Wisconsin. Point sources, such as wastewater treatment plants, are located where people live. Alternatively, Meyer et al. (2024), for example, study the effects of animal agriculture on surface water quality. But much animal agriculture is located in the rural upper reaches of watersheds, with little upstream surface waterbodies. Some applications, therefore, are unable to use upstream water quality as a natural counterfactual. To use this identification strategy, we rely on the fact that HUC12s are the smallest hydrological units–each is roughly 36 mi$^2$ (about the size of a township)–so control areas are directly near treated areas. This identification strategy therefore retains the benefits of the upstream-downstream DD, e.g., upstream and downstream waterbody segments are nearly identical in their unobservable factors.

We construct a treatment indicator that is an interaction of these three components, i.e., DDD estimator. Our treatment, which we denote $Treat$ in the empirical model specification, represents 10-km stretches of streams and rivers upstream and downstream of point sources that are regulated by the phosphorus rule, five-years after their first NPDES permit reissuance date after December 1, 2010, i.e., $Treated * Post * Downstream$.

## 5.3 Estimating equation

Using this definition of treatment, we estimate the following regression specification:

$$TP_{imt} = \beta Treat_{imt} + \mathbf{Prec'}_{imt}\mu + \psi_i + \nu_{mt} + \epsilon_{imt} \tag{4}$$

where our outcome, $TP_{imt}$, is the average surface waterbody concentration of TP (in mg/L) within five km upstream (in the next HUC12) or five km downstream of point source $i$ in month $m$ of year $t$. For some specifications, we log-transform the outcome to follow the literature and normalize the skewed distribution of TP concentrations. $Treat_{imt}$ is the DDD indicator that represents stream reach segments downstream of regulated point sources that must comply with the phosphorus rule. The coefficient of interest, $\beta$, represents the effect of phosphorus rule compliance on average concentrations of TP downstream of treated point sources. Untreated entities consist of point sources that must comply with the phosphorus rule but have not yet reached five years after their post-December 1, 2010 NPDES permit reissuance date (including those that must comply after the end of our panel in 2022); point sources that comply with the phosphorus rule through compliance options other than a treatment technology upgrade or water pollution offset trading; and point sources that are never subject to the phosphorus rule because the waterbody receiving the point source's discharges meets the relevant NR102 criterion. Untreated observations from the second and third groups are fundamentally different than treatment observations and therefore do not represent good controls. We therefore remove observations from these groups from all analyses. We rely primarily on treatment timing and the upstream or downstream relationship between water quality outcomes and point sources for identification.

Next, $\mathbf{Prec}_{imt}$ is a vector of precipitation and its square. We include precipitation controls because precipitation obviously affects TP concentrations in surface waterbodies and the square accounts for dilution that occurs during the runoff process. Inclusion of these control factors can reduce the variance of the error term and improve the precision of our estimates. $\nu_{mt}$ represents month-by-year fixed effects, which capture common trends in surface water quality, such as seasonality and other regulatory policies. We capture facility fixed effects with $\psi_i$. These fixed effects control for time-invariant characteristics of each point source, such as its relative size. Because we use a within estimator, $\psi_i$ also controls for time-invariant characteristics of the determinants of surface water quality surrounding each point source, such as its location in the watershed, the number of river and stream miles surrounding the facility (which are important factors for facilitating offset trades), or other topological factors. $\epsilon_{iwt}$ is the exogenous error term. We cluster standard errors at the facility level.

For our empirical strategy, $\beta$ is the DDD coefficient. Assuming several conditions, such as plausibly exogenous treatment assignment and the standard parallel trends assumption, researchers traditionally consider $\beta$ to identify a causal relationship between the treatment and the outcome for the treated group. Recent work demonstrates that TWFE regressions may not recover causal parameters of interest when there are more than two time periods and units are treated at different times. This empirical setting can particularly lead to bias when there are heterogeneous treatment effects across units or when treatment effects are dynamic (De Chaisemartin and d'Haultfoeuille 2020; Borusyak et al. 2021; Callaway and Sant'Anna 2021; Goodman-Bacon 2021; Sun and Abraham 2021; Roth et al. 2023); Roth et al. (2023) provides a thorough review of these estimators. Applicable to our setting where treatment is an absorbing state so that once a unit is treated it remains treated for the duration of the panel, the proposal of Borusyak et al. (2021) presents an "imputation" estimator. Imputation estimators use a two-step process. First, researchers fit a TWFE regression on not-yet-treated units and time periods. Second, researchers predict never-treated counterfactual outcomes. The researcher uses the predicted outcomes to infer treatment effects for each unit and then aggregates them to produce average parameter estimates. To mitigate possible bias in our estimates, we use the imputation estimator of Gardner (2022).

## 5.4 Primary estimation results

Table 2 presents results for the estimation of equation (4). We include in this table results for several specifications and analysis samples to assess the robustness of our primary results. The first column presents results for the DDD, with a log-transformed outcome. The second column presents results for this same specification while using the imputation estimator of Gardner (2022) to account for the staggered nature of our treatment. Finally, column 3 presents results for the outcome measured in levels, again using Gardner's imputation estimator.

Results across these specifications and analysis techniques are qualitatively and quantitatively similar. We focus our discussion on the results from the imputation estimator of Gardner (2022) [columns 2 and 3], which provides the most precise results and best eliminates bias inherent in TWFE regressions. Estimation results show that, for surface waterbodies directly downstream of affected point sources, Wisconsin's phosphorus rule significantly decreases average concentrations of TP. The column 2 point estimate of -0.309 suggests that the rule decreases average downstream TP concentrations by 26%, relative to the counter-

Table 2: Effect of phosphorus rule compliance on surface water quality

| Variable | (1) | (2) | (3) |
|---|---|---|---|
| Treat | -0.211*** | -0.309*** | -0.0294** |
| | (0.0787) | (0.0771) | (0.0138) |
| Facility FE | X | X | X |
| Month#year FE | X | X | X |
| Precipitation controls | X | X | X |
| Gardner two-stage DD | | X | X |
| Outcome scale | Log | Log | Level |
| Observations | 2,826 | 2,826 | 2,826 |

*Notes:* *** p≤0.01, ** p≤0.05, * p≤0.1. Robust standard errors are clustered at the facility level and located in parentheses. The unit of observation is the stream reach-month. Dependent variables in the first two columns are log-transformed, and are in levels in column 3. Precipitation controls include total monthly precipitation and total monthly precipitation squared at the stream reach.

factual of upstream and not yet regulated surface waterbodies in the upstream HUC12.[16] Column 3 results also show that results are robust to the scaling of the outcome. Facility-level phosphorus rule compliance decreases downstream TP concentrations by 0.0294 mg/L, on average, during the post-treatment period. This value is 18.5% of the sample mean TP concentration, or an improvement from 0.159 mg/L to 0.13 mg/L. Although still impaired, even these surface water quality improvements come with sizable recreation and other monetary benefits (Raff and Meyer 2022).

The potential bias of the standard TWFE estimate is likely lower in our setting than in many others. As described in Roth et al. (2023)'s recent review of the literature, the standard TWFE estimator produces biased estimates primarily when different units are treated at different times and there is heterogeneity in treatment effects over time. However, treatment in our setting is largely lumped toward the latter portion of our sample. No point sources are treated until at least five years after the promulgation of Wisconsin's phosphorus rule on December 1, 2010. As a result, our empirical framework more closely resembles a 2x2 fixed effects model than a setting where treatment occurs throughout the panel.

Contrary to popular opinion before the passing of the phosphorus rule,[17] our estimation

---

[16] Throughout this paper, we interpret the treatment effects of a binary variable on a log-transformed outcome using exp($\beta$)-1%.

[17] See, for example, this Milwaukee Journal Sentinel op-ed: https://archive.jsonline.com/news/opinion/101513839.html.

results present evidence that there are sizable benefits from stringent TP discharge limits in Wisconsin, at least to stream reaches directly downstream of affected point sources. Thus far, however, we have only been able to identify the average effects downstream of all point sources that comply with the phosphorus rule. In the following section, we examine the differential effects by the two primary compliance options: technological upgrade and water pollution offset trading.

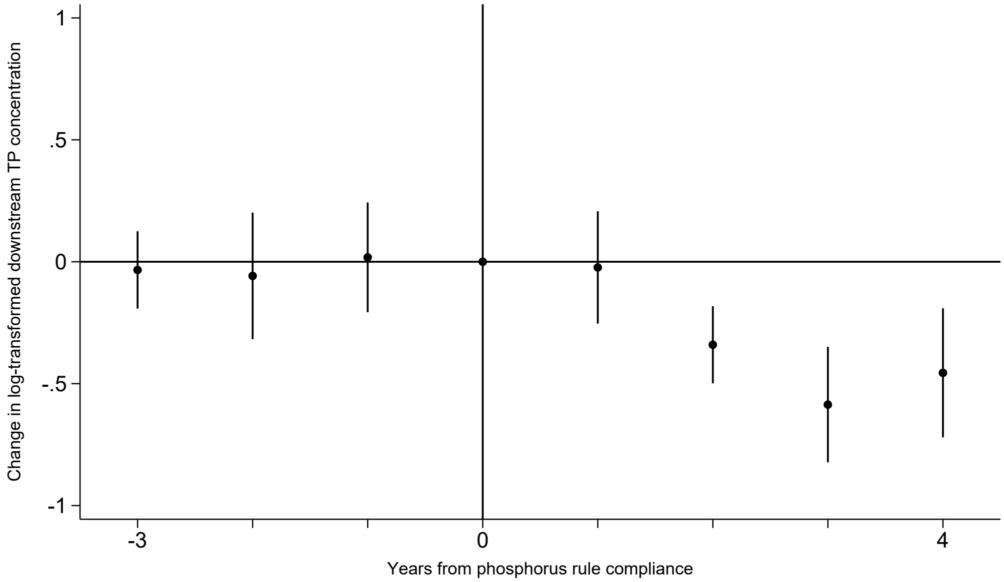
## 5.5 Event study

We next provide results for a standard event study of our primary specification. Figure 4 provides estimation results for this event study. The figure suggests that downstream TP concentrations trend similarly for facilities that comply with the phosphorus rule during our sample period and those that do not comply with the rule. The pre-treatment differences are practically zero and statistically insignificant, which provides suggestive evidence that our empirical framework satisfies the parallel trends assumption. After facilities comply with the phosphorus rule, downstream surface waterbody segments experience significant decreases in TP concentrations, consistent with the results provided in Table 2.

## 5.6 Randomized inference placebo test

In this subsection, we assess whether the coefficient estimates that we present above are due simply to chance by conducting a randomized inference placebo test (Athey and Imbens 2017). Randomized inference placebo tests, which are common in studies that use RCTs (e.g., Kerwin and Thornton 2021), randomize treatment throughout the analysis sample. Based on the randomized treatments, we then estimate our primary regression specification (equation (4)). If the treatment effects occur by chance, it is possible that the randomized treatments also have a statistically significant and meaningful impact on our outcome. Of course, there exists the possibility of a spurious relationship when estimating a single randomized regression, so the placebo tests repeat the randomized treatment estimation many times. We repeat the randomization process 2,500 times. We examine the distribution of the estimated placebo coefficients from the randomized inference placebo test. The test produces a p-value

Figure 4: Effect of phosphorus rule compliance on surface water quality, event study

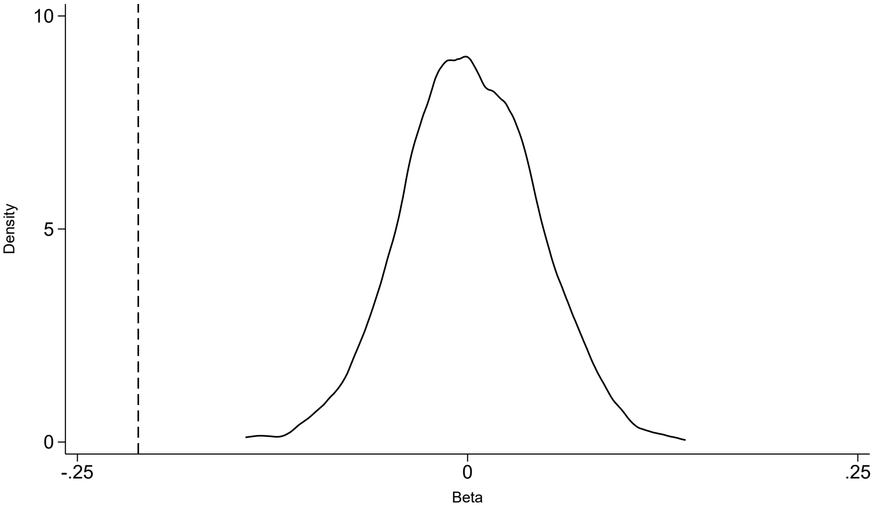


*Notes*: Results provided are the point estimates from the estimation of an event study of equation (4). Standard errors are clustered at the facility level and produce 95% confidence intervals, which are included. Dependent variable is log-transformed.

that reflects the portion of placebo coefficients whose absolute value exceeds the estimated coefficient from our primary TWFE specification ($\beta = -0.211$). The p-value represents a test of the null hypothesis that the treatment effect equals the placebo test coefficient.

We first conduct the randomized inference placebo test using Heß (2017)'s "ritest" Stata procedure. Second, we produce our own test using Monte Carlo simulations. We arrive at a p-value of 0.000 for both methods. Therefore, we strongly reject the null hypothesis that our results are spurious. Of the 2,500 placebo estimations, the mean treatment coefficient is 0.0014, with a mean t-statistic of 0.0495. Both values suggest that placebo treatments do not produce results like those of actual treatment. Figure 5 displays the distribution of the placebo coefficient estimates, with a vertical dashed line at the primary DDD coefficient estimate to compare the values. We again see strong evidence that the placebo estimations do not produce results like our primary treatment estimations, suggesting that the effect that we identify is a direct result of treatment and not simply by chance.

Figure 5: Randomized Inference Placebo Test Result Density



*Notes*: Figure presents the density of the coefficient estimates from a randomized inference placebo test of 2,500 runs. Coefficient estimate from the primary estimation of equation (**??**) is identified by the dashed line.

# 6 The effects of the Wisconsin's phosphorus rule on water quality: Heterogeneity by compliance option

Perhaps more importantly, we examine in this section the effect of point source phosphorus rule compliance option on downstream surface water quality. Wisconsin's phosphorus rule imposes sizable costs on point sources that discharge to surface waterbodies with poor ambient water quality. To meet stringent standards, point sources subject to the rule must upgrade to tertiary wastewater treatment technology. Or point sources can participate in water pollution offset trading to comply with the rule, at substantially lower cost. We also show above that the phosphorus rule itself improves downstream water quality. Here, we are interested in the extent to which compliance option impacts the surface water quality improvements of the rule. If both technological upgrade and offset trading improve downstream surface water quality, we are also interested in which is the cost-effective option.

In this section, we first describe the empirics of the heterogeneity analysis. Then we present results. Finally, we include a discussion on the cost-effectiveness of the compliance options.

## 6.1 Empirical framework

To examine differential impacts, we modify equation (4) to allow for heterogeneous effects of the two primary phosphorus rule compliance options. In this second specification, we interact a factor for water pollution offset trading facility, $Offset_i$, with the treatment dummy. This interaction represents the differential change in downstream TP concentrations between facilities that comply with the rule via treatment technology upgrade and those that comply through offset trading. We can then find the water quality effects of offset trading by examining the linear combination of the main DDD dummy and the interaction.

$Offset_i$ can take one of two forms. First is a simple dummy representing if the facility participates in offset trading. Second, we include a continuous interaction factor of the total credits that each facility receives. As shown in Table 1, there exists variation in the number of credits that each facilities receives in their trading agreements. Depending on the trading ratio, we expect agreements with more credits generated to result in less TP discharges from upstream, nonpoint sources, thus resulting in larger decreases in downstream TP concentrations. We primarily rely on estimates of this continuous measure, rather than the $Offset_i$ dummy.

We estimate the following regression specification for our heterogeneity analysis :

$$TP_{imt} = \beta Treat_{imt} + \gamma(Treat_{imt} * Offset_i) + \mathbf{Prec'}_{imt}\mu + \psi_i + \nu_{mt} + \epsilon_{imt} \tag{5}$$

where all notation follows equation (4). We again cluster standard errors at the facility level and focus on the results from Gardner (2022)'s imputation estimator.

The $Treat$ DDD coefficient estimated in equation (4) identifies the causal effect of phosphorus rule compliance on downstream surface waterbody TP concentrations. For the heterogeneity analysis, it is possible that unobserved managerial factors lead to each facility's compliance option. For example, managerial attitudes toward environmental protection may influence a facility's chosen compliance path. If this is the case, then estimation of equation 5 produces correlational, rather than causal, estimates of compliance option. However, we argue that each facility's compliance option is not a choice for each facility, but rather the result of exogenous factors.

The third- and fourth-year reports exclusively suggest the least-cost compliance option as the path forward for point sources needing to comply with the phosphorus rule. Like above, we calculate the yearly compliance costs of our analysis sample using ex ante estimates and a 5% discount rate over a 20-year time period. The median treatment technology upgrade cost is $123,021 per year. For offset trading, we normalize the cost of each to 100 credits because larger offset trading agreements are obviously more expensive. The median cost for trades of 100 credits is $11,522 per year. We assume that affected facility managers are cost-minimizers and do not choose to comply with the phosphorus rule via a substantially more expensive technological upgrade if water pollution offset trading is a viable alternative. For private facilities, this assumption follows from standard economic theory. For public facilities, such as wastewater treatment plants, managers must recover costs. This results in pass-through in the form of higher sewer utility bills, which are extremely unpopular politically (Meyer and Raff 2024).

Instead, outside factors make compliance option exogenous to each facility, rather than a facility-level decision. We list two key examples. First, consider the geography of each regulated point source. The requirement that trades be within the same HUC12 makes finding trading partners difficult. HUC12 sub-watersheds are roughly the size of a township. If regulated point sources are located in a HUC12 with little agriculture or opportunities to purchase offsets, then expensive technological upgrades are likely the only available compliance option. Even more restrictive is if the point source is at the top of the watershed. If there are few stream and river miles upstream of the point source within the same HUC12, then finding trading partners is virtually impossible, again forcing the facility toward technological upgrade. Second, the additionality requirements of Wisconsin's offset trading program (and most others) can prohibit the choice of compliance option. Especially in Wisconsin, TMDLs are an important impediment to offset trading. According to WDNR, over 40% of the state's surface waterbodies are covered by a TMDL (either approved or in development). For point sources within these TMDLs, agricultural and other BMPs are required to go toward the loading reduction targets of the TMDL before being part of any offset trade.[18] Many acres

---

[18] We had in-depth conversations with one point source manager in Wisconsin whose facility is required to comply with the phosphorus rule and is located in an area covered by a TMDL. Although the facility's watershed has ample agricultural activity, the manager said that his "hand were tied" regarding compliance

in Wisconsin are also covered by working lands programs such as the Conservation Reserve Program. For a relatively new program like WDNR's offset trading program, it can be difficult to find trading partners that are not already under contract with these other, more established, programs.

## 6.2 Estimation results by compliance option

Table 3 presents results for the estimation of equation (5). Panel A presents the regression coefficients from this estimation. Treat represents the effects of phosphorus rule compliance on downstream TP concentrations for facilities that comply with the rule via treatment technology upgrade. The coefficient of the interaction between treat and offset is the differential impacts on downstream TP concentrations between upgrading facilities and facilities that comply with the phosphorus rule via offset trading. Panel B presents the linear combinations of offset trades. Offset and credits represent the the effects of phosphorus rule compliance on downstream TP concentrations for trading facilities.

The results in Table 3 suggest heterogeneity in the water quality benefits of Wisconsin's phosphorus rule along the compliance option dimension. Although the differences between the two compliance options are not statistically significant at conventional levels for some specifications. We primarily discuss estimates from the imputation estimator of Gardner (2022) with a log-transformed outcome and the continuous offset measure (column 4). For point sources complying with Wisconsin's phosphorus rule via a tertiary treatment technology upgrade, compliance decreases average downstream TP concentrations by roughly 30%. And average downstream TP concentrations are 12.2% higher for point sources that comply via water pollution offset trading than for point sources that comply via treatment technology upgrade, for trades of 100 credits. Although offset trades decrease downstream TP concentrations by less than their technological counterpart, the linear combination of the main effect and the interaction effect (Treat + (treat*offset)) is statistically significant and negative. This linear combination suggests that downstream TP concentrations decrease by 20.8% after facilities comply with the phosphorus rule via an offset trade of 100 credits. The results presented in column 6 represent the empirical test of equation (3). The results are

---

via technological upgrade, because "all of the possible offset trades in the watershed go to the TMDL."

Table 3: Effect of phosphorus rule compliance on surface water quality by compliance option

| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| *Panel A. Regression coefficients* | | | | | | |
| Treat | -0.279*** | -0.269*** | -0.346*** | -0.363*** | -0.0244 | -0.0364** |
| | (0.0943) | (0.0847) | (0.0723) | (0.0690) | (0.0156) | (0.0167) |
| Treat*offset | 0.151 | 0.0920** | 0.113 | 0.130*** | -0.0150 | 0.0168 |
| | (0.102) | (0.0350) | (0.110) | (0.0279) | (0.0381) | (0.0110) |
| *Panel B. Linear combinations* | | | | | | |
| Offset | -0.128 | | -0.233** | | -0.0395 | |
| | (0.0806) | | (0.108) | | (0.0333) | |
| Credits (00s) | | -0.177*** | | -0.233*** | | -0.0195* |
| | | (0.0661) | | (0.0624) | | (0.0109) |
| Facility FE | X | X | X | X | X | X |
| Month#year FE | X | X | X | X | X | X |
| Precipitation controls | X | X | X | X | X | X |
| Gardner two-stage DD | | | X | X | X | X |
| Outcome scale | Log | Log | Log | Log | Level | Level |
| Observations | 2,826 | 2,826 | 2,826 | 2,826 | 2,826 | 2,826 |

*Notes:* *** p≤0.01, ** p≤0.05, * p≤0.1. Robust standard errors are clustered at the facility level and located in parentheses. The unit of observation is the stream reach-month. Panel A presents the regression coefficients from the estimation of equation (5). Treat represents the effects of phosphorus rule compliance on downstream TP concentrations for upgrading facilities. Treat*offset is the difference between upgrading facilities and offset facilities. Panel B presents the linear combinations of offset trades. Offset and credits represent the the effects of phosphorus rule compliance on downstream TP concentrations for trading facilities. Dependent variables in the first four columns are log-transformed, and are in levels in columns 5 and 6. Precipitation controls include total monthly precipitation and total monthly precipitation squared at the stream reach.

robust to measuring the outcome in levels. Because the marginal effect of credits is less than the main effect for technology upgrade, we conclude that trading ratios are set too low in WDNR's program, and that greater water quality benefits are possible by incorporating a greater margin of error for offset trades.

The results for the treatment technology upgrade compliance option are unsurprising. After installing tertiary treatment technology, point sources can certainly decrease TP discharges to meet their NPDES-permitted levels because the WQBEL is set at a level that can be met with such technology. Indeed, using EPA's Discharge Monitoring Report (DMR)

Pollutant Loading Tool,[19] we find that the average point source that complies with the phosphorus rule via treatment technology upgrade decreases their effluent discharges of TP by 533 pounds per year.

But the offset trading results are especially important. No prior study identifies a link between water pollution offset trading and surface water quality on a large spatial scale. By restricting our analysis to tight geographies (only five km upstream and downstream of affected point sources), we allow offset trades to cover a larger proportion of the study area than studies that examine nonpoint source pollution and other agricultural BMPs. We show that offset trades produce noticeable downstream water quality improvements, but the level of these improvements is less than that of treatment technology upgrades. We therefore conclude that trading ratios are likely too low in WDNR's program. However, the savings available to regulated point sources via the offset trading program are substantial. The following subsection examines the cost-effectiveness of the program at the micro level.

## 6.3 Cost-effectiveness

We have shown that compliance with Wisconsin's phosphorus rule through both treatment technology upgrade and water pollution offset trading leads to significant downstream water quality improvements. Although compliance via treatment technology upgrade produces larger TP reductions, offset trading is the much cheaper option. In this subsection, we examine the cost-effectiveness of compliance options. Specifically, we examine the cost per unit of surface water TP concentration reduction for each compliance option.

To perform this analysis, we use the same data and costs that we describe above, for each compliance option. In our final analysis sample, the median treatment technology cost to comply with the phosphorus rule is \$123,021 per year. And the median annual cost for trades, normalized to a size of 100 credits, is \$11,522. We calculate the cost per 0.01 mg/L reduction in downstream TP concentrations for each compliance option by dividing the yearly cost by the appropriate value from column 6 of Table 3, because we wish to perform the analysis in levels. For treatment technology upgrade, compliance with the phosphorus rule decreases downstream TP concentrations by 0.0364 mg/L. On a per unit basis, it costs

---

[19] https://echo.epa.gov/trends/loading-tool/get-data.

upgrading facilities \$33,797 per year to decrease downstream TP concentrations by 0.01 mg/L (123,021/(0.0364*100)). It costs facilities complying with the rule via offset trading \$5,909 per year for the same reductions (11,522/(0.0195*100)).

In the aggregate, we show above that water pollution offset trading saves regulated Wisconsin point sources roughly \$6.4 million annually. In the analysis of this section, we show that those savings are translated into real water quality improvements. Importantly, offset trades can produce similar reductions in downstream TP concentrations as treatment technology upgrades, at roughly 17% of the cost. Thus, offset trading represents the cost-effective option for phosphorus rule compliance and a way to improve surface water quality in Wisconsin.

# 7 Conclusion

In this paper, we examine the water quality benefits of Wisconsin's phosphorus rule, which in 2010 made TP discharge limits at NPDES-regulated point sources much more stringent. Overall, we find that Wisconsin's phosphorus rule decreases average downstream surface waterbody concentrations of TP by 26%, compared to the counterfactual of upstream surface waterbodies in the next HUC12 and point sources not yet regulated by the rule. We then show that water pollution offset trading delivers downstream water quality benefits, although the scale of the effects are not as large as those from the treatment technology upgrade.

EPA and others promote market-based approaches to water pollution control. However, because of the few water quality markets where trades occur, there is little evidence on the effectiveness of such programs. In our setting, the stringency and high costs of complying with the phosphorus rule via treatment technology upgrade led many point sources in Wisconsin to participate in water pollution offset trading to comply. We therefore use the comparatively many trades to quantify the differential benefits of offset trading compared to a technological option. We find that, in aggregate, water pollution offset trading presents compliance cost savings of roughly \$6.4 million per year compared to the treatment technology upgrade option. Perhaps more important, the per unit reduction costs for offsets are a fraction of those for the technological upgrade. To improve downstream TP concentrations by 0.01 mg/L, the annual costs for a treatment technology upgrade are $34,000 and for offset trades are $5,900. Water pollution offset trading therefore represents an efficacious and cost-effective compliance option.

Our results shed light on the efficacy of Wisconsin's phosphorus rule and are important when designing water pollution offset programs. The decrease in pollutant loadings and subsequent water quality improvements from technological compliance come at high cost to regulated point sources and customers (Meyer and Raff 2024). But the benefits of the cost-effective approach are less (although not zero), suggesting that the trading ratios in Wisconsin's program are below their optimal level. The programs must carefully consider ex ante trade amounts and ratios to ensure that the cost-effective compliance option produces benefits commensurate with those of the technological option. Finally, offset trading has additional benefits for the local area, in addition to water quality improvements. The costs of offset trades typically stay in the same locality as the regulated facility in the form of payments to farmers or other land users. Compared to technological costs that go to companies outside of the local area, offset trading costs can therefore help local, oftentimes rural, communities.

# References

**Athey, Susan and Guido W Imbens**, "The econometrics of randomized experiments," in "Handbook of economic field experiments," Vol. 1, Elsevier, 2017, pp. 73–140.

**Borusyak, Kirill, Xavier Jaravel, and Jann Spiess**, "Revisiting event study designs: Robust and efficient estimation," *arXiv preprint arXiv:2108.12419*, 2021.

**Breetz, Hanna L, Karen Fisher-Vanden, Hannah Jacobs, and Claire Schary**, "Trust and communication: mechanisms for increasing farmers' participation in water quality trading," *Land Economics*, 2005, *81* (2), 170–190.

**Callaway, Brantly and Pedro HC Sant'Anna**, "Difference-in-differences with multiple time periods," *Journal of econometrics*, 2021, *225* (2), 200–230.

**Chaisemartin, Clément De and Xavier d'Haultfoeuille**, "Two-way fixed effects estimators with heterogeneous treatment effects," *American economic review*, 2020, *110* (9), 2964–2996.

**Chakraborti, Lopamudra**, "Do plants' emissions respond to ambient environmental quality? Evidence from the clean water act," *Journal of Environmental Economics and Management*, 2016, *79*, 55–69.

**Cohen, Alex and David A Keiser**, "The effectiveness of incomplete and overlapping pollution regulation: Evidence from bans on phosphate in automatic dishwasher detergent," *Journal of Public Economics*, 2017, *150*, 53–74.

**DeBoe, Gwendolen and Kurt Stephenson**, "Transactions costs of expanding nutrient trading to agricultural working lands: A Virginia case study," *Ecological Economics*, 2016, *130*, 176–185.

**Doyle, Martin W, Lauren A Patterson, Yanyou Chen, Kurt E Schnier, and Andrew J Yates**, "Optimizing the scale of markets for water quality trading," *Water Resources Research*, 2014, *50* (9), 7231–7244.

**Earnhart, Dietrich and Donna Ramirez Harrington**, "Effect of audits on the extent of compliance with wastewater discharge limits," *Journal of Environmental Economics and Management*, 2014, *68* (2), 243–261.

**Fang, Feng, K William Easter, and Patrick L Brezonik**, "Point-nonpoint source water quality trading: A case study in the Minnesota River Basin," *JAWRA Journal of the American Water Resources Association*, 2005, *41* (3), 645–657.

**Fisher-Vanden, Karen and Sheila Olmstead**, "Moving pollution trading from air to water: potential, problems, and prognosis," *Journal of Economic Perspectives*, 2013, *27* (1), 147–172.

**Fleming, Patrick M, Erik Lichtenberg, and David A Newburn**, "Water quality trading in the presence of conservation subsidies," *Land Economics*, 2020, *96* (4), 552–572.

**Gardner, John**, "Two-stage differences in differences," *arXiv preprint arXiv:2207.05943*, 2022.

**Ghosh, Gaurav, Marc Ribaudo, and James Shortle**, "Baseline Requirements can hinder trades in water quality trading programs: Evidence from the Conestoga watershed," *Journal of environmental management*, 2011, *92* (8), 2076–2084.

**Goodman-Bacon, Andrew**, "Difference-in-differences with variation in treatment timing," *Journal of econometrics*, 2021, *225* (2), 254–277.

**Goulder, Lawrence H**, "Markets for pollution allowances: what are the (new) lessons?," *Journal of economic perspectives*, 2013, *27* (1), 87–102.

**Heß, Simon**, "Randomization inference with Stata: A guide and software," *The Stata Journal*, 2017, *17* (3), 630–651.

**Hsieh, Hsin-Chieh and Benjamin M Gramig**, "Estimating the impact of cover crop adoption on ambient nitrogen concentration in the upper Mississippi River drainage," *Applied Economic Perspectives and Policy*, 2024, *46* (2), 609–626.

**Jarvie, Michelle and Barry Solomon**, "Point-nonpoint effluent trading in watersheds: a review and critique," *Environmental impact assessment review*, 1998, *18* (2), 135–157.

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

**_ , Catherine L Kling, and Joseph S Shapiro**, "The low but uncertain measured benefits of US water quality policy," *Proceedings of the National Academy of Sciences*, 2019, *116* (12), 5262–5269.

**Kerwin, Jason T and Rebecca L Thornton**, "Making the grade: The sensitivity of education program effectiveness to input choices and outcome measures," *Review of Economics and Statistics*, 2021, *103* (2), 251–264.

**Mason, Charles F. and Andrew J. Plantinga**, "The additionality problem with offsets: Optimal contracts for carbon sequestration in forests," *Journal of Environmental Economics and Management*, 2013, *66* (1), 1–14.

**Meyer, Andrew and Zach Raff**, "The effect of water pollution regulation on prices: Evidence from Wisconsin's phosphorus rule and sewer utility bills," *Journal of Environmental Economics and Management*, 2024, *124*, 102914.

**_ , _ , and Sarah Porter**, "Remotely sensed imagery reveals animal feeding operations increase downstream dissolved reactive phosphorus," *JAWRA Journal of the American Water Resources Association*, 2024.

**Morgan, Cynthia and Ann Wolverton**, "Water quality trading in the United States: trading programs and one-time offset agreements," *Water policy*, 2008, *10* (1), 73–93.

**Newburn, David A and Richard T Woodward**, "An Ex Post Evaluation of Ohio's Great Miami Water Quality Trading Program 1," *JAWRA Journal of the American Water Resources Association*, 2012, *48* (1), 156–169.

**Olmstead, Sheila M**, "The economics of water quality," *Review of Environmental Economics and Policy*, 2010.

**Paudel, Jayash and Christine L Crago**, "Environmental externalities from agriculture: evidence from water quality in the united states," *American Journal of Agricultural Economics*, 2021, *103* (1), 185–210.

**Probst, Benedict, Malte Toetzke, Andreas Kontoleon, Laura Diaz Anadon, and Volker H. Hoffmann**, "Systematic review of the actual emissions reductions of carbon offset projects across all major sectors," *Working Paper*, 2023.

**Rabotyagov, Sergey S, Adriana M Valcu, and Catherine L Kling**, "Reversing Property Rights: Practice-Based Approaches for Controlling Agricultural Nonpoint-source Water Pollution When Emissions Aggregate Nonlinearly," *American Journal of Agricultural Economics*, 2014, *96* (2), 397–419.

**Raff, Zach**, "Delivering on the Promise of Water Quality Markets," *The Future of Waqter Markets*, 2022, p. 37.

**_ and Andrew Meyer**, "CAFOs and surface water quality: Evidence from Wisconsin," *American Journal of Agricultural Economics*, 2022, *104* (1), 161–189.

**Ribaudo, Marc and Jeffrey Savage**, "Controlling non-additional credits from nutrient management in water quality trading programs through eligibility baseline stringency," *Ecological Economics*, 2014, *105*, 233–239.

**Ribaudo, Marc O and Cynthia J Nickerson**, "Agriculture and water quality trading: Exploring the possibilities," *Journal of Soil and Water Conservation*, 2009, *64* (1), 1–7.

**Rodríguez, Fernando**, "On the use of exchange rates as trading rules in a bilateral system of transferable discharge permits," *Environmental and Resource Economics*, 2000, *15*, 379–395.

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

**Saby, Linnea, Jacob D Nelson, Lawrence E Band, and Jonathan L Goodall**, "Nonpoint Source Water Quality Trading outcomes: Landscape-scale patterns and integration with watershed management priorities," *Journal of Environmental Management*, 2021, *294*, 112914.

**Schmalensee, Richard and Robert N Stavins**, "The SO2 allowance trading system: The ironic history of a grand policy experiment," *Journal of Economic Perspectives*, 2013, *27* (1), 103–122.

**Selman, Mindy, Evan Branosky, and Cy Jones**, "Water quality trading programs: An international overview," 2009.

**Shortle, James, Markku Ollikainen, and Antti Iho**, "Water quality trading," in "Water Quality and Agriculture: Economics and Policy for Nonpoint Source Water Pollution," Springer, 2021, pp. 269–317.

**Skidmore, Marin, Tihitina Andarge, and Jeremy Foltz**, "Effectiveness of local regulations on nonpoint source pollution: Evidence from Wisconsin dairy farms," *American Journal of Agricultural Economics*, 2023, *105* (5), 1333–1364.

**Stephenson, Kurt and Gwen DeBoe**, "Transaction Costs for Nonpoint Source Water Quality Trading Credits: Implications for the Chesapeake Bay," 2016.

_ **and Leonard Shabman**, "Can water quality trading fix the agricultural nonpoint source problem?," *Annual Review of Resource Economics*, 2017, *9*, 95–116.

**Sun, Liyang and Sarah Abraham**, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of econometrics*, 2021, *225* (2), 175–199.

**Taylor, Charles A and Hannah Druckenmiller**, "Wetlands, flooding, and the clean water act," *American Economic Review*, 2022, *112* (4), 1334–1363.

**USEPA**, "Final water quality trading policy," Technical Report, Washington, DC 2003.

_ , "Updating the EPA's water quality trading policy to promote market-based mechanisms for improving water quality," Technical Report, Washington, DC 2019.

**Wardle, Arthur R.**, "Compliance Markets Without Romance: Lessons from the Renewable Fuel Standard," in Roberta Q. Herberg, Gavin Roberts, and Brianne Wolf, eds., *Social Coordination and Public Policy*, Lanham, MD: Lexington Books, 2023, pp. 171–192.

**Woodward, Richard T**, "Lessons about effluent trading from a single trade," *Applied Economic Perspectives and Policy*, 2003, *25* (1), 235–245.

_ **and Ronald A Kaiser**, "Market structures for US water quality trading," *Applied Economic Perspectives and Policy*, 2002, *24* (2), 366–383.