The Missing Benefits of Clean Water and the Role of Mismeasured Pollution

David A. Keiser

Working Paper 18-WP 581
June 2018

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

David A. Keiser is assistant professor of economics, Iowa State University, 468F Heady Hall, Ames, IA 50011.
E-mail: *dkeiser@iastate.edu*

Iowa State University does not discriminate on the basis of race, color, age, ethnicity, religion, national origin, pregnancy, sexual orientation, gender identity, genetic information, sex, marital status, disability, or status as a U.S. veteran. Inquiries can be directed to the Interim Assistant Director of Equal Opportunity and Compliance, 3280 Beardshear Hall, (515) 294-7612.

The Missing Benefits of Clean Water and the Role of Mismeasured Pollution*

David A. Keiser

Iowa State University and CARD

dkeiser@iastate.edu

Abstract

Although the U.S. spends billions of dollars a year controlling water pollution, there is little empirical evidence of comparable benefits. This study argues that measurement error in pollution data causes benefits to be underestimated. Using upstream concentrations as instrumental variables for local concentrations, the study finds substantial benefits from reducing nutrient pollution. Instrumental variable estimates of the effects of phosphorus on recreational use are an order of magnitude larger than conventional estimates. The study uses a long-term pollution dataset from Iowa to show that this difference is consistent with estimates of measurement error in several U.S. water pollution datasets.

JEL: C20; C26; Q26, Q50, Q51, Q53

Keywords: measurement error; non-market valuation; nutrient pollution; recreational demand; water pollution

*I thank a number of individuals for feedback on this work. This includes Otávio Bartalotti, Jude Bayham, Claire Brennecke, Peter Christensen, Alex Cohen, Eli Fenichel, Kenneth Gillingham, Namrata Kala, Brent Kreider, Harvey Lapan, Robert Mendelsohn, Nicholas Muller, James Saiers, Joseph Shapiro, and William Wheeler. I also thank Sheila Olmstead and three anonymous reviewers for constructive comments. I thank the United States Environmental Protection Agency (grant # 83456501-0) for funding. Todd Campbell and Yongjie Ji provided excellent research assistance. A previous working version of this paper was titled, "Estimating the Effects of Nutrients on Water-Based Recreational Uses in the United States: An Instrumental Variables Approach". All errors are my own.

1 Introduction

Spending on clean air and clean water programs in the United States has averaged tens of billions of dollars a year over much of the last four decades (USEPA, 1997, 1999, 2011; Van Houtven et al., 2000). While a number of studies find that air pollution control programs are worth these costs, a recurrent finding is that water pollution control programs are unlikely to pass a cost-benefit test (Freeman, 1982; Carson and Mitchell, 1993; Lyon and Farrow, 1995; Bingham et al., 2000; Hansen, 2007; Olmstead, 2010). Across a number of different settings, measured benefits are often small relative to their costs. This paper provides evidence that the noisy nature of water pollution data has caused some of these benefits to be underestimated.

Unlike the dominating role that human health plays in economic analyses of air pollution, studies of water pollution control have largely focused on the importance of water-based recreational use. The majority of measured benefits rely upon stated preference surveys. In these surveys, individuals state what they are willing to pay to move up a step function of water quality levels described as boatable, fishable, and swimmable (Carson and Mitchell, 1993). As an alternative, revealed preference methodologies infer the relationship between water quality and recreation using data on actual visitations and water quality measurements. When revealed preference approaches have been used, prior analyses have implicitly assumed that water pollution is exogenous and perfectly measured. This paper relaxes these assumptions and provides new estimates of the effects of water pollution on recreation. In particular, I develop a reduced-form model to estimate the effects of nutrient pollution (phosphorus) on water-based recreational use.

Nutrient pollution provides an excellent setting to examine the effects of water quality on recreational use. These pollutants result in very visible, undesirable changes to the aquatic system that include decreased water clarity, odor, loss of biodiversity, and harmful algal blooms that threaten drinking water and human health (USEPA, 2000; Smith and Schindler, 2009; Garnache et al., 2016). The issue is widespread, and the USEPA considers nutrient pollution to have, "...the potential to become one of the costliest, most difficult environmental problems we face in the 21st century" (USEPA, 2009).

This paper's primary research design uses upstream concentrations of phosphorus as an instrumental variable for local county-level pollution concentrations. The validity of this approach rests on the assumption that pollution from far away distances affects recreation, but only through its effect on local pollution concentrations.[1] To employ this strategy, this paper matches individual responses from a national survey on recreation with local and upstream water quality measurements from two national datasets. These data provide broad cross-sectional coverage of recreation and water quality across much of the U.S. In an ordinary least squares (OLS)-type model, the effects of

[1] Bayer et al. (2009) utilize a similar strategy to identify the effects of air quality on home prices and wages. Ebenstein (2012) uses distance from a headwater stream to identify the effect of water pollution on health in China.

phosphorus on water-based recreation are generally negative, small and at times only marginally significant. When local concentrations are instrumented with upstream pollution levels, the effect of phosphorus on water-based recreation is negative, large in magnitude, and highly significant. Given the cross-sectional nature of the data, I interpret these estimates as representing recreational responses to long-run water quality conditions.[2]

Naturally, these results beg the question as to why the OLS and instrumental variables estimates diverge. I pay particular attention to the role of measurement error in the water quality data as a potential source of bias. The concentrations that I use from the national datasets may be subject to measurement error for a number of reasons. One dataset, a nationally representative survey by the USEPA, provides only one measurement of phosphorus per lake for a particular year (2007). Due to the variable nature of water quality, these single measurements of phosphorus are likely noisy measures of the true, long-run concentrations in these lakes. The other dataset, provided by the USGS, uses a water quality model to produce lake-level estimates of concentrations for a given year (2002) assuming long-run stream flow conditions. Thus, measurement error may be introduced if this model poorly predicts concentrations, or if concentrations for this "base year" are a noisy measure of long-run conditions. In addition, measurement error in both datasets could be introduced when averaging concentrations across lakes to obtain an estimate of county-level water quality conditions. This would be the case if water quality varies substantially within counties and these datasets are not comprehensive.[3]

To explore the potential role of measurement error, I exploit a separate, unique water quality dataset from the state of Iowa. These monitoring data are collected several times a year from a large number of lakes within the state and cover the period from 2001 to 2012. I assume these data provide a close approximation to the "true" long-run concentration of nutrients at lakes within Iowa. I assume the Iowa Lakes data provide a superior measurement of water quality since these data are collected more frequently and over a longer time period than either the USEPA or USGS data. For example, in contrast to the single USEPA measurement per lake, the Iowa Lakes data provide approximately 28 measurements per lake over the span of a decade. They are also more comprehensive. For the set of counties that are in both the Iowa Lakes and USEPA datasets, the Iowa Lakes data includes 2.8 lakes per county compared to USEPA's 1.5 lakes per county. Likewise, for the same set of counties that are in both the Iowa Lakes and USGS datasets, the Iowa

[2] An interesting extension would be to compare these findings to a study that examines short-run responses to changes in water quality using a panel data setting.

[3] In practice, additional measurement error may arise if there is a mismatch between local county-level concentrations and water quality that drives behavior (e.g., water quality at other sites or in other counties that may influence behavior). This potential source of measurement error does not contribute to this paper's estimates of measurement error. This paper's estimates are based on the difference between "true" and noisy local county-level concentrations. Nonetheless, the instrumental variables strategy, if valid, would correct for these additional sources of measurement error.

Lakes data includes 1.7 lakes per county compared to USGS's 1.3 lakes per county.

I quantify measurement error in the "noisy" national datasets by taking the difference between them and the "true" Iowa concentrations. These calculations provide estimates of measurement error for a subset of roughly 60 counties. The measurement error estimates are then used for three main purposes: 1) to estimate attenuation factors, 2) to test classical measurement error assumptions, and 3) to provide assurances that the instrumental variables are uncorrelated with the measurement error. The first step gauges the importance of measurement error in the data. The second and third steps provide indirect tests of the validity of the instruments. The estimated attenuation factors suggest that measurement error would reduce, or attenuate, true parameter estimates by a factor of ten. This difference is similar to ratios that I find between several OLS and instrumental variables specifications. Importantly, the instrumental variables are also uncorrelated with the measurement error estimates. These findings do not imply that other potential sources of bias should be ignored. Rather, they highlight the important role of measurement error in pollution data.

Finally, to understand how these results may inform prior cost-benefit studies of water pollution control programs, I perform a replication exercise that examines how the upstream instrumental variable research design affects estimates of the benefits of the Conservation Reserve Program (CRP). As a large federal program to control sediment erosion and nutrient pollution, the U.S. has spent roughly $58B in total on the CRP over the last twenty-seven years.[4] However, current estimates place the benefits of the program at approximately 70 to 85 percent of its costs (Hansen, 2007). These estimates are based, in part, on a prior revealed preference study that estimates the effects of sediment erosion on recreational use (Feather and Hellerstein, 1997). With conventional OLS and count data models, I find nearly identical point estimates of the effects of sediment erosion on recreation. However, when local pollution is instrumented with upstream concentrations, parameter estimates are negative, significant, and much larger in magnitude. Using the instrumental variables results, a back of the envelope calculation suggests that the benefits of the CRP equal or exceed costs by factor of 2 to 1. This finding is the first empirical evidence to suggest a favorable cost-benefit relationship for this program.

This paper builds on several strands of literature. First, this paper extends what is known about the role of measurement error in pollution data and the appropriateness to which an instrumental variable can be used to correct for it. Prior studies importantly raise measurement error as a potential source of bias. However, the type of measurement error (i.e., classical versus non-classical) and the degree to which it biases estimates of the effects of pollution on economic outcomes has not been thoroughly explored.[5] For example, Schlenker and Walker (2016) study how carbon monox-

[4] Author calculation based on annual rental payments from 1987-2014, adjusted to 2014 dollars (USDA, 1986-2014).
[5] See Atkinson and Crocker (1992) for a similar observation two decades earlier.

ide exposure impacts respiratory and heart-related hospitalization rates. In their study, Schlenker and Walker (2016) note that their, "instrumental variables setting allows (them) to simultaneously address issues pertaining to both avoidance behavior and classical forms of measurement error, each of which lead to significant downward bias in conventional dose-response estimates." Similarly, Chay and Greenstone (2003) note that their instrumental variable estimates of the effects of air pollution on infant health, "may suffer from less attenuation bias than (their) fixed effects estimates." More recently, Aizer et al. (2018) argue that measurement error in blood tests lead one to significantly underestimate the impact of lead pollution on educational outcomes. In this setting, Aizer et al. (2018) implement an instrumental variables strategy to correct for this source of bias. However, a valid instrumental variables approach requires that the instrument must be uncorrelated with the measurement error. This assumption, while often held, is rarely tested.[6] In this paper, I use the Iowa Lakes data to test this critical assumption. Furthermore, outside of economics, this work contributes to recent studies that have called attention to the role of measurement error in water quality data and the implications for trends analysis, calibration of hydrological models, and policy responses such as the implementation of Total Maximum Daily Loads (Harmel et al., 2006; McMillan et al., 2012).

Second, there have been a number of recent studies that address endogenous site quality in recreational demand modeling (Murdock, 2006; Timmins and Murdock, 2007). Much of this attention has focused on the potential of omitted variable bias related to the travel cost parameter. However, a number of prior studies suggest these concerns apply more broadly (von Haefen and Phaneuf, 2008; Abidoye et al., 2012). In a recent review of recreational demand models, Moeltner and von Haefen (2011) state that the concern of omitted variables is, "likely germane to most recreation demand applications and ... somewhat surprisingly has received direct attention only in recent years." This paper builds on the existing literature by relaxing the assumption that water pollution is exogenous.[7] It may be the case that other variables such as site characteristics are correlated with water quality and also influence recreation. If these variables are unobserved, their absence would bias estimates of the effects of water quality on recreation in a standard OLS-type regression. This concern has been expressed in studies of the effects of water pollution on other outcomes of interest, including human health (Cutler and Miller, 2005; Kremer et al., 2011; Currie et al., 2013) and housing (Mendelsohn et al., 1992; Leggett and Bockstael, 2000; Keiser and Shapiro, 2017). Although this paper focuses on the role of measurement error, this does not preclude the possibility of these other sources of bias. It also does not preclude the instrument, if valid, from correcting these potential sources of bias.

[6] In other economic fields, particularly labor economics, understanding measurement error and its implications for the use of instrumental variables is a critical concern (Bound et al., 2001).

[7] One exception is a panel data study by Hausman et al. (1995) that examines the effects of the Exxon Valdez oil spill.

Third, this paper builds on a long history of recreational demand modeling (Phaneuf and Smith, 2005). This paper's approach most closely resembles what Phaneuf and Smith (2005) describe as some of the literature's earliest "reduced form" activity participation models. The recreational data used in this paper provide a measure of participation in a recreational activity and the associated number of days per year. However, these data do not provide a site destination for the recreationist. One advantage of using these data is that I am able to implement a research design that takes advantage of considerable variation in water quality across space to identify the effect of water quality on recreation. As discussed in Section 2, clearly identifying how recreation responds to changes in water quality can be extremely useful for a benefit-transfer approach to estimate spatially-explicit marginal damages of water pollution. However, it is important to note that a key weakness of this approach is that site choice is not observed. In many modern structural approaches, observing site choice allows the analyst to recover a welfare measure of both marginal and non-marginal changes in site quality (Moeltner and von Haefen, 2011). This paper provides a potential starting point (i.e., instrumental variable) for future explorations with site choice models.

Fourth, these results extend recent studies that examine the relationship between nutrient pollution and economic uses. Although conceptual models outline the general impacts of nutrients on water uses, there is limited information on the precise link between concentrations and use (Dodds et al., 2009; Egan et al., 2009; Keeler et al., 2012). This paper provides a response function that links water-based recreational use to nutrient pollution. These estimates can be used in future studies to estimate spatially-explicit marginal and total benefits of pollution control.

Finally, these results may also hint at why some similar revealed preference studies have struggled to establish a strong relationship between recreation and water quality. For example, in one of the first cost-benefit analyses of the Clean Water Act (CWA), Freeman (1982) considered a number of early reduced-form recreational studies to assess benefits.[8] Consistent with the OLS findings in this paper, some of these early studies note weak relationships between water quality and recreation. For example, Freeman (1982) notes that in one of the studies he reviews, "In a second-stage equation to estimate the frequency of participation of those who participated one or more times, it was found that water quality variables (were) not statistically significant" (Peskin and Harrington, 1981). Roughly a decade after Freeman (1982) and several years before Feather and Hellerstein

[8] To estimate recreational benefits of the CWA, Freeman (1982) reviewed four studies commissioned by the National Commission on Water Quality (1976) and two analyses performed at Resources for the Future. Several of the studies that Freeman (1982) describes use quantitative measures of water quality (Bell and Canterbery, 1975; National Planning Association, 1975; Peskin and Harrington, 1981). However, it is important to note that Freeman (1982) places a significant amount of weight on Russell and Vaughan (1982) to estimate recreational benefits of the CWA. Their study uses a survey of fish and game officials to gauge the extent of waters that are suitable for fishing within a state. This is not a direct measure of water pollution. It is not clear that Freeman's cost-benefit analysis of the CWA would substantially change, but measurement error may have contributed to weaker relationships between water quality and recreation in these earlier studies, potentially biasing benefits downward.

(1997), Ribaudo and Piper (1991) note similar struggles to establish a reduced-form relationship between water quality and recreation in an effort to measure the benefits of the CRP. Ribaudo and Piper (1991) state, "A major surprise of the results was that water quality appeared to play no role in the number of fishing trips made."

The rest of the paper proceeds as follows. Section 2 motivates the empirical approach with a theoretical framework used to measure spatially-explicit marginal damages from water pollution. Section 3 details the primary datasets. In Section 4, I outline an instrumental variables estimator for count data and examine the relationship between nutrients and water-based recreational use. Section 5 explores the role of measurement error. Section 6 describes implications of these results for the CRP. Section 7 concludes.

2 Theoretical Motivation

This section motivates the empirical approach with a simple model of recreational demand. This section also shows how the empirical estimates provided by this paper can be used in future research to estimate the marginal and total damages of water pollution.

Let utility for a representative individual be a function of total recreational days (R) in a year and a composite good (B) with a price of 1. Each trip has an implicit price (p). The individual has an annual income (Y). The utility gained from recreational days is influenced by site characteristics. Some of these characteristics are water pollution concentrations $(\mathbf{C})$, whereas others $(\mathbf{W})$ are not. An exogenous vector of socioeconomic characteristics $(\mathbf{S})$ influence the utility received from R and B. The individual's maximization problem is:

$$Max_{R,B} U(R,B;\mathbf{C},\mathbf{W},\mathbf{S})$$

$$s.t.\ p \cdot R + B \leq Y$$

Maximizing utility subject to the budget constraint yields the Marshallian demand function:

$$R = R(p,Y,\mathbf{C},\mathbf{W},\mathbf{S}) \tag{1}$$

If this function is known, the change in consumer surplus $(\triangle CS)$ from a change in concentration of a particular pollutant (C) is captured by the area between two demand functions at concentration levels C^0 and C^1:

$$\triangle CS = \int R\left(p,Y,\mathbf{C^1},\mathbf{W},\mathbf{S}\right) dp - \int R\left(p,Y,\mathbf{C^0},\mathbf{W},\mathbf{S}\right) dp \tag{2}$$

When these demand functions are not known for a particular study area, a "benefit transfer" approach approximates equation (2) by multiplying an estimate of the average individual consumer

surplus of a recreational day $(\bar{CS})$ by an estimate of the marginal effect of concentrations on recreational demand $\partial R(\cdot)/\partial C$. Marginal damages for a particular pollutant i from an emissions source j are then equal to the sum of marginal changes that occur to n individuals in each location l as a result of a marginal change in emissions E:

$$MD_{ij} = \sum_{l=1}^{L} \frac{\partial C_{il}}{\partial E_{ij}} \frac{\partial R(\cdot)}{\partial C_{il}} \cdot \bar{CS} \cdot n_l \tag{3}$$

Similar spatially-explicit marginal damages of air pollution have played an important role in recent discussions of optimal air pollution policies (Muller and Mendelsohn, 2007, 2009; Muller et al., 2011). However, obtaining such estimates of marginal damages for water pollution has been challenging (Keiser and Muller, 2017). In order to estimate marginal damages, as is seen from equation (3), one must estimate the marginal effect of pollution concentrations on recreational demand, $\partial R(\cdot)/\partial C$. This paper provides this critical element. Other key components have been estimated by others (College of Forestry, Oregon State University, 1981-2010; Smith et al., 1997; Bennear et al., 2005).

3 Data

This paper brings together several datasets on water quality, recreational use, and other geospatial data. Table 1 presents summary statistics for recreational use and water quality. Figure 1, Panels A and B provide visual representations of these data. Each source is described in more detail in the following sub-sections.

3.1 Water Pollution

I use two separate sources of water pollution data to construct two estimating samples. Each sample provides broad coverage of the U.S. The first source of data, from the USEPA, consists of measured water pollution concentrations in lakes across the U.S. The second source, from the USGS, provides imputed concentrations. I use both types of data since the structure of the measurement error may be different for measured and imputed concentrations, which may affect the validity of the instrumental variables strategy. These two distinct types of pollution data have been used in prior recreational demand studies (see, for example, Feather and Hellerstein (1997); Egan et al. (2009)).[9]

The USEPA data include phosphorus concentrations collected at 1,028 lakes throughout the

[9] The use of predicted concentrations has featured heavily in prior cost-benefit studies of national water quality policies. Measured concentrations have featured more prominently in localized studies of the effects of water quality on economic outcomes. Current work by Keiser and Shapiro (2017) utilizes a much more dense network of stream and lake monitors to study the consequences of the U.S. Clean Water Act.

U.S. as part of USEPA's National Lakes Assessment (USEPA, 2007). These collection efforts were designed to provide a representative sample of water quality in U.S. lakes. Single sample measurements were taken during the summer of 2007.[10] The second dataset is constructed from lake concentrations obtained from the USGS SPARROW Decision Support System (USGS, 2002).[11] The SPARROW model uses measured concentrations of water pollutants at select locations across the U.S. to examine how emissions, land use, climate, and landscape characteristics explain pollution concentrations across a study area using a regression-based approach (Smith et al., 1997). Calibrated equations are then used to predict pollutant loads and concentrations to and throughout the U.S. for a given "baseline" year, assuming long-term stream flow conditions. I use phosphorus concentration estimates for a baseline year of 2002 from six large river basin models that comprise most of the U.S.

For each of these data sources, I construct a measure of local phosphorus by taking the average concentration of these pollutants at lakes within each county. For counties without a lake, I construct a measure of local water quality by averaging the concentration of adjacent counties. I construct these measures at the county-level to match the level of geographic detail provided in the recreational use dataset. In constructing these averages, I drop outliers from the top and bottom one percentile to account for any extremely inaccurate measurements or predictions.[12] For this analysis, I focus on lake water quality to define local water quality. This focus on lakes is primarily driven by the availability of long-term monitoring data at lakes within the state of Iowa. These data allow me to examine the role of measurement error in this study. In Section 4, I briefly discuss regressions that use county-level river and stream concentrations from the USGS SPARROW data rather than lakes. As with lakes, I find larger effects with the instrumental variable specifications.

Table 1 shows average phosphorus concentrations by census division for each estimating sample. Panel A of Figure 1 displays average phosphorus concentrations by county for the USEPA estimating sample. The USEPA sample represents 1,846 counties in the U.S. The geographic extent of the USGS sample is slightly larger with 2,295 counties. Both estimating samples provide broad coverage of the U.S., except for the Southwest and most of California.[13] As shown in Panel A, phosphorus concentrations tend to be lower on the coast, particularly for New England. Concentrations are higher in the Midwest, likely due to greater agricultural activity and fertilizer use. For example, using the USGS SPARROW regional models, Robertson and Saad (2011) estimate

[10] For 95 lakes, a second sample is obtained on a return visit. I use measurements from the first visit.

[11] These data originate from the following studies: Hoos and McMahon (2009); Brown et al. (2011); Garcia et al. (2011); Moore et al. (2011); Rebich et al. (2011); Robertson and Saad (2011); Wise and Johnson (2011).

[12] Mean concentrations are two to three times higher in the USGS SPARROW sample. Given these large differences, I further restrict the maximum and minimum concentrations in the USGS sample to the maximum and minimum from the USEPA sample.

[13] These areas of the U.S. are excluded because corresponding USGS models have not been developed. Since the USGS concentrations are used as instrumental variables, these areas are also excluded from the USEPA sample.

that manure and farm fertilizer contribute 50 percent of the total phosphorus load in the upper Midwest, while Moore et al. (2011) find similar sources contribute approximately 30 percent of the total phosphorus load in the Northeast and Mid-Atlantic. In the Northeast and Mid-Atlantic, forested land is a much larger contributor (approximately 43 percent). The variation in concentrations across the U.S. motivates the use of state fixed-effects in the empirical analysis to control for other unobserved factors that may be correlated with these concentrations and also affect recreational use. I also control for land cover and metropolitan and micropolitan statistical areas to control for characteristics of sources of phosphorus that vary within a state and may also influence recreation. The Appendix displays a map of phosphorus concentrations for the USGS data (Appendix Figure 2). The USGS data follow a similar pattern. These maps suggest that while some differences exist, the two estimating samples appear to provide similar representations of nutrient concentrations across the US.

It is worth noting that in addition to phosphorus, there can be other determinants of local water quality. For nutrient pollution in particular, nitrogen also plays an important role. In freshwater settings, phosphorus is often the limiting nutrient for plant growth and thus is the main cause of adverse changes to the aquatic system. Nitrogen plays a similar role in saltwater systems (Smith and Schindler, 2009). Furthermore, sediment pollution, also from agriculture, is an important determinant of water clarity and may be correlated with these measures of phosphorus. Phosphorus may attach to sediment that runs off into streams and lakes, and thus could reflect contemporaneous measures of erosion. Longer-term, legacy effects are another possibility in which phosphorus concentrations may reflect phosphorus that has accumulated in lakes over time and is being released in the water column (Garnache et al., 2016). Deciding what measure of water quality to use in recreational demand models remains an important topic, with some authors opting to estimate parameters for particular pollutants (e.g., Bockstael et al., 1987); Egan et al., 2009), others defining water quality using an index (e.g, Phaneuf et al., 1998), and others using some other measure of aquatic characteristics that are themselves impacted by water quality such as fish catch rates (e.g., Bockstael et al., 1989). Furthermore, many non-pollution factors such as temperature and flow can impact how changes in pollutants translate into changes in the characteristics of aquatic systems (e.g, Smith and Schindler, 2009); Keeler et al., 2012). This study controls for a number of these parameters with controls for climate, predominant land uses, and general geographic differences across states. With this in mind, the analysis provides an average effect of phosphorus pollution for the study sample conditional on controls (e.g., an average effect across the US). Furthermore, it is important to keep in mind that to the extent that phosphorus pollution is proxying for other measures of pollution (e.g., nitrogen, sediment, toxins), the study's findings serve as a proxy for water quality if variation in those pollutants are correlated with the upstream instrumental variable

concentrations.[14]

3.2 Recreational Use

Recreational use data are taken from the National Survey on Recreation and the Environment (NSRE) for the years 1999 to 2009 (USDA, 2009).[15] The NSRE is generated by a randomized telephone survey of a cross-section of individual households in the U.S. The survey reports individual-level responses for both the extensive margin (participation) and intensive margin (the number of days) for a variety of outdoor recreational activities. The survey also captures a range of demographic information of the respondents and includes the county of residence. I use survey responses for activities that include swimming in natural bodies of water, coldwater fishing, warmwater fishing, boating (with a motor), and visitations to non-coastal watersides for other recreational purposes. I construct a water-based recreational day category that represents the sum of days from all of these categories.

A few key aspects of the data are worth mentioning. For approximately 35 percent of the sample, respondents only answered questions regarding participation in water-based recreation activities (e.g., yes or no to the activities listed above). These respondents did not provide information on the number of days they participated. It is not clear which respondents were only asked this extensive margin question. Thus, I remove a random sample of non-participants such that the final estimating sample contains the same balance of participants and non-participants as the original sample.[16] In the Appendix, I show that this data selection step has little effect on the results. In addition, the survey reports post-stratification weights that consider age, gender, and race in a similar manner as the U.S. Census survey design. These weights are further refined to reflect educational and urban/rural classifications. I control for individual demographic and residential features in the analyses. In the Appendix, I explore the effect of using sampling weights as a robustness check and find that these weights do not significantly impact the results.

Table 1 shows summary statistics of recreational use by census division for the two estimat-

[14] To account for other measures of water quality, I also examined the effects of nitrogen on recreation. When nitrogen is included in the analysis, the phosphorus results in Section 4 remain largely unchanged. However, the upstream instrumental variables for nitrogen were too weak to be informative and thus are not included. Using a more limited geographic sample that includes fish toxins yielded similar phosphorus results.

[15] The NSRE is a partnership between the USDA Forest Service Southern Research Station, the National Oceanic and Atmospheric Administration, the University of Georgia, the University of Tennessee and other federal, state or private sponsors.

[16] For example, my starting sample for the US EPA analysis contains 48,632 observations with 67 percent of respondents answering that they had participated in one of the above water-based recreational activities. Of these respondents, 11,369 answered that they had participated in a water-based recreational activity, but did not answer questions about the total number of participation days. To construct the estimating sample, I randomly remove 5,624 respondents that did not participate in water-based recreation. This data step leaves an estimating sample where 67 percent of the sample participates in one of the above activities. This sample is slightly restricted in the subsequent analyses by limiting the sample to counties where I am able to track upstream to construct an instrumental variable.

ing samples. Panel B of Figure 2 provides a map of average county-level participation for the USEPA estimating sample. Not too surprisingly, participation in water-based recreational activities is higher on the coasts. These higher rates of participation appear weakly correlated with lower pollution concentrations. However, the coasts may also have other factors that are attractive to recreationists such as easier access to waterbodies for recreating. Similar recreational patterns are seen in the USGS estimating sample (Appendix Figure 1).

3.3 Upstream Concentrations

In addition to lake concentrations, the USGS SPARROW data provide concentrations of rivers and streams. I use these river and stream concentrations to construct upstream concentrations as instrumental variables for local county-level concentrations. Once they enter surface waters, nutrients such as phosphorus are known to travel great distances. For example, nutrient pollution issues in the Gulf of Mexico and the Chesapeake Bay are caused by nutrient loads as far away as Minnesota and New York respectively (USEPA, 2010; Robertson and Saad, 2013). This aspect of nutrient pollution suggests that upstream concentrations may induce variation in local water quality. In Section 4, I perform standard formal tests of the strength of the instruments to confirm this assumption.

To track upstream, I use a recently developed program from Keiser and Shapiro (2017). This program uses a stream routing network from the National Hydrography Dataset Plus, Version 2.1 (NHD) to track upstream and downstream through surface waters in the US. The NHD provides an electronic map of surface waters of the US that include rivers, streams, and lakes. In total, this network includes 3.5 million river miles that cover nearly the entire U.S. (see Keiser and Shapiro (2017) for more details). To create the instrumental variables, I use this program to first link a county with any other county that is upstream from it as connected by this stream network.[17] I then construct my instrumental variables for a particular downstream county by averaging the concentrations of phosphorus in the upstream counties. In calculating these averages, I exclude any upstream counties that are adjacent to the local downstream county to help satisfy the necessary exclusion restriction (see Section 4.1 for further details). My preferred specification uses non-adjacent counties that are up to 200 miles upstream from a local county. To create instruments at varying distances, I also calculate the average concentration of counties that are 30 to 200 and 60 to 200 miles upstream of the local county. These distances are calculated from the center of each local downstream county to the center of upstream counties. I use the county to county distances rather than the stream distances to account for driving distances between counties.

[17] More precisely, I start with all streams of at least Strahler Order 3 (all but very small streams) within a county. I then search upstream through the stream network to identify upstream counties. This method links over 85 percent of counties in the US with an upstream county.

3.4 Other Geospatial Data

I construct a number of additional county-level characteristics as controls. These are average precipitation and temperature data obtained from the PRISM Climate Group (PRISM Climate Group, 1981-2010), shoreline delineation provided by the National Oceanic and Atmospheric Administration (NOAA, 2000), average county elevation data from the National Elevation Dataset (USGS, 2015), the predominant land cover classification provided by the 2001 National Land Cover Database (USGS, 2001), and 2009 metropolitan and micropolitan statistical area designations provided by the U.S. Census Bureau (U.S. Census Bureau, 2009).

3.5 Iowa Lakes Data

To estimate measurement error in the USEPA and USGS datasets, I utilize water quality monitoring data from the state of Iowa. These data provide measured concentrations of phosphorus collected on average three times per year from 2001 to 2007 and 2009 to 2012 for 136 lakes in the state (IA DNR and ISU Limnology Laboratory, 2001-2012). These data represent a unique long-term monitoring program with broad coverage of the state.

4 The Effects of Nutrients on Water-based Recreational Use

4.1 Econometrics

To estimate equation (1), I use a conventional count data model as well as an instrumental variables estimator for count data. I define the annual number of recreational days for individual i in county k and state s (R_{iks}) as a function of county-level regressors such as average pollution concentrations ($\mathbf{X_{ks}}$) and individual demographic variables ($\mathbf{D_{iks}}$). In addition, β_0 is an unknown parameter and $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ are unknown parameter vectors.[18] Individual unobserved factors ($\boldsymbol{\Theta}$) that determine the demand for recreational days enter the regression framework in a similar manner as the other regressors with $\eta = exp(\boldsymbol{\Theta}) > 0$. Mullahy (1997) shows that this setup implies the following traditional exponential regression framework:

$$R_{iks} = exp(\beta_0 + \mathbf{X_{ks}}\boldsymbol{\beta_1} + \mathbf{D_{iks}}\boldsymbol{\beta_2}) \cdot \eta_{iks} + \varepsilon_{iks}$$

Count data models such as the Poisson or Negative Binomial assume a certain distribution for η which satisfies the key assumption that $E(\varepsilon|\mathbf{X},\mathbf{D},\eta) = 0$. However, if the expectation of η conditional on $\mathbf{X}$ and $\mathbf{D}$ does not equal the unconditional expectation of η, then the parameter esti-

[18] It is important to reiterate that a number of factors other than phosphorus can determine water quality and influence recreation. To that end, I also control for county-level climate variables, metropolitan and micropolitan statistical area designations, coastal and Great Lakes shorelines, average elevation, and predominant land cover type. To the extent that phosphorus proxies for other pollutants, the coefficient on pollution would reflect not just phosphorus, but these other determinants of water quality.

mates will be inconsistent (Mullahy, 1997). Issues of measurement error and omitted variables are examples where this assumption of traditional count models may not hold. In these circumstances, Mullahy (1997) shows that with a suitable vector of instruments $\mathbf{Z_{iks}}$, the following transformed moment condition can be used to recover consistent parameter estimates through GMM estimation:

$$E\left(\mathbf{Z_{iks}}\left(R_{iks}/exp(\beta_0+\mathbf{X_{ks}}\beta_\mathbf{1}+\mathbf{D_{iks}}\beta_\mathbf{2})-1\right)\right)=0$$

The key assumptions for the consistency of the instrumental variables estimator are $E\left(\varepsilon|\mathbf{X},\mathbf{D},\mathbf{Z}\right)=0$ and $E\left(\eta|Z\right)=\alpha$, where α is a constant (Mullahy, 1997). The instruments that I use in this paper are upstream county-level concentrations of phosphorus. Thus, these assumptions require that the variation in local water quality induced by these upstream concentrations should be mean independent of other unobserved county-level or individual factors that influence recreation. For example, one may be concerned if the upstream concentrations are highly correlated with unobserved preferences for recreation or local amenities that encourage recreation. I perform a number of placebo checks in Section 4.3 to explore these types of threats to identification. These tests examine how phosphorus pollution affects another 25 different types of non-water based recreational activities. As a second test of these identifying assumptions, in Section 5, I examine the correlation between the instruments and my estimate of measurement error using the Iowa Lakes data as a validation dataset. Finally, one additional concern to note would be if the upstream counties also serve as substitute sites for local recreation. In this instance, the instrumental variables estimates would be inconsistent since they would be highly correlated with unobserved quality at substitute sites. This is one reason why I exclude surrounding counties when constructing the instrumental variables. Data collected from the National Household Travel Survey (NHTS) provides some additional reassurances that recreation is concentrated locally. In the 2009 NHTS, approximately 94 percent of all social and recreational trips were to destinations less than 30 miles from home.[19] This suggests that these upstream concentrations (or upstream sites) do not factor into the vast majority of recreational trips. I further explore this issue by examining the robustness of the results when I only use counties that are 30 to 200 miles and then 60 to 200 miles upstream to construct the instrumental variables.

[19] This number rises to 98 percent for destinations less than 50 miles and 99 percent for destinations less than 75 miles from home. Author calculations using 2009 person trips from U.S. Department of Transportation, Federal Highway Administration (2009). Calculations exclude unreported trips.

4.2 Results

As a first step, Tables 2 and 3 report parameter estimates for phosphorus using the USEPA and USGS estimating samples respectively. Column 1 reports conventional cross-sectional results, and columns 2 through 6 report instrumental variables results. For each table, Panel A shows results for the count data model, and Panel B shows results for the closely related semi-log model.[20] The dependent variable for each regression is the individual number of water-based recreational trips in a year. All regressions also include county-level concentrations of phosphorus, individual-level controls for demographic characteristics (race, gender, age, income, education), county-level climate variables (30-year climate normals for average temperature and precipitation and temperature squared)[21], and additional county-level characteristics that may influence recreational use (metropolitan and micropolitan statistical area designations, coastal and Great Lakes shorelines, designations for counties with a lake, average elevation, and predominant land cover type). In addition, each regression except column 3 includes state fixed-effects to control for unobserved factors at the state-level that may influence recreation and be correlated with nutrient levels. Column 3 includes census division fixed effects.

For both estimating samples, the results from conventional models generally show a negative relationship between phosphorus and recreational use (column 1 in Tables 2 and 3).[22] The results are generally of a similar magnitude and are highly significant except for the count model in the USGS sample, which is significant at the 10 percent level.

Column 2 reports results using upstream concentrations of phosphorus as an instrumental variable for downstream concentrations. Across both estimating samples and specifications, the coefficient estimate of the effect of phosphorus on recreational use is highly significant and similar in magnitude. Strikingly, the magnitude of the coefficient on phosphorus is substantially larger at approximately 9 to 15 times the magnitude of the coefficient from the conventional models. Columns 4 to 5 in Tables 2 and 3 show how varying the distances used to define the instrumental variables (e.g., 30 to 200 and 60 to 200 miles upstream) affects the results. Generally, the magnitude of the effect of phosphorus on recreational use remains very similar across these different specifications. However, the standard errors grow larger with instruments from further distances upstream. The Kleibergen-Paap rk Wald F Statistic (Kleibergen and Paap, 2006) shows that this result is likely due to the instrument becoming weaker the further upstream I go to define the instrumental variables concentrations. For example, in the USEPA sample (Table 2), the Kleibergen-Paap rk Wald

[20] I use the negative binomial model to account for over-dispersion in the data. The semi-log model includes 1+ the number of days. Mullahy (1997) shows how a semi-log instrumental variables model should provide similar estimates to the count data instrumental variables model.

[21] I exclude a squared term for precipitation since this causes convergence issues with the GMM count model. However, results are similar when the quadratic term is included in the semi-log models.

[22] Standard errors are clustered at the county level since variation in water quality occurs at the county level.

F Statistic shrinks from 13.452 in column 2 to 6.361 and 2.800 when the instrumental variables are defined at 30 to 200 and 60 to 200 miles upstream respectively. The F Statistics for column 2 is slightly less than the Stock-Yogo (Stock and Yogo, 2005) weak identification critical value for 10% maximal instrumental variables size.[23] Thus, in column 3, I report estimates using census division fixed effects rather than state fixed effects. In these regressions, the instrument is slightly stronger and exceeds the most stringent Stock-Yogo weak identification critical value. The issue of weak instruments appears to be a slightly larger concern when examining counties that are further upstream, where the F Statistic shrinks substantially as the distance upstream used to construct the instrumental variables increases. These results suggest a trade-off between more plausibly exogenous instruments that are further away (e.g., 60 to 200 miles) versus instruments that are stronger, but closer and less plausibly exogenous.

Table 4 reinforces these findings. This table reports the reduced form and first stage results for the semi-log specifications for the USEPA and USGS estimating samples. The closest upstream concentrations (non-adjacent, but >0 to 200 miles) are significant in the first stage for both the USEPA and USGS samples. For the USEPA sample, the instrumental variable concentrations remain strongly significant at 30 to 200 miles, but are only marginally significant at 60 to 200 miles. For the USGS sample, the instrumental variable concentrations are only marginally significant at 30 to 200 miles and 60 to 200 miles. This finding suggests that specifications using further upstream concentrations in USGS sample may be subject to weak instruments.[24]

In addition to assessing the strength of the instruments, Tables 2 and 3 provide some statistics to support the exclusion restriction. The null hypothesis for the overidentification test is not rejected in either estimating sample (see Hansen's J statistic in column 6 of Tables 2 and 3). If one believes that further upstream instruments are more plausibly exogenous, the failure to reject the null hypothesis lends some credibility to the validity of these instruments (Greenstone and Gallagher, 2008; Bazzi and Clemens, 2013). However, in this setting, one concern with these tests is that the different instruments are based on similar sources of variation (upstream water quality). Thus, I cautiously interpret these overidentification tests as providing support for the exclusion restriction and further subject these results to additional robustness checks in Section 4.3.

[23] See Baum et al. (2007) and Bazzi and Clemens (2013) for helpful explanations of these tests.

[24] If measurement error is a key concern, it would seem natural to use one water pollution dataset as an instrument for the other and vice versa. However, I find that these data serve as very weak instruments for each other when using a full suite of controls. The estimated coefficients are negative, large in magnitude, and insignificant. However, in models with less demanding geographic and demographic controls, the coefficient estimates of phosphorus are similar in magnitudes to those using the upstream instrumental variables. A regression of county-level concentrations from the USEPA data on the USGS SPARROW data yields an R-squared value of approximately 0.08. Milstead et al. (2013) similarly compare the USEPA data to the USGS SPARROW data for Northeastern lakes. Their lake-level comparison of the same data yields an R-squared of approximately 0.35. These differences are likely due to averaging at the county-level versus lake level and comparing USEPA and USGS SPARROW data for the entire U.S.

In addition, I test the robustness of these results to the inclusion of outlier concentrations (the top and bottom 1 percentile of concentrations). Results using the USEPA sample change very little with outlier concentrations included. The magnitude of the effect of phosphorus on recreational use actually slightly increases. Outliers are more important for the USGS estimating sample. The inclusion of outliers weakens the instruments and decreases the magnitude of the coefficients. However, it is worth noting that the full USGS sample includes phosphorus concentrations up to 70.68 mg/l, which is much larger than any measured concentration in the USEPA data (4.68 mg/l).

Furthermore, I use these results to estimate average marginal effects. For conventional models, average marginal effects range from -0.299 to -0.562 days per 0.1 mg/l change in phosphorus in the USGS and USEPA samples respectively. For the corresponding instrumental variables estimates, the average marginal effects range from -3.990 to -5.315 days per 0.1 mg/l change in phosphorus.[25] As shown in equation 3 in Section 2, these average marginal effects serve as a critical component for measuring spatially-explicit marginal damages. These results suggest that average marginal damages from the instrumental variables approach are 9 to 13 times larger than those implied by conventional models.

Last, I examine the choice of using lakes versus rivers and streams to measure local water quality. In the conventional models, I still find significant and negative effects of phosphorus on recreational use with a similar magnitude as shown in Tables 2 and 3. For example, for the semi-log models, the coefficient on phosphorus is -0.189 in both the USEPA and USGS samples. For the instrumental variable results, I find effects of phosphorus that are larger in magnitude and highly significant as before. However, these effects are smaller in magnitude than those using lake water quality. For the semi-log models, they range in magnitude from -0.512 to -0.863. These results may suggest that water quality in lakes drives a stronger behavioral response than water quality in rivers. This finding is consistent with other recent recreational demand studies (Egan et al., 2009; Ji et al., 2016).

4.3 Placebo Tests

A concern that one might have regarding these results is that phosphorus pollution could be correlated with other unobserved factors that affect recreational use. For example, one could imagine a scenario in which phosphorus concentrations happen to be lower in areas in which individuals have stronger preferences for all types of recreational activities. If the upstream instrumental variables are also correlated with these preferences, the results in Section 4.2 could reflect an associational (not causal) impact of phosphorus pollution on water-based recreational use. As an indirect test of this concern, I perform a number of placebo tests that take advantage of the extensive recreational

[25] For the USEPA and USGS samples, average marginal effects are reported using regressions from Columns 1 and 2 of Tables 2 and 3.

use database. These tests examine the effects of phosphorus on participation in the individual activities that make up the water-based recreational category in Section 4.2 (swimming, cold and warm water fishing, boating, and visiting a waterbody) as well as 25 other activities.[26]

These results provide two main insights.[27] The first is that swimming and coldwater fishing appear to be the two water-based recreational uses that are most sensitive to changes in phosphorus concentrations. This result is perhaps not too surprising given that water quality is often ranked by its ability to support boatable, fishable, and swimmable water, where swimmable water corresponds to the best water quality and boatable is the worst. The second insight is that the placebo tests provide strong support for the results found in Section 4.2. Across 25 different activities, only two activities display a significant relationship with phosphorus concentrations. One of these activities is swimming in an outdoor pool. In this case, the impact of phosphorus is actually positive, which lends further support to the findings in Section 4.2. If nutrient pollution is worse in a county, swimming in an outdoor pool may serve as a potential substitute activity. The second activity is driving off-road. It is not clear why this activity would be impacted by phosphorus pollution. It could be the case that areas that allow off-road driving experience more erosion, which could increase phosphorus pollution. In this circumstance, the significant finding would reflect a case of reverse causality. However, it may also be the case that this relationship is a spurious correlation that happens by chance. For example, at a 5 percent significance level, one in twenty coefficients will be statistically significant even when no relationship exists between phosphorus and the recreational activity.[28]

5 Measurement Error

This section explores measurement error as one potential mechanism behind the large differences in the OLS and instrumental variable estimates from Section 4.2. Section 5.1 provides a simple statistical framework for exploring this issue. Section 5.2 uses data from the state of Iowa to obtain an estimate of measurement error in each dataset to explore the potential for this source of bias.

5.1 Measurement Error Framework

I use the Iowa lakes water quality data to gauge the extent of measurement error present in the USEPA and USGS pollution datasets. As mentioned in Section 3.5, the Iowa Lakes data are from

[26] These activities include swimming in an outdoor pool, bicycling, hiking, snow and ice activities, outdoor team sports, horseback riding, picnicking, visiting a nature center, visiting a prehistoric site, visiting a historic site, walking for pleasure, visiting a wilderness area, gathering mushrooms, hunting, mountain biking, backpacking, developed camping, primitive camping, viewing birds, viewing fish, viewing wildlife, viewing flowers, viewing scenery, driving for pleasure, and driving off-road.

[27] The results and a more detailed discussion are provided in the Appendix.

[28] I thank an anonymous reviewer for making this point.

a unique long-term monitoring dataset of 136 lakes within the state of Iowa. Since these data are collected frequently and over a long time horizon, I treat phosphorus concentrations from the Iowa data as reflective of true, long-run average concentrations. The difference between the Iowa Lakes and the USEPA and USGS concentrations is defined as an estimate of the measurement error.

Following a standard additive measurement error framework, let C_k^m equal the mismeasured long-run average concentration for a particular county k (i.e., average phosphorus concentration for Story County, IA given by the USEPA or USGS data), C_k^* the true value (i.e., average phosphorus concentration for Story County, IA given by the Iowa lakes long-term monitoring data), and u_k the difference (i.e., the measurement error)[29]:

$$C_k^m = C_k^* + u_k$$

A classical measurement error framework assumes the measurement error is independent of the true concentration, which implies[30]:

$$E(u|C^*) = 0 \tag{4}$$

A well-known result from the classical measurement error framework is that measurement error attenuates the true coefficient from an OLS regression towards zero.[31] In a linear regression with a single regressor, the estimated coefficient $\hat{\beta}$ converges in probability to $\lambda \cdot \beta$, where β is the true coefficient. In this case, the classical measurement error (CME) attenuation factor λ_{CME} is a function of the variance of the true concentration $\sigma^2_{C^*}$ and the variance of the measurement error σ^2_u:

$$\lambda_{CME} = \frac{\sigma^2_{C^*}}{\sigma^2_{C^*} + \sigma^2_u} \tag{5}$$

For non-classical measurement error (NCME) with a single regressor, the error is at least partially correlated with the true concentration. The attenuation factor is now dependent upon the covariance between the measurement error and the true concentration:

[29] In addition to the additivity assumption, I assume that the measurement error is non-differential (e.g, the measurement error is uncorrelated with R_{iks} and ε_{iks}). Furthermore, this analysis implicitly assumes zero measurement error in R_{iks}. Generally, measurement error in the dependent variable will not bias coefficient estimates as long as this measurement error is uncorrelated with the independent variable of interest. This discussion benefits from Bound and Krueger (1991), Bound et al. (1994), Black et al. (2000), and Hyslop and Imbens (2001).

[30] With covariates, the CME framework assumes that the measurement error is also independent of the covariates, which implies $E(u|C^*, \mathbf{X}, \mathbf{D}) = 0$.

[31] The attenuation factor that holds in a single regressor model does not necessarily hold with multiple regressors. In general, it depends on the regression model, the structure of the measurement error, and whether the error is correlated with additional covariates (see Carroll et al. (2006)).

$$\lambda_{NCME} = \frac{\sigma^2_{C^*} + \sigma_{u,C^*}}{\sigma^2_{C^*} + \sigma^2_u + 2\sigma_{u,C^*}} \tag{6}$$

Using the subset of Iowa counties that have long-term monitoring data and are present in the national estimating samples, I calculate λ_{CME}, and λ_{NCME}. These attenuation factors provide useful information on the degree to which measurement error may affect conventional OLS results. I also regress estimates of the measurement error (u_k) on local concentration estimates $\left(C^*_k, C^m_k\right)$ to test the mean independence condition implied by equation (4). Finally, I regress measurement error estimates on instrumental variables concentrations to lend support to the validity of the instrumental variables.

5.2 Measurement Error Results

As a first step, I plot frequency distributions of the measurement errors for the USEPA and USGS samples. The unit of observation is the county-level measurement error, u_k, as defined in Section 5.1. These distributions are shown in Figure 2, Panels A and B. Panel A shows that the main mass of the measurement error for the USEPA distribution is centered around zero. However, the right tail of the distribution extends much farther out than the left resulting in a mean measurement error of 0.060 mg/l. By contrast, the distribution of the measurement error of the USGS sample is centered around 0.2 mg/l with a mean of 0.227 mg/l.[32]

As discussed in Section 5.1, a key factor for understanding the implication of these errors is to understand how they are correlated with the noisy and true concentrations, as well as the relative magnitudes of the variances and covariances given in equations (5) and (6). To better understand these relationships, I first calculate equations (5) and (6) with the Iowa Lakes and corresponding USEPA and USGS data. Table 5, columns 1 and 2 report these values. The estimated attenuation factors are very small and range from 0.01 to 0.11. These factors imply that measurement error in the USEPA and USGS datasets could result in OLS coefficient estimates that are 10 to 100 times smaller than their true values. The size of this attenuation bias is tremendous. If measurement error is the only source of bias, OLS estimates would significantly underestimate the benefits of phosphorus control.

For the USEPA sample, the underlying sample variances used to calculate equations (5) and (6) are $\sigma^2_{C^*} = 0.0033$ and $\sigma^2_u = 0.0293$. The covariance between the measurement error and Iowa Lakes concentrations is -0.0031. For the USGS sample, these sample variances are $\sigma^2_{C^*} = 0.0035$

[32] For the USEPA data, Shapiro and Wilk (1965) and Shapiro and Francia (1972) tests reject the normality of the distribution. For the USGS data, these tests fail to reject the normality of the distribution. Although the shape and location of these distributions are interesting, this analysis focuses on how these errors are correlated with the measured and true concentrations and the resulting attenuation factors. This follows similar analyses such as Bound et al. (1994).

and $\sigma_u^2 = 0.0284$. The covariance between the measurement error and Iowa Lakes concentrations is -0.0025. A few points are worth noting regarding these estimates. First, for both samples, the signal (e.g., variance of the true concentrations) is small relative to the noise (e.g., variance of the measurement error). These large differences lead to the small attenuation factors. These measurement errors are large, but others have noted significant magnitudes of measurement error in nutrient concentrations. For example, in a comprehensive review of the literature, Harmel et al. (2006) note that errors in total phosphorus measurements may range up to 250 percent of their true values. Second, the magnitude of the measurement error variances in the USEPA and USGS samples are very similar. Although the measurement error is generated by two different processes (e.g., infrequent sampling with the USEPA data versus estimation error from the USGS model), the similar measurement error variances lead to similar classical attenuation factors. Furthermore, it is interesting to note that the covariance between the Iowa Lakes concentrations and measurement error estimates is negative in both samples. In labor applications, this negative relationship often reduces attenuation bias and thus is described as "mean reverting" (Bound and Krueger, 1991). However, in this analysis, the large variance in the measurement error relative to the variance in the signal leads to a larger attenuation bias. If this negative covariance was large enough, it could flip the sign of the effect of water quality on recreation in an OLS type regression. This covariance is slightly smaller in the USGS sample versus the USEPA sample. Thus, the USGS sample has slightly larger non-classical attenuation factors.

Next, to further explore the classical versus non-classical structure of the measurement error, I regress measurement error on the mismeasured concentrations (USEPA and USGS estimates of phosphorus) as well as the "true" concentrations (Iowa Lakes). This step essentially tests equation (4).[33] Table 6 reports these results. Panel A displays the USEPA results and Panel B presents USGS results. Column 1 shows results from a regression that includes phosphorus as the sole regressor. Column 2 adds additional controls. For both the USEPA and USGS samples, there is a strong association between the measurement error and the mismeasured concentrations. The coefficients on phosphorus are highly significant. The large R-squared values in column 1 show that a lot of the variation in the measurement error is explained by the variation in mismeasured concentrations. These results are consistent with a classical measurement error structure. Columns 3 and 4 repeat these analyses with the Iowa Lakes data. The significance of the coefficient estimates in columns 3 and 4 suggests that the measurement error is at least partially non-classical. However, the relationship between the measurement error and the phosphorus concentration is much weaker than in the case of the USEPA and USGS data.

[33] The main focus of this regression is to analyze whether a strong relationship exists between the measurement error and the mismeasured concentrations. A non-zero mean would shift the intercept term in an OLS regression, but does not affect whether the error is classical or non-classical in nature.

Figure 3, Panels A through D provide further visual evidence of these relationships. Panels A and B display scatter plots of measurement error observations versus their corresponding "noisy" measure in the USEPA (Panel A) and USGS data (Panel B). Panels C and D display similar scatter plots of measurement error observations versus their corresponding "true" concentrations in the Iowa Lakes data for the USEPA sample (Panel C) and USGS sample (Panel D). These figures reinforce the findings in Table 6. There appears to be a very strong, positive correlation between the measurement error and the noisy measures in both the USEPA and USGS data. These relationships appear much stronger than the corresponding relationships between the measurement error observations and the Iowa Lakes data. These results suggest that the measurement error is much more classical in nature. Others have made this assumption regarding pollution data in other settings to support an instrumental variables strategy (Chay and Greenstone, 2003; Schlenker and Walker, 2016; Aizer et al., 2018). In this paper, the Iowa Lakes data offers an opportunity to take this one step further and provide evidence for this assumption.

Next, I examine how the instrumental variables approach potentially corrects for measurement error. If one is willing to assume that the instrumental variables estimate of the phosphorus coefficient (from Section 4.2) is equal to its true value, then the ratio of the OLS estimate to the instrumental variables estimate should closely resemble these attenuation factors. This follows from the discussion in Section 5.1 since $\lambda = \frac{\hat{\beta}}{\beta}$. Columns 3 through 6 of Table 5 report these OLS to instrumental variables ratios for both the count data models (columns 3 and 5) and the semi-log models (columns 4 and 6). I also report these ratios with only phosphorus as a regressor (columns 3 and 4) and with the full suite of controls (columns 5 and 6). I report these ratios in Table 5 so that they can be closely compared to the attenuation factors shown in columns 1 and 2. For both the USEPA and USGS data, when controls are considered, these ratios are remarkably similar to the CME attenuation factors. The CME attenuation factor is approximately 0.10 and the coefficient ratios range from 0.07 to 0.11. In the case of a single regressor, the ratio estimates for the USEPA data are higher, but still similar (0.173 to 0.258). For the USGS data, these ratios are even higher (0.500 to 0.579) but still suggest large differences. The finding that the ratios closely resemble the CME attenuation factors is consistent with the classical nature of measurement error shown in Figure 3 and Table 6.

I also use these data as an indirect check on the validity of the instrumental variables strategy. For consistent estimates, the measurement error must be independent of the instrumental variables, which implies $E(u|Z) = 0$. Table 7 reports results of linear regressions of measurement error estimates on the instrumental variables concentrations. Panel A shows results with USEPA data and Panel B shows results with the USGS data. There is no significant association between the upstream concentration and the measurement error for either dataset, with or without controls. Panels E and F from Figure 3 provide visual evidence to support these regressions. These figures

display scatter plots of measurement error estimates versus their corresponding non adjacent, upstream instrumental variable concentrations at >0 to 200 miles. Panel E displays these results for the USEPA data. Panel F show these results for the USGS data. These figures show no association between the measurement error estimate and the corresponding upstream instrumental variable concentrations. These results lend further support to the necessary exclusion restriction.

In summary, these results suggest that 1) measurement error is a concern in the national water quality datasets, 2) the structure of this measurement error appears more classical than non-classical in nature, and 3) the instrumental variables are largely uncorrelated with the measurement error, which lends support for their use. Taken together, these results suggest that the instrumental variables strategy corrects for significant attenuation bias that plagues the OLS research design.

6 Implications for the Conservation Reserve Program

The prior sections shed light on large differences between estimates of the effects of water quality on recreational demand in OLS versus instrumental variable research designs. These findings suggest that some prior studies may have undervalued the benefits of water pollution control. To assess the implications of these findings, I examine a prior cost-benefit study of the CRP. In a similar effort 20 years ago, Feather and Hellerstein (1997) examined how water quality affects water-based recreational use. Similar to the current study, the authors discuss that their original intent was to include certain measures of ambient water quality (total nitrates, total phosphorus, and dissolved oxygen). However, the limited and noisy nature of the water quality data at the time led the authors to use a prediction of soil erosion as a proxy for ambient water quality. The predicted nature of these data most closely resembles the USGS estimating sample used in the prior sections of this paper. The authors combine recreational use data from the 1992 version of the NSRE with estimates of soil erosion. Thus, the similar nature of the water quality and recreational data of Feather and Hellerstein (1997) make this study very suitable for comparison. Furthermore, the results of Feather and Hellerstein (1997) have subsequently been used in cost-benefit studies of the CRP (Hansen, 2007).

Feather and Hellerstein (1997) utilize a random utility model to first capture an individual's site choice. They then use a double-hurdle Poisson model to estimate the effects of erosion on the number of lake and river recreational days.[34] To reexamine their findings, I estimate my empirical model using NSRE data from 1999 to 2009 and county-level average estimates of soil erosion (NRCS, 1997). Results are shown in Appendix Table 5. For the intensive margin component of their double-hurdle Poisson model, Feather and Hellerstein's (1997) parameter estimates of the effects of water erosion (tons/acre) on recreational days range from -0.0309 to -0.0462. Using

[34] The 1992 version of the NSRE included site choices for a limited number of respondents. The authors control for travel costs, percent of land in forest cover, percent of land privately owned, lake area, and family income.

minimal controls for a similar version of the count and semi-log models from Section 4, I estimate parameter estimates of -0.0392 to -0.0542. These results are nearly identical to those found by Feather and Hellerstein. Since soil erosion acts as a potential proxy for nutrient pollution, I utilize the upstream concentrations of phosphorus as instruments for local erosion estimates.[35] For the instrumental variable specifications, the effect of soil erosion on recreational use is negative and highly significant. Parameter estimates range from -0.176 (no controls) to -0.545 (a full suite of controls). These instrumental variables parameter estimates are 4 to 14 times larger than the corresponding conventional parameter estimates. The Kleibergen-Paap rk Wald F Statistics show that these instruments are very strong.

Turning to Hansen's (2007) cost-benefit analysis, his estimates place water-based recreational use at approximately 11 percent of the total benefits of the CRP. Although this fraction seems small, it is the largest benefit of any single soil conservation category.[36] Considering all uses, Hansen (2007) places total benefits of the CRP at approximately 70 to 85 percent of costs. Adjusting water-based recreational benefits by 4 to 14 times the prior estimate would, at a minimum, bring total benefits in line with costs. At the upper end of estimates, the benefit to cost ratio is nearly 2 to 1.[37] Other estimates of water quality benefits in Hansen (2007) are based on similar sources of data. It is possible that total benefits would be even higher.

7 Conclusions

This paper develops an instrumental variables approach to estimate the effects of nutrient pollution on water-based recreational use. Upstream pollution concentrations serve as instrumental variables for local pollution concentrations. I report three main findings. First, the instrumental variables approach reveals a significant, negative effect of phosphorus pollution on water-based recreational use. This estimate is an order of magnitude larger than estimates from conventional OLS-type models. Second, I find strong evidence that this discrepancy is due, at least in part, to measurement error in two national-scale water pollution datasets. I use a long-term monitoring pollution dataset from the state of Iowa to validate this claim. Finally, these results have important policy implications. Nearly all studies of water pollution control programs suggest that the total costs of these programs outweigh their benefits. I find that this result does not necessarily hold for the CRP. This paper provides new estimates that suggest the benefits of this program may very

[35] Again, it is worth mentioning that these estimates could potentially reflect other pollutants (i.e., nutrients) correlated with current or prior erosion levels.

[36] Other categories include reservoir services, navigation, municipal water treatment, dust-cleaning, irrigated agriculture, irrigation ditches and canals, soil productivity, marine fisheries, freshwater fisheries, marine recreational fishing, municipal and industrial water use, steam electric power plants, flood damages, and road drainage ditches.

[37] This calculation uses annual cost of $1.7B for the CRP reported by Hansen (2007). Adjusting water-based recreational use benefits upwards by a factor of 4 to 14 increases annual benefits from $140.99M to $563.96M to $1.973B. This adjustment increases total annual benefits of the CRP from $1.3B to $1.9-$3.3B.

well outweigh its costs. However, many other water pollution control efforts, such as the Clean Water Act, target other pollutants and other sources of pollution. Future work is needed to examine whether similar cost-benefit relationships hold for other programs.[38]

References

Abidoye, Babatunde O., Joseph A. Herriges, and Justin L. Tobias. 2012. Controlling for observed and unobserved site characteristics in rum models of recreation demand. *American Journal of Agricultural Economics* 94 (5): 1070–1093.

Aizer, Anna, Janet Currie, Peter Simon, and Patrick Vivier. 2018. Do low levels of blood lead reduce children's future test scores? *American Economic Journal: Applied Economics* 10 (1): 307–341.

Atkinson, Scott E. and Thomas D. Crocker. 1992. Econometric health production functions: Relative bias from omitted variables and measurement error. *Journal of Environmental Economics and Management* 22 (1): 12–24.

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *The Stata Journal* 7 (4): 465–506.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management* 58(1): 1–14.

Bazzi, Samuel and Michael A. Clemens. 2013. Blunt instruments: Avoiding common pitfalls in indentifying the causes of economic growth. *American Economic Journal: Macroeconomics* 5 (2): 152–186.

Bell, Frederick W. and E. Ray Canterbery. 1975. An assessment of the economic benefits which will accrue to commercial and recreational fisheries from incremental improvements in the quality of coastal waters.Technical report, Tallahassee, FL.

Bennear, Lori S., Robert N. Stavins, and Alexander F. Wagner. 2005. Using revealed preferences to infer environmental benefits: Evidence from recreational fishing licenses. *Journal of Regulatory Economics* 28 (2): 157–179.

[38] In a current paper, Keiser and Shapiro (2017) find that the costs of the CWA municipal grants program are larger than the corresponding local changes in home values. The authors discuss a number of potential reasons for this finding. It is important to note that the CRP focuses largely on unregulated agricultural pollutants and is more targeted in nature than the CWA.

Bingham, Tayler H., Timothy R. Bondelid, Brooks M. Depro, Ruth C. Figueroa, A. Brett Hauber, Suzanne J. Unger, and George L. Van Houtven. 2000. A benefits assessment of water pollution control programs since 1972: Part 1, the benefits of point source controls for conventional pollutants in rivers and streams: Final report.Technical report, Prepared by Research Triangle Institute for the Office of Water and Office of Policy, Economics, and Innovation, U.S. Environmental Protection Agency, Washington DC.

Black, Dan A., Mark C. Berger, and Frank A. Scott. 2000. Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association* 95 (451): 739–748.

Bockstael, Nancy E., Michael W. Hanemann, and Catherine L. Kling. 1987. Estimating the value of water quality improvements in a recreation demand framework. *Water Resources Research* 23 (5): 951–960.

Bockstael, Nancy E., Kenneth E. Mcconnell, and Ivar E. Strand. 1989. Measuring the benefits of improvements in water quality: The chesapeake bay. *Marine Resource Economics* 6 (1): 1–18.

Bound, John, Charles Brown, Greg J. Duncan, and Willard L. Rodgers. 1994. Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics* 9 (1): 1–24.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. Measurement error in survey data. In *Handbook of Econometrics*, ed. Heckman, James J. and Edward Leamer.Amsterdam: Elsevier.

Bound, John and Alan B. Krueger. 1991. The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics* 9 (1): 1–24.

Brown, Juliane B., Lori A. Sprague, and Jean A. Dupree. 2011. Nutrient sources and transport in the missouri river basin, with emphasis on the effects of irrigation and reservoirs. *Journal of the American Water Resources Association* 47 (5): 1034–1060.

Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. 2006. *Measurement error in nonlinear models.*Boca Raton:Taylor and Francis Group, LLC.

Carson, Richard T. and Robert C. Mitchell. 1993. The value of clean water: The public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resources Research* 29 (7): 2445–2454.

Chay, Kenneth Y. and Michael Greenstone. 2003. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *Quarterly Journal of Economics* 118 (3): 1121–1167.

College of Forestry, Oregon State University. 1981-2010. Recreational Use Values Database. Available online at http://recvaluation.forestry.oregonstate.edu/ (accessed September 10, 2013).

Currie, Janet, Joshua Graff Zivin, Katherine Meckel, Matthew Neidell, and Wolfram Schlenker. 2013. Something in the water: contaminated drinking water and infant health. *Canadian Journal of Economics* 46 (3): 791–810.

Cutler, David and Grant Miller. 2005. The role of public health improvements in health advances: The twntieth-century united states. *Demography* 42 (1): 1–22.

Dodds, Walter K., Wes W. Bouska, Jeffrey L. Eitzmann, Tyler J. Pilger, Kristen L. Pitts, Alyssa J. Riley, Joshua T. Schloesser, and Darren J. Thornbrugh. 2009. Eutrophication of u.s. freshwaters: Analysis of potential economic damages. *Environmental Science and Technology* 43 (1): 12–19.

Ebenstein, Avraham. 2012. The consequences of industrialization: Evidence from water pollution and digestive cancers in china. *Review of Economics and Statistics* 94 (1): 186–201.

Egan, Kevin J., Joseph A. Herriges, Catherine L. Kling, and John A. Downing. 2009. Valuing water quality as a function of water quality measures. *American Journal of Agricultural Economics* 91 (1): 106–123.

Feather, Peter and Daniel Hellerstein. 1997. Calibrating benefit function transfer to assess the conservation reserve program. *American Journal of Agricultural Economics* 79 (1): 151–162.

Freeman, Myrick A. 1982. *Air and Water Pollution Control.*John Wiley & Sons.

Garcia, Ana Maria, Anne B. Hoos, and Silvia Terziotti. 2011. A regional modeling framework of phosphorus sources and transport in streams of the southeastern united states. *Journal of the American Water Resources Association* 47 (5): 991–1010.

Garnache, Cloe, Scott M. Swinton, Joseph A. Herriges, Frank Lupi, and R. Jan Stevenson. 2016. Solving the phosphorus pollution puzzle: Synthesis and directions for future researc. *American Journal of Agricultural Economics* 98 (5): 1334–1359.

Greenstone, Michael and Justin Gallagher. 2008. Does hazardous waste matter? evidence from the housing market and the superfund program. *Quarterly Journal of Economics* 123 (3): 951–1003.

von Haefen, Roger H. and Daniel J. Phaneuf. 2008. Identifying demand parameters in the presence of unobservables: A combined revealed and stated preference approach. *Journal of Environmental Economics and Management* 56 (1): 19–32.

Hansen, LeRoy. 2007. Conservation reserve program: Environmental benefits update. *Agricultural and Resource Economics Review* 36 (2): 1–14.

Harmel, R.D., R.J. Cooper, R.M. Slade, R.L. Haney, and J.G. Arnold. 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Transactions of the American Society of Agricultural and Biological Engineers* 49 (3): 689–701.

Hausman, Jerry A., Gregory K. Leonard, and Daniel McFadden. 1995. A utility-consistent, combined discrete choice and count data model assessing recreational use losses due to natural resource damage. *Journal of Public Economics* 56: 1–30.

Hoos, Anne B. and Gerard McMahon. 2009. Spatial analysis of instream nitrogen loads and factors controlling nitrogen delivery to streams in the southeastern united states using spatially referenced regression on watershed attributes (sparrow) and regional classification frameworks. *Hydrological Processes* 23 (16): 2275–2294.

Hyslop, Dean R. and Guido W. Imbens. 2001. Bias from classical and other forms of measurement error. *Journal of Business and Economic Statistics* 19 (4): 475–481.

IA DNR and ISU Limnology Laboratory. 2001-2012. Iowa Lakes Monitoring Data. Available online at http://limnology.eeob.iastate.edu/lakereport/chemical_report.aspx (accessed December 1, 2014).

Ji, Yongjie, Joseph A. Herriges, and Catherine L. Kling. 2016. Modeling recreation demand when the access point is unknown. *American Journal of Agricultural Economics* 98 (3): 860–880.

Keeler, Bonnie L., Stephen Polasky, Kate A. Brauman, Kris A. Johnson, Jacques C. Finlay, Ann O'Neill, Kent Kovacs, and Brent Dalzell. 2012. Linking water quality and well-being for improved assessment and valuation of ecosystem services. *Proceedings of the National Academy of Sciences of the United States of America* 109 (45): 18619–18624.

Keiser, David A. and Nicholas Z. Muller. 2017. Air and water: Integrated assessment models for multiple media. *Annual Review of Resource Economics* 9: 165–184.

Keiser, David A. and Joseph S. Shapiro. 2017. Consequences of the clean water act and the demand for water quality. NBER Working Paper 23070, National Bureau of Economic Research, Cambridge, MA.

Kleibergen, Frank and Richard Paap. 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* 133 (1): 97–126.

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane. 2011. Spring cleaning: Rural water impacts, valuation, and property rights institutions. *Quarterly Journal of Economics* 126 (1): 145–205.

Leggett, Christopher G. and Nancy E. Bockstael. 2000. Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management* 39: 121–144.

Lyon, Randolph M. and Scott Farrow. 1995. An economic analysis of clean water act issues. *Water Resources Research* 31 (1): 213–223.

McMillan, Hilary, Tobias Krueger, and Jim Freer. 2012. Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. *Hydrological Processes* 26 (26): 4078–4111.

Mendelsohn, Robert, Daniel Hellerstein, Michael Huguenin, Robert Unsworth, and Richard Brazee. 1992. Measuring hazardous waste damages with panel models. *Journal of Environmental Economics and Management* 22: 259–271.

Milstead, W. Bryan, Jeffrey W. Hollister, Richard B. Moore, and Henry A. Walker. 2013. Estimating summer nutrient cconcentration in northeastern lakes from sparrow load predictions and modeled lake depth and volume. *Public Library of Science One* 8 (11): e81457.

Moeltner, Klaus and Roger von Haefen. 2011. Microeconometric strategies for dealing with unobservables and endogenous variables in recreation demand models. *Annual Reviews of Resource Economics* 3: 375–96.

Moore, Richard B., Craig M. Johnston, Richard A. Smith, and Bryan Milstead. 2011. Source and delivery of nutrients to receiving waters in the northeastern and mid-atlantic regions of the united states. *Journal of the American Water Resources Association* 47 (5): 965–990.

Mullahy, John. 1997. Instrumental-variable estimation of count data models: Applications to models of cigarette smoking behavior. *The Review of Economics and Statistics* 79(4): 1649–1675.

Muller, Nicholas Z. and Robert Mendelsohn. 2007. Measuring the damages of air pollution in the united states. *Journal of Environmental Economics and Management* 54 (1): 1–14.

——— 2009. Efficient pollution regulation: Getting the prices right. *American Economic Review* 99 (5): 1714–1739.

Muller, Nicholas Z., Robert Mendelsohn, and William Nordhaus. 2011. Environmental accounting for pollution in the united states economy. *American Economic Review* 101 (5): 1649–1675.

Murdock, Jennifer. 2006. Handling unobserved site characteristics in random utility models of recreational demand. *Journal of Environmental Economics and Management* 51 (1): 1–25.

National Commission on Water Quality. 1976. Staff report.Technical report, Washington DC.

National Planning Association. 1975. Water-related recreation benefits resulting from p.l. 92-500..Technical report, Washington DC.

NOAA. 2000. Medium Resolution Shoreline. Available online at http://shoreline.noaa.gov/data/datasheets/medres.html (accessed August 14, 2014).

NRCS. 1997. National Resources Inventory.

Olmstead, Sheila M. 2010. The economics of water quality. *Review of Environmental Economics & Policy* 4 (1): 44–61.

Peskin, Henry M. and Winston Harrington. 1981. Policy-relevant applications of a system of expanded national accounts: effects of water quality on recreation behavior.Technical report.

Phaneuf, Daniel J., Catherine L. Kling, and Joseph A. Herriges. 1998. Valuing water quality improvements using revealed preference methods when corner ssolution are present. *American Journal of Agricultural Economics* 80 (5): 1025–1031.

Phaneuf, Daniel J. and V. Kerry Smith. 2005. Recreation demand models. In *Handbook of Environmental Economics*, ed. Mäler, Karl-Göran and Jeffrey R. Vincent.Amsterdam: Elsevier.

PRISM Climate Group. 1981-2010. 30-Year Normals. Available online at http://www.prism.oregonstate.edu/normals/ (accessed February 11, 2015).

Rebich, Richard A., Natalie A. Houston, Scott V. Mize, Daniel K. Pearson, Patricia B. Ging, and C. Evan Horning. 2011. Sources and delivery of nutrients to the northwestern gulf of mexico from streams in the south-central united states. *Journal of the American Water Resources Association* 47 (5): 1061–1086.

Ribaudo, Marc O. and Steven L. Piper. 1991. Estimating changes in recreational fishing participation from national water quality policies. *Water Resources Research* 27 (7): 1757–1763.

Robertson, Dale M. and David A. Saad. 2011. Nutrient inputs to the laurentian great lakes by source and watershed estimated using sparrow watershed models. *Journal of the American Water Resources Association* 47 (5): 1011–1033.

——— 2013. Sparrow models used to understand nutrient sources in the mississippi/atchafalaya river basin. *Journal of Environmental Quality* 42: 1422–1440.

Russell, Clifford S. and William J. Vaughan. 1982. The national recreational fishing benefits of water pollution control. *Journal of Environmental Econonomics and Management* 9 (4): 328–354.

Schlenker, Wolfram and W. Reed Walker. 2016. Airports, air pollution, and contemporaneous health. *The Review of Economic Studies* 83 (2): 768–809.

Shapiro, Samuel S. and R. S. Francia. 1972. An approximate analysis of variance test for normality. *Journal of the American Statistical Association* 67 (337): 215–216.

Shapiro, Samuel S. and Martin B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52 (3-4): 591–611.

Smith, Richard A., Gregory E. Schwarz, and Richard B. Alexander. 1997. Regional interpretation of water-quality monitoring data. *Water Resources Research* 33(12): 2781–2798.

Smith, Val H. and David W. Schindler. 2009. Eutrophication science: Where do we go from here? *Trends in Ecology and Evolution* 24 (4): 201–207.

Stock, James H. and Motohiro Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. Donald W. K. Andrews and James H. Stock.Cambridge University Press.

Timmins, Christopher and Jennifer Murdock. 2007. A revealed preference approach to the measurement of congestion in travel cost models. *Journal of Environmental Economics and Management* 53 (2): 230–249.

U.S. Census Bureau. 2009. Metropolitan and micropolitan statistical areas and components. Available online at http://www.census.gov/population/metro/data/defhist.html (accessed August 13, 2014).

U.S. Department of Transportation, Federal Highway Administration. 2009. 2009 National Household Travel Survey. Available online at http://nhts.ornl.gov/det/ (accessed March 21, 2016).

——— 1986-2014. CRP Enrollment and Rental Payments by State, 1986-2014. Available online at http://www.fsa.usda.gov/programs-and-services/conservation-programs/reports-and-statistics/conservation-reserve-program-statistics/index (accessed April 17, 2015).

USDA. 2009. National Survey on Recreation and the Environment.

USEPA. 1997. The benefits and costs of the clean air act, 1970 to 1990.Technical report, USEPA.

——— 1999. The benefits and costs of the clean air act, 1990 to 2010.Technical report, USEPA.

——— 2000. Nutrient criteria technical guidance manual: Lakes and reservoirs.Technical report, USEPA.

——— 2007. National Lakes Assessment. Available online at http://water.epa.gov/type/lakes/lakessurvey_index.cfm (accessed March 13, 2015).

——— 2009. An Urgent Call to Action: Report of the State-EPA Nutrient Innovations Task Group.

——— 2010. Chesapeake bay tmdl.Technical report, USEPA.

——— 2011. The benefits and costs of the clean air act from 1990 to 2020.Technical report, USEPA.

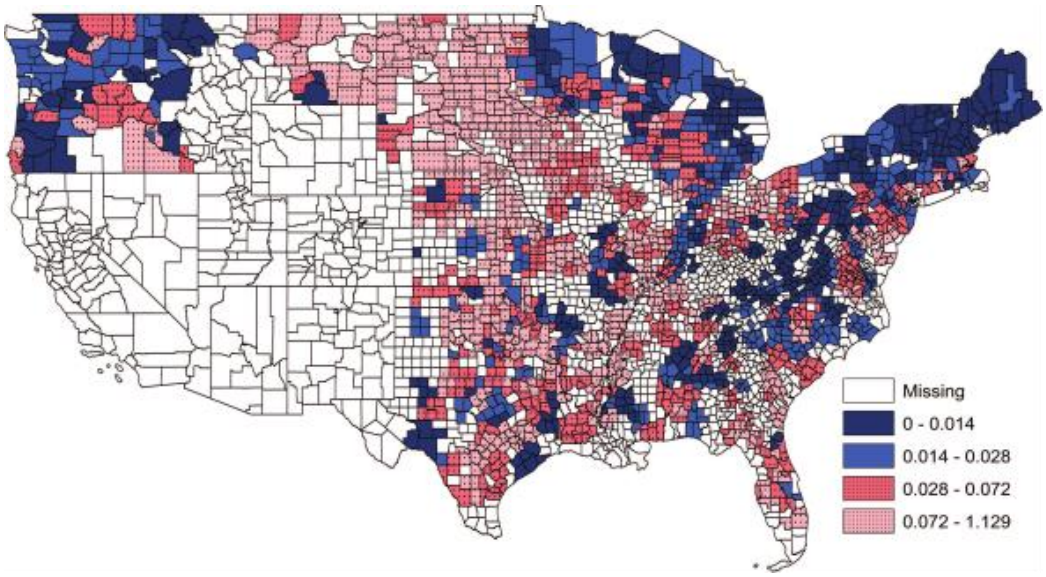
USGS. 2001. National Land Cover Database 2001. Available online at http://www.mrlc.gov/nlcd2001.php (accessed March 11, 2015).

——— 2002. SPARROW Decision Support System. Available online at http://cida.usgs.gov/sparrow/ (accessed April 9, 2015).

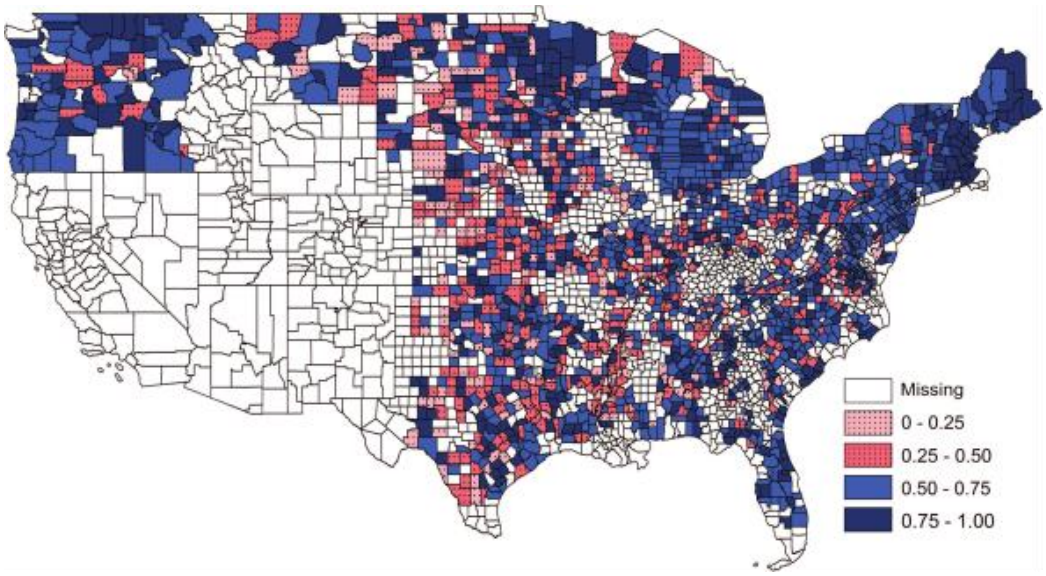
USGS. 2015. National Elevation Dataset, 1 arc-second. Available online at http://nationalmap.gov/ (accessed February 5, 2015).

Van Houtven, George L., Smita B. Brunnermeier, and Mark C. Buckley. 2000. A retrospective assessment of the costs of the clean water act: 1992 to 1997.Technical report, Prepared by Research Triangle Institute for the Office of Water and Office of Policy, Economics, and Innovation, U.S. Environmental Protection Agency, Washington, DC.

Wise, Daniel R. and Henry M. Johnson. 2011. Surface-water nutrient conditions and sources in the united states pacific northwest. *Journal of the American Water Resources Association* 47 (5): 1110–1135.



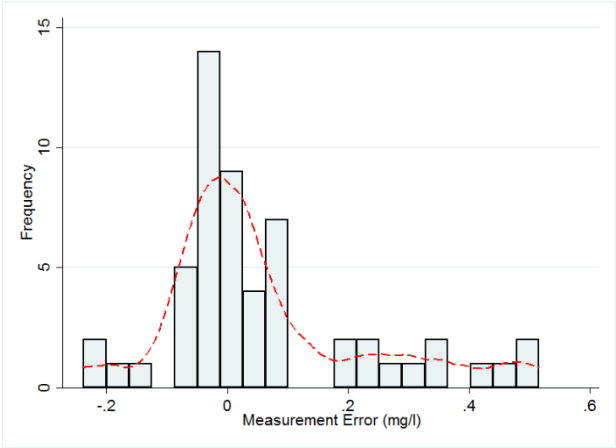
PANEL A. PHOSPHORUS CONCENTRATIONS (MG/L)



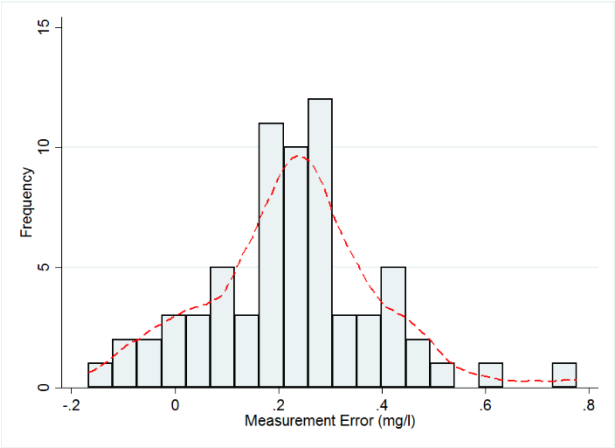
PANEL B. WATER-BASED RECREATIONAL PARTICIPATION RATES BY COUNTY

FIGURE 1 – PHOSPHORUS CONCENTRATIONS AND RECREATION

Note. - Figure 1 shows average county-level phosphorus concentrations (Panel A) and average county-level participation rates for water-based recreational use (Panel B) for the USEPA estimating sample. Concentration levels are shown by quantile. The source of recreational data is the National Survey on Recreation and the Environment (NSRE) from 1999 to 2009.



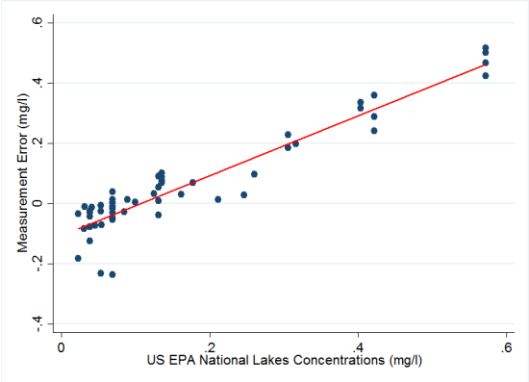
PANEL A. US EPA NATIONAL LAKES



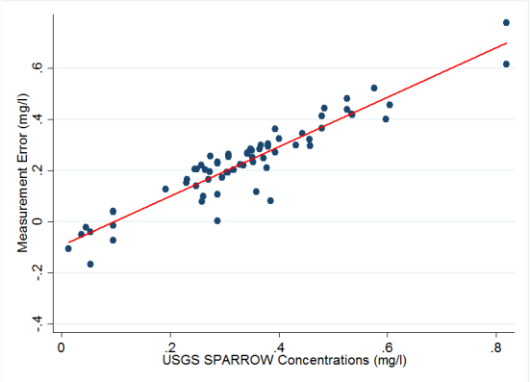
PANEL B. USGS SPARROW

FIGURE 2 – MEASUREMENT ERROR DISTRIBUTIONS

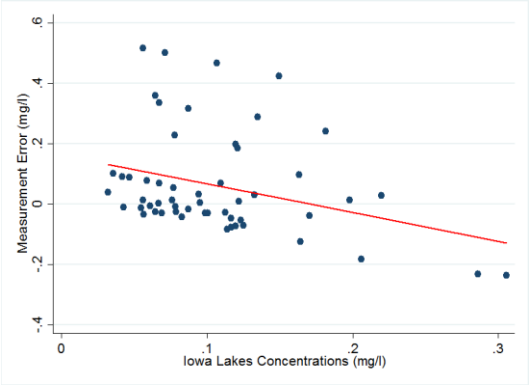
Notes: Figure 2, Panels A and B show measurement error distributions for average county-level phosphorus concentrations in the US EPA National Lakes (Panel A) and USGS SPARROW samples (Panel B). Epanechnikov kernel density function shown with dashed red line.



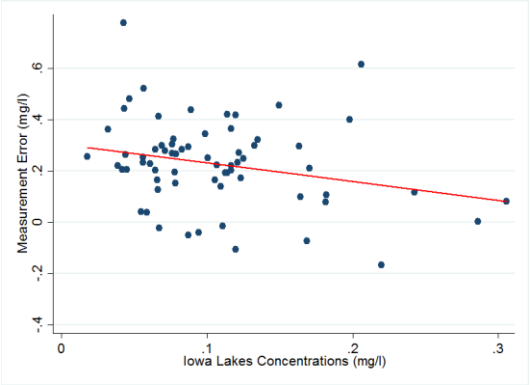
PANEL A. MEASUREMENT ERROR VS. US EPA NATIONAL LAKES



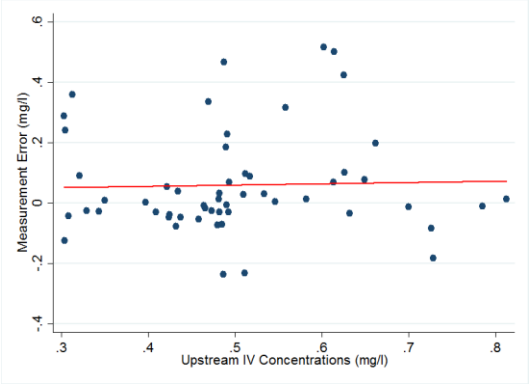
PANEL B. MEASUREMENT ERROR VS. USGS SPARROW



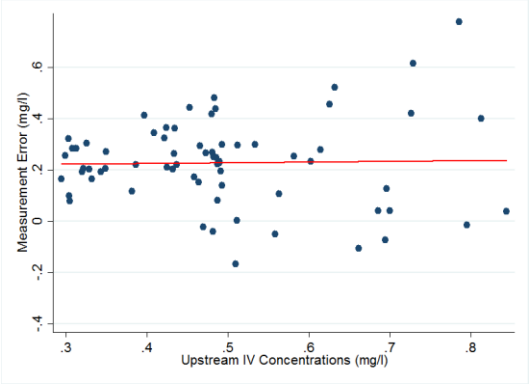
PANEL C. MEASUREMENT ERROR VS. IOWA LAKES (US EPA NATIONAL LAKES SAMPLE)



PANEL D. MEASUREMENT ERROR VS. IOWA LAKES (USGS SPARROW SAMPLE)



PANEL E. MEASUREMENT ERROR VS. UPSTREAM IV (30 MILES, US EPA NATIONAL LAKES SAMPLE)



PANEL F. MEASUREMENT ERROR VS. UPSTREAM IV (30 MILES, USGS SPARROW SAMPLE)

FIGURE 3 – MEASUREMENT ERROR DISTRIBUTIONS AND SCATTER PLOTS

Notes: Figure 3, Panels A through F display scatter plots of measurement error estimates versus noisy measurements (e.g., US EPA (Panel A) and USGS (Panel B) samples), "true" estimates (e.g., Iowa Lakes from the US EPA (Panel C) and USGS (Panel D) samples), and upstream instrumental variables (IV) concentrations at >0 to 200 miles in the US EPA (Panel E) and USGS (Panel F) samples. Fitted line from ordinary least squares regression of measurement error on corresponding independent variable shown with solid red line. All observations reflect county-level average phosphorus concentrations.

TABLE 1 – PARTICIPATION RATES AND NUTRIENT CONCENTRATIONS BY CENSUS DIVISION

	US EPA National Lakes Sample			**USGS SPARROW Sample**		
Census Division	Participation Rate	Phosphorus (mg/l)	Observations	Participation Rate	Phosphorus (mg/l)	Observations
New England	78%	0.043	3,177	77%	0.035	3,313
Mid Atlantic	68%	0.032	4,216	67%	0.087	4,819
East North Central	65%	0.046	5,909	64%	0.195	6,916
West North Central	63%	0.141	3,444	62%	0.285	3,842
South Atlantic	68%	0.059	5,609	66%	0.092	6,130
East South Central	60%	0.056	1,894	60%	0.124	2,952
West South Central	60%	0.072	3,270	61%	0.297	3,907
Mountain	70%	0.157	413	70%	0.121	426
Pacific	71%	0.018	1,455	71%	0.069	1,552
Total	67%	0.061	29,387	66%	0.154	33,857

Notes: Census divisions and corresponding states used in this analysis are New England (CT, ME, MA, NH, RI, VT), Middle Atlantic (NJ, NY, PA), East North Central (IN, IL, MI, OH, WI), West North Central (IA, KS, MN, MO, NE, ND, SD), South Atlantic (DE, DC, FL, GA, MD, NC, SC, VA, WV), East South Central (AL, KY, MS, TN), West South Central (AR, LA, OK, TX), Mountain (ID, MT), and Pacific (OR, WA).

TABLE 2 – THE EFFECT OF NUTRIENTS ON WATER-BASED RECREATIONAL USE
(US EPA NATIONAL LAKES)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Negative Binomial and GMM Count Models						
Phosphorus (mg/l)	-0.379**	-3.405***	-3.482***	-3.797***	-2.756	-4.090***
	(0.168)	(0.990)	(0.785)	(1.346)	(1.709)	(1.069)
Instruments	None	>0 - 200	>0 - 200	30 - 200	60 - 200	> 0 - 200; 30 - 200
Fixed Effects	State	State	Cen. Div.	State	State	State
Observations	29,387	29,387	29,387	26,073	17,677	26,073
Hansen's J Stat						0.145
						(p = 0.703)
Panel B: Semi-log Models						
Phosphorus (mg/l)	-0.319***	-4.874***	-4.915***	-5.659**	-5.495	-5.941***
	(0.102)	(1.583)	(1.372)	(2.411)	(3.543)	(2.221)
Instruments	None	>0 - 200	>0 - 200	30 - 200	60 - 200	> 0 - 200; 30 - 200
Fixed Effects	State	State	Cen. Div.	State	State	State
Observations	29,387	29,387	29,387	26,073	17,677	26,073
Hansen's J Stat						0.093
						(p = 0.761)
Kleibergen-Paap		13.452	16.411	6.361	2.800	4.364
rk Wald F Stat		(F = 16.38)	(F = 16.38)	(F = 16.38)	(F = 16.38)	(F = 19.93)

Notes: Dependent variable in Panel A is the number of water-based recreational days taken in a year. Dependent variable in Panel B is ln(recreational days + 1). All regressions include state fixed effects (except column 3, which includes census division fixed effects) and controls for metropolitan and micropolitan statistical areas, coastal and Great Lakes shorelines, 30-year climate normals for temperature and precipitation and temperature squared, indicator for counties with a lake, race, gender, age, income, education, elevation, and predominant land cover type. Estimate for Panel A columns (4) and (5) include a dummy variable for caucasian instead of dummy variables for all races so that the model converges. Instrumental variables are concentrations of phosphorus from upstream counties at >0 to 200 miles, 30 to 200 miles, and/or 60 to 200 miles. For parameter estimates, robust standard errors, clustered at the county level, in parentheses. For Kleibergen-Paap rk Wald F Stat, Stock-Yogo weak identification critical value for 10% maximal instrumental variable size in parentheses. Critical value for 15% maximal instrumental variable size equals 8.96 for specifications (2) - (5) and 11.59 for specification (6).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 3 – THE EFFECT OF NUTRIENTS ON WATER-BASED RECREATIONAL USE
(USGS SPARROW)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Negative Binomial and GMM Count Models						
Phosphorus (mg/l)	-0.205*	-2.561***	-1.935***	-2.785***	-3.399***	-2.381***
	(0.119)	(0.696)	(0.405)	(0.803)	(0.985)	(0.658)
Instruments	None	>0 - 200	>0 - 200	30 - 200	60 - 200	> 0 - 200; 30 - 200
Fixed Effects	State	State	Cen. Div.	State	State	State
Observations	33,857	33,857	33,857	29,689	20,406	29,689
Hansen's J Stat						0.803
						(p = 0.370)
Panel B: Semi-log Models						
Phosphorus (mg/l)	-0.247***	-3.119***	-2.405***	-3.774**	-5.218*	-2.934***
	(0.0702)	(1.150)	(0.600)	(1.781)	(3.041)	(1.014)
Instruments	None	>0 - 200	>0 - 200	30 - 200	60 - 200	> 0 - 200; 30 - 200
Fixed Effects	State	State	Cen. Div.	State	State	State
Observations	33,857	33,857	33,857	29,689	20,406	29,689
Hansen's J Stat						0.942
						(p = 0.332)
Kleibergen-Paap		6.718	16.570	3.640	3.062	5.000
rk Wald F Stat		(F = 16.38)	(F = 16.38)	(F = 16.38)	(F = 16.38)	(F = 19.93)

Notes: Dependent variable in Panel A is the number of water-based recreational days taken in a year. Dependent variable in Panel B is ln(recreational days + 1). All regressions include state fixed effects (except column 3, which includes census division fixed effects) and controls for metropolitan and micropolitan statistical areas, coastal and Great Lakes shorelines, 30-year climate normals for temperature and precipitation and temperature squared, indicator for counties with a lake, race, gender, age, income, education, elevation, and predominant land cover type. Estimate for Panel A, columns (4) and (5) include a dummy variable for caucasian instead of dummy variables for all races so that the model converges. Instrumental variables are concentrations of phosphorus from upstream counties at >0 to 200 miles, 30 to 200 miles, and/or 60 to 200 miles. For parameter estimates, robust standard errors, clustered at the county level, in parentheses. For Kleibergen-Paap rk Wald F Stat, Stock-Yogo weak identification critical value for 10% maximal instrumental variable size in parentheses. Critical value for 15% maximal instrumental variable size equals 8.96 for specifications (2) - (5) and 11.59 for specification (6).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 4 – REDUCED FORM AND FIRST STAGE PHOSPHORUS RESULTS
(US EPA NATIONAL LAKES AND USGS SPARROW)

	US EPA National Lakes Sample			USGS SPARROW Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Reduced Form						
IV >0 - 200	-0.375***			-0.423***		
	-0.087			(0.0737)		
IV 30 - 200		-0.332***			-0.414***	
		(0.111)			(0.0939)	
IV 60 - 200			-0.251**			-0.412***
			(0.124)			(0.114)
Panel B: First Stage						
IV >0 - 200	0.0769***			0.136***		
	(0.0210)			(0.0524)		
IV 30 - 200		0.0586**			0.110*	
		-0.0233			(0.0575)	
IV 60 - 200			0.0454*			0.0789*
			(0.0271)			(0.0451)
Instruments	>0 - 200	30 - 200	60 - 200	>0 - 200	30 - 200	60 - 200
Fixed Effects	State	State	State	State	State	State
Observations	29,387	26,073	17,677	33,857	29,689	20,406

Notes: Dependent variable in Panel A is ln(recreational days + 1). Dependent variable in Panel B is local phosphorus concentration. All regressions include state fixed effects and controls for metropolitan and micropolitan statistical areas, coastal and Great Lakes shorelines, 30-year climate normals for temperature and precipitation and temperature squared, indicator for counties with a lake, race, gender, age, income, education, elevation, and predominant land cover type. Instrumental variables are concentrations of phosphorus from upstream counties at >0 to 200 miles, 30 to 200 miles, and/or 60 to 200 miles. For parameter estimates, robust standard errors, clustered at the county level, in parentheses.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 5 – ESTIMATED ATTENUATION FACTORS AND RATIOS OF COEFFICIENT ESTIMATES

	Attenuation Factors		Ratio of Estimates			
			Single Regressor		With Controls	
	Classical	Non-Classical	Count Models	Semi-Log	Count Models	Semi-Log
	(1)	(2)	(3)	(4)	(5)	(6)
US EPA National Lakes	0.101	0.008	0.258	0.173	0.111	0.065
USGS SPARROW	0.108	0.036	0.579	0.500	0.080	0.079

Notes: The ratios of parameter estimates in columns (3) to (6) equal the negative binomial (count model) or OLS (semi-log) parameter estimate of phosphorus divided by the corresponding instrumental variable parameter estimate. Instrumental variable parameter estimates are from regressions using upstream instrument concentrations at >0-200 miles upstream. Controls include state fixed effects and controls for metropolitan and micropolitan statistical areas, coastal and Great Lakes shorelines, 30-year climate normals for temperature and precipitation and temperature squared, indicator for counties with a lake, race, gender, age, income, education, elevation, and predominant land cover type. Attenuation factor calculations (columns (1) to (2)) use 55 counties (US EPA National Lakes) or 68 counties (USGS SPARROW).

TABLE 6 –MEASUREMENT ERROR VERSUS LOCAL CONCENTRATIONS

	Measured Concentration (1)	Measured Concentration (2)	IA Lakes (3)	IA Lakes (4)
Panel A: US EPA National Lakes				
Phosphorus (mg/l)	0.994***	1.100***	-0.954***	-2.180**
	(0.0409)	(0.0674)	(0.329)	(0.898)
Controls	N	Y	N	Y
R^2	0.887	0.982	0.103	0.79
Counties	55	55	55	55
Panel B: USGS SPARROW				
Phosphorus (mg/l)	0.965***	0.866***	-0.731**	-0.412
	(0.0463)	(0.0921)	(0.366)	(0.466)
Controls	N	Y	N	Y
R^2	0.880	0.951	0.065	0.786
Counties	68	68	68	68

Notes: Coefficient estimates are from linear regressions of measurement error estimates on concentrations from the IA Lakes data or measured concentration (US EPA National Lakes or USGS SPARROW). Additional controls include dummy variables for metropolitan and micropolitan statistical areas, 30-year climate normals for temperature and precipitation and temperature squared, indicators for counties with a lake, race, gender, age, income, education, and elevation. Cropland is predominant land cover type in all counties. Robust standard errors in parentheses.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 7 –MEASUREMENT ERROR
VERSUS INSTRUMENTAL VARIABLE CONCENTRATIONS

	(1)	(2)
Panel A: US EPA National Lakes		
Phosphorus IV >0 - 200	0.0427	-1.009
	(0.186)	(0.801)
Controls	N	Y
R^2	0.001	0.733
Counties	55	55
Panel B: USGS SPARROW		
Phosphorus IV >0 - 200	0.0301	0.350
	(0.199)	(0.271)
Controls	N	Y
R^2	0.001	0.786
Counties	68	68

Notes: Coefficient estimates are from linear regressions of measurement error estimates on instrumental variable concentrations in the two primary estimating samples (US EPA National Lakes or USGS SPARROW). Controls include dummy variables for metropolitan and micropolitan statistical areas, 30-year climate normals for temperature and precipitation and temperature squared, indicators for counties with a lake, race, gender, age, income, education, and elevation. Cropland is predominant land cover type in all counties. Robust standard errors in parentheses.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Appendix

This appendix describes a number of robustness checks and placebo tests.

As an indirect test of the validity of the empirical approach in Section 4, I examine the estimated effects of phosphorus on individual water-based recreational uses (e.g., swimming, cold and warm water fishing, boating (with a motor) and visiting a waterbody for any other recreational purpose) and 25 additional activities reported in the NSRE. These additional activities serve as placebo tests since they are not obviously defined as activities where water quality would affect use. For this analysis, I examine the effect of phosphorus on the extensive margin (decision to participate in recreation) for three reasons. The first is to examine the effects of the sample selection step that drops a random sample of "no" respondents as described in Section 3.2. The second reason is to take advantage of all possible observations since the intensive margin (number of days) is not asked of all respondents that participate in an activity. Lastly, I use the full survey respondent samples to examine the effects of including or excluding survey weights. Appendix Tables 1 and 2 present these results. Appendix Tables 3 and 4 report results using survey weights.

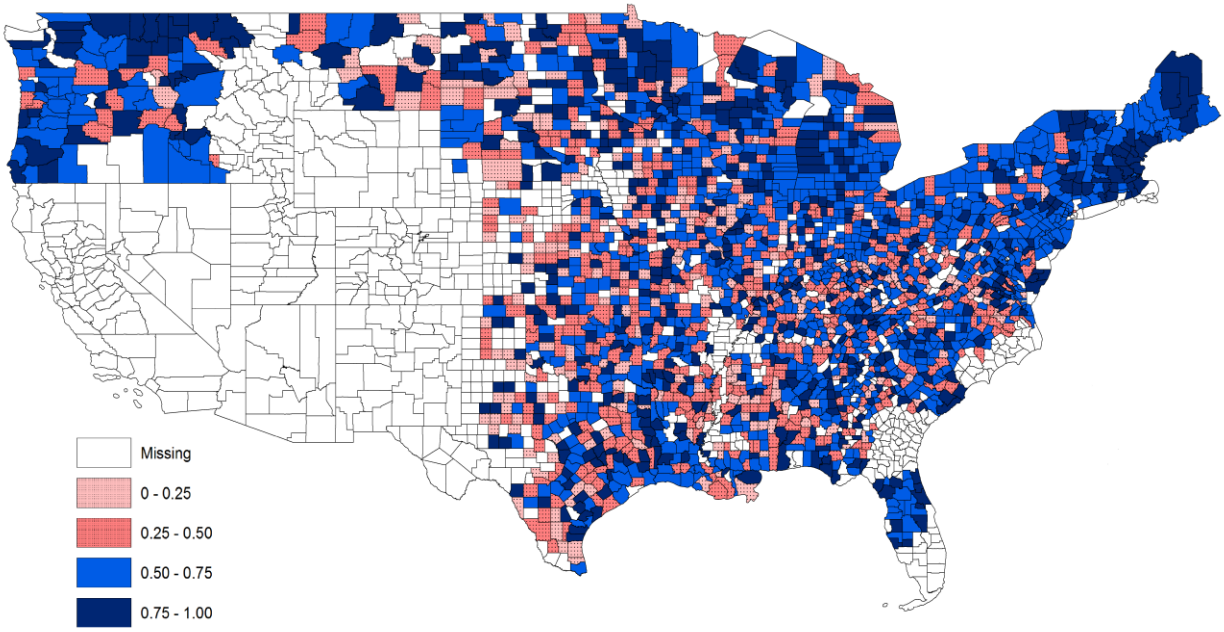
In Appendix Tables 1 and 2, for each activity, I first report a specification that uses the "Normal" set of controls from section 4.2. This includes state fixed effects and other demographic, geographic, and climate controls. For each activity, I also report a second set of results that show "+Activ." controls. These specifications include controls for participation in every additional non-water based recreational activity. These are included in these specifications to control for preferences for other recreational activities. The first two specifications shown in Panel A for Appendix Tables 1 and 2 examine the extensive margin for water-based recreational days using the estimating samples from Section 4.2. Consistent with the results in Tables 2 and 3, phosphorus has a negative and significant effect on the decision to participate in water-based recreation. Adding controls for other recreational activities yields very similar point estimates that are all highly significant except for the USEPA sample. In this specification the point estimate is nearly identical but no longer significant. The next two columns reflect all responses from the NSRE regardless of whether the respondent answered questions regarding the intensive margin (i.e., the number of days for a particular activity). These coefficient estimates are negative and similar in magnitude to the estimating samples used in Section 4.2.

The next five activities are the activities that comprise the measure of water-based recreational use. For these activities, swimming and coldwater fishing are negative and highly significant in both estimating samples with and without the addition of controls for other recreational activities. Warmwater fishing, boating, and visiting a waterbody are negative and significant in the samples without the additional recreational controls, but become insignificant and of a smaller magnitude when adding controls for additional recreational activities. These results suggest that swimming and coldwater fishing are the most responsive water-based recreational uses to changes in phosphorus concentrations.

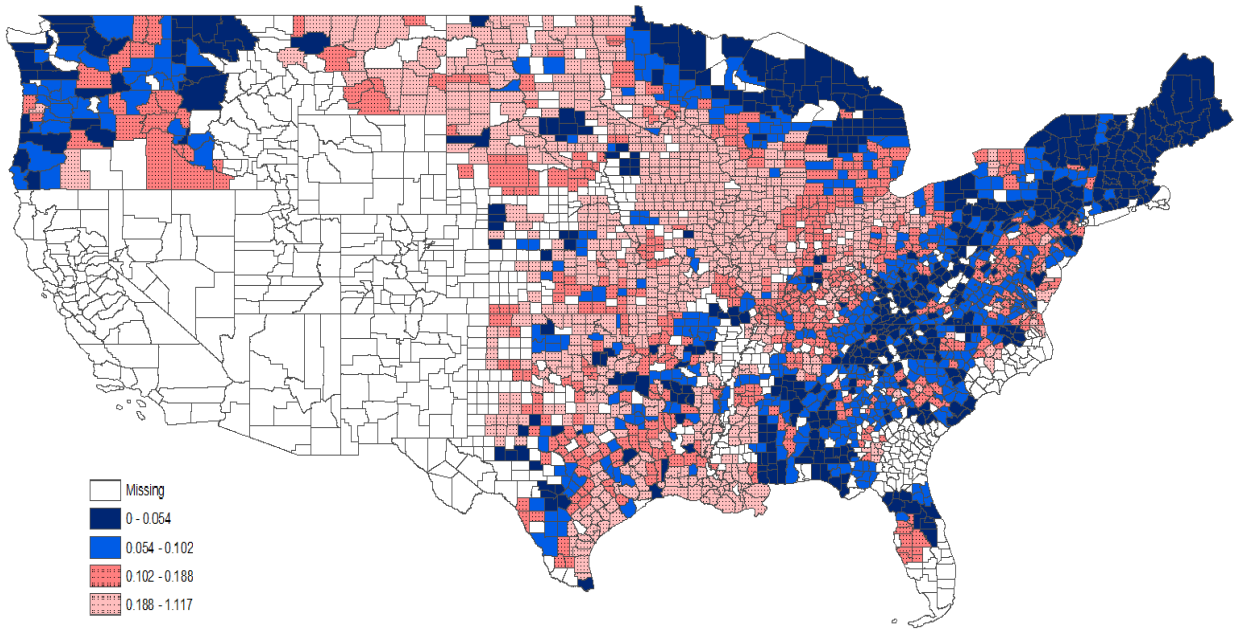
The next 25 columns in Appendix Tables 1 and 2 display the results for the placebo regressions. Of these 25 activities, a number of activities have negative and significant signs when considering the "Normal" controls. For instance, the coefficients on visiting a wilderness place and gathering mushrooms are negative and significant in both the USEPA and USGS samples. However, when controls are added for other recreational activities, the only activities (other than water-based recreational activities) that are significant are swimming in a pool and driving off road. Interestingly, and perhaps supportive of the empirical approach, is that the sign of swimming in an outdoor pool is positive. Intuitively, this suggests that as water quality in natural bodies of water decreases, individuals substitute to man-made pools for swimming. It is unclear why driving

off-road remains negative and highly significant in these specifications.[1] I find similar effects when considering survey weights (see Appendix Tables 3 and 4). In summary, the placebo tests are generally supportive of the empirical approach. These results add some reassurance to the validity of the findings in section 4.2.

[1]One possible explanation is that individuals drive off-road for water-based recreational activities. However, the correlation coefficient between water-based recreation and driving off-road is approximately 0.2, which does not suggest an incredibly strong relationship between the two activities. Similar correlation coefficients exist between water-based recreation and other activities ranging from approximately 0.1 for viewing birds to 0.40 for swimming in an outdoor pool.



APPENDIX FIGURE 1 – WATER-BASED RECREATIONAL PARTICIPATION RATE BY COUNTY (USGS SPARROW)



APPENDIX FIGURE 2 – PHOSPHORUS CONCENTRATIONS (MG/L)
(USGS SPARROW)

APPENDIX TABLE 1 – EXTENSIVE MARGIN AND PLACEBO TESTS (US EPA NATIONAL LAKES)

Panel A: Recreation Activities - Part 1

	Water-based Recreation		Water-based Recreation		Swim Outdoors		Coldwater Fishing		Warmwater Fishing		Motor Boating		Visit Waterbody		Swimming Outdoor Pool	
Phosphorus	-2.771**	-2.740	-3.165***	-2.909**	-3.120***	-4.069***	-4.185***	-5.466***	-2.546**	-0.132	-1.707**	-0.351	-2.492***	-1.986	1.328	3.736**
(mg/l)	(1.201)	(1.959)	(0.968)	(1.415)	(0.865)	(1.487)	(1.191)	(1.896)	(0.999)	(1.285)	(0.721)	(1.178)	(0.905)	(1.351)	(0.832)	(1.491)
Observations	29,387	9,574	45,412	14,521	59,541	14,961	56,729	14,916	51,852	14,916	59,531	14,948	47,351	14,600	33,320	14,964
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Panel B: Recreation Activities - Part 2

	Bicycling		Hiking		Snow and Ice Activities		Outdoor Team Sports		Horseback Riding		Picnicking		Visit Nature Center		Visit Prehistoric Site	
Phosphorus	0.329	1.397	-1.463*	-0.425	-0.519	0.327	0.367	0.634	-0.916	-0.771	-0.229	0.108	-0.0231	-2.498*	-0.419	0.701
(mg/l)	(0.660)	(1.541)	(0.762)	(1.235)	(0.703)	(1.241)	(0.950)	(1.414)	(0.845)	(1.558)	(0.613)	(1.122)	(0.648)	(1.421)	(0.840)	(1.358)
Observations	54,198	11,427	54,118	14,964	53,311	14,964	35,033	14,964	52,800	14,964	53,849	14,964	53,850	14,964	51,206	14,964
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Panel C: Recreation Activities - Part 3

	Visit Historic Site		Walk for Pleasure		Visit Wilderness		Gather Mushrooms		Hunt		Mountain Biking		Backpacking		Developed Camping	
Phosphorus	1.192	1.616	-0.0754	0.0553	-1.476**	0.789	-2.608***	-1.511	-1.372	0.900	-0.0832	2.262	-2.711***	-1.914	-0.592	0.0418
(mg/l)	(0.821)	(1.350)	(0.857)	(1.721)	(0.685)	(1.277)	(0.859)	(1.204)	(1.035)	(1.714)	(0.753)	(1.967)	(1.031)	(2.047)	(0.665)	(1.269)
Observations	47,920	14,964	54,763	14,985	54,056	14,985	52,001	14,964	62,280	14,967	52,483	6,439	54,793	14,964	54,181	14,964
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Panel D: Recreation Activities - Part 4

	Primitive Camping		View Birds		View Fish		View Wildlife		View Flowers		View Scenery		Drive for Pleasure		Drive Off-road	
Phosphorus	-1.459*	-1.185	-0.450	1.818	-1.883**	-1.296	-1.584**	-0.872	-0.686	-0.136	-0.759	0.712	-1.355*	-0.842	-2.528**	-4.445***
(mg/l)	(0.790)	(1.314)	(0.665)	(1.187)	(0.788)	(1.419)	(0.702)	(1.206)	(0.654)	(1.314)	(0.630)	(1.270)	(0.759)	(1.158)	(0.989)	(1.693)
Observations	50,488	14,964	60,061	14,964	60,356	14,964	59,614	14,964	59,748	14,964	59,784	14,964	48,582	14,964	48,582	14,964
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Notes: Dependent variable is participation in specified activity. All estimates are control function estimates using Stata's IV Probit command. "Normal" controls include state fixed effects and all standard controls used throughout the paper. "+ Activ" controls also include participation in non-water based activities. See text for more details. Robust standard errors, clustered at the county level, in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

APPENDIX TABLE 2 – EXTENSIVE MARGIN AND PLACEBO TESTS (USGS SPARROW)

Panel A: Recreation Activities - Part 1

	Water-based Recreation		Water-based Recreation		Swim Outdoors		Coldwater Fishing		Warmwater Fishing		Motor Boating		Visit Waterbody		Swimming Outdoor Pool	
Phosphorus	-2.048***	-2.099**	-1.957***	-1.603*	-1.764***	-3.496***	-2.684***	-3.035**	-1.702**	-0.535	-0.779*	-0.122	-1.340**	-1.053	0.964*	1.902**
(mg/l)	(0.782)	(1.012)	(0.702)	(0.820)	(0.581)	(1.010)	(0.942)	(1.258)	(0.748)	(0.740)	(0.433)	(0.677)	(0.560)	(0.778)	(0.508)	(0.875)
Observations	33,857	10,936	52,358	16,682	68,695	17,206	65,422	17,150	59,767	17,146	68,676	17,187	54,602	16,785	38,353	17,210
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Panel B: Recreation Activities - Part 2

	Bicycling		Hiking		Snow and Ice Activities		Outdoor Team Sports		Horseback Riding		Picnicking		Visit Nature Center		Visit Prehistoric Site	
Phosphorus	0.325	0.687	-0.741*	0.416	-0.193	0.421	0.280	0.259	-0.800	-0.181	-0.225	0.160	0.180	-0.192	-0.421	-0.300
(mg/l)	(0.360)	(0.834)	(0.388)	(0.725)	(0.375)	(0.732)	(0.490)	(0.780)	(0.490)	(0.905)	(0.309)	(0.660)	(0.370)	(0.741)	(0.460)	(0.750)
Observations	62,545	13,191	62,465	17,210	61,477	17,210	40,497	17,210	60,955	17,210	62,148	17,210	62,147	17,210	59,112	17,210
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Panel C: Recreation Activities - Part 3

	Visit Historic Site		Walk for Pleasure		Visit Wilderness		Gather Mushrooms		Hunt		Mountain Biking		Backpacking		Developed Camping	
Phosphorus	0.465	1.063	-0.134	0.0829	-1.130***	-0.318	-1.605**	-1.139	-1.297*	-0.335	-0.0869	1.159	-1.639***	-0.597	-0.204	0.269
(mg/l)	(0.398)	(0.792)	(0.415)	(0.891)	(0.407)	(0.712)	(0.625)	(0.763)	(0.747)	(0.947)	(0.397)	(1.282)	(0.604)	(1.238)	(0.345)	(0.731)
Observations	55,262	17,210	63,186	17,234	62,392	17,234	59,998	17,210	71,849	17,214	60,580	7,333	63,200	17,210	62,544	17,210
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Panel D: Recreation Activities - Part 4

	Primitive Camping		View Birds		View Fish		View Wildlife		View Flowers		View Scenery		Drive for Pleasure		Drive Off-road	
Phosphorus	-0.906*	-0.661	-0.166	0.860	-1.088**	-0.938	-0.829*	-0.0870	-0.446	-0.357	-0.524	-0.344	-1.188**	-0.837	-1.433**	-1.874*
(mg/l)	(0.546)	(0.866)	(0.313)	(0.743)	(0.509)	(0.852)	(0.464)	(0.676)	(0.350)	(0.755)	(0.337)	(0.752)	(0.483)	(0.722)	(0.600)	(0.958)
Observations	58,249	17,210	69,284	17,210	69,650	17,210	68,795	17,210	68,943	17,210	68,995	17,210	56,081	17,210	57,782	17,210
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.

Notes: Dependent variable is participation in specified activity. All estimates are control function estimates using Stata's IV Probit command. "Normal" controls include state fixed effects and all standard controls used throughout the paper. "+ Activ" controls also include participation in non-water based activities. See text for more details. Robust standard errors, clustered at the county level, in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

APPENDIX TABLE 3 – EXTENSIVE MARGIN AND PLACEBO TESTS (US EPA NATIONAL LAKES – SURVEY WEIGHTS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Panel A: Recreation Activities - Part 1																		
	Water-based Recreation		Water-based Recreation		Swim Outdoors		Coldwater Fishing		Warmwater Fishing		Motor Boating		Visit Waterbody		Swimming Outdoor Pool		Bicycling	
Phosphorus	-3.955**	-4.124	-3.701***	-4.183**	-4.277***	-5.554**	-3.652**	-4.073*	-2.439*	-1.075	-2.115**	-0.403	-2.386*	-1.699	0.986	3.585*	-0.335	-0.0115
(mg/l)	(1.802)	(2.565)	(1.423)	(2.046)	(1.498)	(2.181)	(1.468)	(2.460)	(1.390)	(1.819)	(1.056)	(1.680)	(1.244)	(1.807)	(1.290)	(2.171)	(0.948)	(2.030)
Observations	29,387	9,574	45,412	14,521	59,541	14,961	56,729	14,916	51,852	14,916	59,531	14,948	47,351	14,600	33,320	14,964	54,198	11,427
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.
Panel B: Recreation Activities - Part 2																		
	Hiking		Snow and Ice Activities		Outdoor Team Sports		Horseback Riding		Picnicking		Visit Nature Center		Visit Prehistoric Site		Visit Historic Site		Walk for Pleasure	
Phosphorus	-1.018	-0.772	-0.911	0.947	1.542	2.262	-0.673	-0.0446	-0.522	1.496	-0.950	-2.142	-0.267	2.081	1.236	0.00530	-0.511	0.719
(mg/l)	(1.122)	(1.655)	(1.020)	(1.766)	(1.297)	(2.052)	(1.152)	(2.245)	(0.896)	(1.614)	(1.077)	(1.853)	(1.229)	(2.020)	(1.298)	(1.926)	(1.336)	(2.086)
Observations	54,118	14,964	53,311	14,964	35,033	14,964	52,800	14,964	53,849	14,964	53,850	14,964	51,206	14,964	47,920	14,964	54,763	14,985
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.
Panel C: Recreation Activities - Part 3																		
	Visit Wilderness		Gather Mushrooms		Hunt		Mountain Biking		Backpacking		Developed Camping		Primitive Camping		View Birds		View Fish	
Phosphorus	-0.808	-0.396	-3.282**	-2.661	-2.144	1.562	-0.695	5.016*	-0.898	0.378	-1.208	-1.446	-0.101	-2.738	0.543	2.722	0.196	1.256
(mg/l)	(0.982)	(1.750)	(1.311)	(1.701)	(1.547)	(2.216)	(1.014)	(2.925)	(1.317)	(2.826)	(0.948)	(1.751)	(1.246)	(2.043)	(1.155)	(1.706)	(1.027)	(2.092)
Observations	54,056	14,985	52,001	14,964	62,280	14,967	52,483	6,439	54,793	14,964	54,181	14,964	50,488	14,964	60,061	14,964	60,356	14,964
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.
Panel D: Recreation Activities - Part 4																		
	View Wildlife		View Flowers		View Scenery		Drive for Pleasure		Drive Off-road									
Phosphorus	-1.499	-0.166	-0.531	-0.267	-1.199	1.046	-1.195	-1.407	-3.373**	-4.381**								
(mg/l)	(1.131)	(1.909)	(0.943)	(1.778)	(0.948)	(1.756)	(1.015)	(1.750)	(1.488)	(2.121)								
Observations	59,614	14,964	59,748	14,964	59,784	14,964	48,582	14,964	50,099	14,964								
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.								

Notes: Dependent variable is yes/no for participation in specified activity type. All reported estimates are control function estimates produced using Stata's IV Probit command. All regressions include state fixed effects and controls for metropolitan and micropolitan statistical areas, coastal and Great Lakes shorelines, 30-year climate normals for temperature and precipitation and temperature squared, reservoir county, race, gender, age, income, education, elevation, and predominant land cover type. "+ Activ" controls include all the above controls plus controls for participation in all other individual non-water based recreational activities. Instrumental variables are concentrations of phosphorus at >0-200 miles upstream. For parameter estimates, robust standard errors, clustered at the county level, in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

APPENDIX TABLE 4 – EXTENSIVE MARGIN AND PLACEBO TESTS (USGS SPARROW – SURVEY WEIGHTS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Panel A: Recreation Activities - Part 1																		
	Water-based Recreation		Water-based Recreation		Swim Outdoors		Coldwater Fishing		Warmwater Fishing		Motor Boating		Visit Waterbody		Swimming Outdoor Pool		Bicycling	
Phosphorus	-2.906	-2.689*	-2.318	-1.664*	-2.309*	-3.609**	-2.683	-2.801	-1.591	-0.407	-1.294	-0.267	-1.265	-0.312	0.315	1.384	-0.251	-0.220
(mg/l)	(1.976)	(1.462)	(1.511)	(0.973)	(1.208)	(1.428)	(1.714)	(1.817)	(1.280)	(0.881)	(1.039)	(0.879)	(0.899)	(0.865)	(0.722)	(1.078)	(0.563)	(0.973)
Observations	33,857	10,936	52,358	16,682	68,695	17,206	65,422	17,150	59,767	17,146	68,676	17,187	54,602	16,785	38,353	17,210	62,545	13,191
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.
Panel B: Recreation Activities - Part 2																		
	Hiking		Snow and Ice Activities		Outdoor Team Sports		Horseback Riding		Picnicking		Visit Nature Center		Visit Prehistoric Site		Visit Historic Site		Walk for Pleasure	
Phosphorus	-0.315	1.691	-0.253	0.652	1.007	1.441	-0.361	1.509	-0.504	0.339	-0.428	-0.293	-0.198	0.207	-0.0137	0.306	-1.209	-0.456
(mg/l)	(0.594)	(1.489)	(0.537)	(0.986)	(0.943)	(1.375)	(0.638)	(1.784)	(0.550)	(0.887)	(0.600)	(1.193)	(0.644)	(0.940)	(0.559)	(1.009)	(0.991)	(1.042)
Observations	62,465	17,210	61,477	17,210	40,497	17,210	60,955	17,210	62,148	17,210	62,147	17,210	59,112	17,210	55,262	17,210	63,186	17,234
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.
Panel C: Recreation Activities - Part 3																		
	Visit Wilderness		Gather Mushrooms		Hunt		Mountain Biking		Backpacking		Developed Camping		Primitive Camping		View Birds		View Fish	
Phosphorus	-1.176	-1.197	-2.089	-1.526	-1.859	-0.336	-0.330	2.891	-0.907	1.749	-0.370	0.297	-0.672	-1.545	0.0888	1.416	-0.782	-0.149
(mg/l)	(0.779)	(1.096)	(1.303)	(1.129)	(1.645)	(1.260)	(0.472)	(2.220)	(0.855)	(1.581)	(0.540)	(0.883)	(0.939)	(1.616)	(0.487)	(1.124)	(0.966)	(1.157)
Observations	62,392	17,234	59,998	17,210	71,849	17,214	60,580	7,333	63,200	17,210	62,544	17,210	58,249	17,210	69,284	17,210	69,650	17,210
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.
Panel D: Recreation Activities - Part 4																		
	View Wildlife		View Flowers		View Scenery		Drive for Pleasure		Drive Off-road									
Phosphorus	-1.318	0.386	-0.725	-1.003	-0.836	-0.557	-1.339	-0.951	-1.677*	-1.535								
(mg/l)	(1.118)	(0.928)	(0.752)	(1.038)	(0.701)	(0.922)	(0.888)	(0.944)	(0.947)	(1.075)								
Observations	68,795	17,210	68,943	17,210	68,995	17,210	56,081	17,210	57,782	17,210								
Controls	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.	Normal	+ Activ.								

Notes: Dependent variable is yes/no for participation in specified activity type. All reported estimates are control function estimates produced using Stata's IV Probit command. All regressions include state fixed effects and controls for metropolitan and micropolitan statistical areas, coastal and Great Lakes shorelines, 30-year climate normals for temperature and precipitation and temperature squared, reservoir county, race, gender, age, income, education, elevation, and predominant land cover type. "+ Activ" controls include all the above controls plus controls for participation in all other individual non-water based recreational activities. Instrumental variables are concentrations of phosphorus at >0-200 miles upstream. For parameter estimates, robust standard errors, clustered at the county level, in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

APPENDIX TABLE 5 – THE EFFECTS OF SOIL LOSS ON RECREATIONAL USE (INTENSIVE MARGIN)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Negative Binomial and GMM Count Models						
Soil Loss (tons/acre/yr)	-0.0446***	-0.0392**	-0.00865	-0.176***	-0.193***	-0.429**
	(0.0156)	(0.0174)	(0.0196)	(0.0366)	(0.0489)	(0.171)
Controls	None	Partial	Full	None	Partial	Full
Instruments	None	None	None	>0 - 200	>0 - 200	>0 - 200
Observations	19,544	19,544	19,544	19,544	19,544	19,544
Panel B: Semi-log Models						
Soil Loss (tons/acre/yr)	-0.0542***	-0.0395***	-0.00901	-0.221***	-0.217***	-0.545***
	(0.0123)	(0.0139)	(0.0141)	(0.0364)	(0.0463)	(0.207)
Controls	None	Partial	Full	None	Partial	Full
Instruments	None	None	None	>0 - 200	>0 - 200	>0 - 200
Observations	19,544	19,544	19,544	19,544	19,544	19,544
Kleibergen-Paap				106.498	75.711	11.697
rk Wald F Stat				(F = 16.38)	(F = 16.38)	(F = 16.38)

Notes: Dependent variable in Panel A is the number of water-based recreational days taken in a year conditional on taking at least one trip. Dependent variable in Panel B is ln(recreational days + 1) for the same sample. Column (5) includes controls for income, age, gender, education, and forest land cover. Column (6) includes state fixed effects and controls for metropolitan and micropolitan statistical areas, coastal and Great Lakes shorelines, 30-year climate normals for temperature and precipitation and temperature squared, indicator for counties with a lake, race, gender, age, income, education, elevation, and predominant land cover type. Instrumental variables are concentrations of phosphorus (columns (4) to (6)) at >0 - 200 miles upstream. For parameter estimates, robust standard errors, clustered at the county level, in parentheses. For Kleibergen-Paap rk Wald F Stat, Stock-Yogo weak identification critical value for 10% maximal IV size in parentheses. Critical value for 15% maximal instrumental variable size equals 8.96 for specifications (4) - (6).

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.