# CAFOs and surface water quality: Evidence from Wisconsin

Zach Raff [a]
University of Wisconsin-Stout

Andrew Meyer [b]
Marquette University

January 25, 2020

**Abstract**

Concentrated Animal Feeding Operations (CAFOs) – animal feeding operations with over 1,000 animal units in confined spaces – have proliferated over the past 30 years in the United States. CAFOs provide operational cost savings, but higher animal concentrations in confined spaces can generate external costs, e.g., non-point source water pollution. In this study, we improve on previous research designs to estimate the relationship between the growth in CAFOs and surface water quality using longitudinal data on a large spatial scale. We use a panel dataset from 1995-2017 that links CAFO intensity with nearby surface water quality readings in Wisconsin to perform our analysis. Leveraging variation in CAFO intensity within hydrological regions over time, we find that increasing CAFO intensity increases the levels of nutrients, specifically total phosphorus and ammonia, in surface water; adding one CAFO to a Hydrologic Unit Code-8 (HUC8) region leads to a 1.7% increase in total phosphorus levels and a 2.7% increase in ammonia levels, relative to sample mean levels. As an important contribution of our work, we use these results to calculate the external costs of surface water quality damages from CAFOs in Wisconsin. Our results imply that the marginal CAFO in Wisconsin produces non-market surface water quality damages of at least $203,541 per year.

**JEL Classification:** H23, Q15, Q53

**Keywords:** ammonia, Concentrated Animal Feeding Operations, non-point source pollution, surface water quality, total phosphorus

[a] Corresponding Author: University of Wisconsin-Stout, 712 Broadway St. S, Menomonie, WI 54751. Telephone: (715) 232-2493. Email: raffz@uwstout.edu

[b] Economics Department, Marquette University, 1225 W. Wisconsin Ave., Milwaukee, WI 53233. Telephone: (414) 288-5489. Email: andrew.g.meyer@marquette.edu

# CAFOs and surface water quality: Evidence from Wisconsin

January 25, 2021

**Abstract**

Concentrated Animal Feeding Operations (CAFOs) – animal feeding operations with over 1,000 animal units in confined spaces – have proliferated over the past 30 years in the United States. CAFOs provide operational cost savings, but higher animal concentrations in confined spaces can generate external costs, e.g., non-point source water pollution. In this study, we improve on previous research designs to estimate the relationship between the growth in CAFOs and surface water quality using longitudinal data on a large spatial scale. We use a panel dataset from 1995-2017 that links CAFO intensity with nearby surface water quality readings in Wisconsin to perform our analysis. Leveraging variation in CAFO intensity within hydrological regions over time, we find that increasing CAFO intensity increases the levels of nutrients, specifically total phosphorus and ammonia, in surface water; adding one CAFO to a Hydrologic Unit Code-8 (HUC8) region leads to a 1.7% increase in total phosphorus levels and a 2.7% increase in ammonia levels, relative to sample mean levels. As an important contribution of our work, we use these results to calculate the external costs of surface water quality damages from CAFOs in Wisconsin. Our results imply that the marginal CAFO in Wisconsin produces non-market surface water quality damages of at least $203,541 per year.

**JEL Classification:** H23, Q15, Q53

**Keywords:** ammonia, Concentrated Animal Feeding Operations, non-point source pollution, surface water quality, total phosphorus

Over the past several decades, technological change and economies of scale have led to significant changes in livestock operations in the United States (Osterberg and Wallinga 2004; Sneeringer 2009). The industry has steadily moved from small, individual operations to larger and more concentrated operations. U.S. Environmental Protection Agency (EPA) estimates show that the number of animal feeding operations (AFOs) has declined considerably over this time period, while the number of animals at each operation has increased (EPA 2001; Copeland 2010). In U.S. dairy production for example, the midpoint herd size – where half of all cows are in larger herds and half are in smaller herds – increased from 101 cows in 1992 to 900 cows by 2012 (MacDonald and Newton 2014).

Concentrated Animal Feeding Operations (CAFOs) are large AFOs that do not store or grow crops on any part of the lot/facility and have over 1,000 animal units[1] onsite which are covered for at least 45 days per year (EPA 2004). However, CAFOs can impose more substantial external costs on outside parties than small-scale AFOs (EPA 2002a,b; Copeland 2010). Specifically, CAFOs produce large amounts of animal waste in confined spaces (Hribar 2010), which must be stored onsite or spread onto agricultural fields. The waste can then leach into groundwater or, when precipitation events occur, run off the land into surface waterbodies (EPA 2004; Burkholder et al. 2006). Importantly, the runoff is considered a non-point source pollutant, which is largely exempt from Clean Water Act (CWA) regulation (Olmstead 2009). Indeed, existing ecological studies document a correlation between CAFOs and lower ambient water quality at the individual waterbody level. For example, water samples near CAFOs often contain elevated levels of pollutants such as fecal coliform and nutrients such as nitrates (Weldon and Hornbuckle 2006; Mallin et al. 2015). However, previous studies do not establish or quantify a link between CAFOs and surface water pollution on a large spatial scale; nearly all studies examine water quality near

individual CAFOs without a suitable comparison group. Additionally, government reports that link CAFOs and surface water quality (e.g., EPA 2000b; EPA 2002a,b) rely on ecological modeling to predict these effects, rather than observational data to examine the effects, *ex post*.

In this paper, we improve upon and add to previous analyses by identifying a link between the size and scope of CAFOs and ambient surface water quality on a large spatial scale. To identify this link, we use longitudinal data on of the size and location of the universe of permitted CAFOs in Wisconsin and proximate water quality data. Specifically, we use a difference-in-differences (DD) framework to examine the effects of CAFOs on nearby ambient concentrations of total phosphorus and ammonia (jointly hereafter "nutrients").[2] This approach allows us to compare changes in nutrient concentrations among watersheds with large expansions in CAFOs with changes in nutrient concentrations in control watersheds with no or low expansion in CAFOs.[3] We also monetize these effects to calculate the loss in non-market water quality benefits, i.e., damages, associated with CAFOs.

The connection between CAFOs and water quality has been well-studied, particularly in the ecological literature. The animal waste produced by CAFOs is not treated like that of humans (EPA 2001), so the excessive nutrients present in animal waste can increase eutrophication in surface waterbodies via discharge events (Hooda et al. 2000; Weldon and Hornbuckle 2006). As a result, EPA regulates CAFOs under the CWA more strictly than other AFOs. There is evidence that regulation decreases surface water pollution from CAFOs (Chen et al. 2019), but may lead to regulatory avoidance by AFOs, i.e., bunching at the size threshold (Sneeringer and Key 2011). Manure from CAFOs also contains pathogens that can be harmful to humans, e.g., E. coli; these pathogens can live in surface water and impair its quality (Spellman and Whiting 2007; Heaney et al. 2015). Hormones and heavy metals injected into animals at CAFOs and present in manure can

also impair water quality (Barker and Zublena 1995; Boxall et al. 2003). Hooda et al. (2000) and Hu et al. (2017) provide thorough reviews of the literature connecting CAFOs and water pollution.

Another relevant literature focuses on other external costs associated with CAFOs. First, the large amounts of animal waste in confined spaces at large AFOs produce sulfur-related air pollution (Sneeringer 2010). Next, a sizable epidemiological literature shows that proximity to large AFOs negatively affects human health, both of the employees at AFOs and of the general public (Sigurdarsen and Kline 2005; Radon et al. 2007; Greger and Koneswaran 2010; Ramos et al. 2018).[4] In the economics literature, Sneeringer (2009) uses longitudinal data to find a causal relationship between human health and animal concentrations. The author finds that larger concentrations of animals at feeding operations increase infant mortality. Finally, CAFOs are associated with external costs in the housing market through property value decreases (Herriges et al. 2005; Ready and Abdalla 2005; Isakson and Ecker 2008). Importantly, the literature lacks any study that uses longitudinal data to identify the effects of the growth of CAFOs on surface water quality on a large spatial scale.

This study adds to these literatures in several ways. First, we provide more credible *ex post* estimates of the effects of CAFOs on ambient surface water quality than previous studies. The literature, including several EPA reports, lacks identified econometric studies that connect CAFOs and surface water quality on a large spatial scale. Instead, prior research uses ecological modeling and case studies to document this association. We rely on plausibly exogenous variation in the timing of CAFO expansion within a watershed to address these potential differences between nutrient levels in areas with differing CAFO intensity. Second, we add to a growing literature that examines non-point source pollution using a rich water quality dataset on a large spatial scale (Grenestam and Nordin 2018; Meyer 2018; Behrer et al. 2019; Chen et al. 2019; Grant and

Langpap 2019; Keiser and Shapiro 2019; Paudel and Crago 2020).[5] The literature is rife with studies that examine point source pollution,[6] as discharge data are easier to identify from point sources than non-point sources. However, quantifying non-point discharges is difficult and the literature on the longitudinal effects of various sources of non-point source pollution is scarce. Our use of water quality measures as an outcome helps to close this gap in the literature. Third, our study is important from an economic perspective. Elevated levels of nutrients in surface waterbodies can lead to eutrophication and algal blooms, which are costly.[7] Previous work estimates the total annual costs associated with damages from freshwater eutrophication in the United States to be $2.2 billion (Dodds et al. 2009). Additionally, algal blooms are becoming more prevalent with climate change (Wells et al. 2015) so this issue is likely to be even more important in the future.[8] Finally, we monetize the estimated non-market water quality damages from CAFOs, so our study adds to the non-market valuation literature and is useful for policymakers.

To address our research question, we create a panel dataset that links CAFO presence and intensity with nearby surface water quality readings of total phosphorus and ammonia for the state of Wisconsin between 1995 and 2017. We use Wisconsin's hydrological network to identify the level of CAFO exposure at each water quality monitoring location. The USGS methodology that identifies hydrological networks divides and sub-divides the United States into successively smaller networks nested within larger networks. USGS gives each area a hydrological unit code (HUC) with two (largest areas) to twelve (smallest sub-basins) digits. For our analysis, we use the eight-digit HUCs (HUC8), which are called "cataloging units" or "watersheds" and correspond to a drainage basin or distinct hydrologic feature; HUC8s also follow natural boundaries for surface water flow (Seaber et al. 1987).[9] Our primary empirical strategy leverages within HUC8 variation in CAFO presence and intensity over time to estimate the effects on surface water quality from

these operations. We find that increasing CAFO presence and intensity leads to significantly higher levels of nutrients in surface water readings. Estimation results show that adding one CAFO to a HUC8 region leads to a 1.7% increase in average total phosphorus levels and a 2.7% increase in average ammonia levels, while controlling for several time-varying characteristics such as precipitation, land use patterns, and land cover. We then monetize these effects by applying benefit transfer functions to estimated changes in a water quality index. We find that the average Wisconsin household's willingness to pay (WTP) to eliminate the water quality damages produced by a marginal CAFO in each HUC8 is roughly $3-$12 per year. Collectively, we estimate Wisconsin's total annual willingness to pay to avoid the water quality damages from one CAFO would be in the range of $203,541-$821,533, which compares to approximately 4-17% of the associated sales revenue of an average dairy CAFO in Wisconsin. Policymakers can use this information about lost non-market water quality benefits to quantify how much external damage should be priced into production decisions.

The paper proceeds as follows. The next section provides background information on CAFOs and non-point source pollution. We then describe the data, including statistical summaries. The following section provides the empirical analysis, which discusses the empirical results and provides robustness checks. The final section concludes.

## Background

In this section, we first discuss the specifics of CAFOs, water pollution stemming from CAFOs, and the regulatory efforts aimed at preventing this externality. We then discuss our selection of research site.

### *CAFOs and water pollution*

The distribution of animals on AFOs has trended away from traditional, small-scale AFOs over the past several decades; this is evidenced by the decline in the number of livestock farms at a time when livestock inventory has remained constant (Sneeringer 2009; Copeland 2010) or, in Wisconsin's case, declined (USDA 2020). In Wisconsin, this trend has resulted in the rapid growth in the number of CAFOs in the state. Figure 1 shows Wisconsin CAFO locations in 1994 and 2017, respectively, which confirms this growth. CAFOs still comprise a relatively small percentage of the total AFOs in Wisconsin, but a growing percentage of livestock are located on CAFOs. In 2019, about 3.5% of all dairy operations in Wisconsin were CAFOs but nearly 25% of dairy cows were located on CAFOs (Cushman 2019).

Growth in CAFOs can pose a threat to human health and the environment because of the extreme concentration of animal waste at CAFOs. In 1997, CAFOs in the United States produced over 291 billion pounds of wet manure (EPA 2001).[10] According to the U.S. Government Accountability Office (GAO), a dairy farm with 1,200 dairy cows can produce almost 30,500 tons of manure per year. This amount of animal waste is roughly equivalent to the amount of annual human sanitary waste produced by a U.S. city with 46,000 people (GAO 2008). Manure is most often dealt with at CAFOs in one of two ways. First, manure can be stored onsite, typically within surface "lagoons", in large piles, or under buildings or tarps. Second, manure can be spread onto farmland, often at agronomically inappropriate rates (Osterberg and Wallinga 2004; Hu et al. 2017). The potential for water pollution is significant for either storage method. Regarding the former, manure lagoons are typically insecure and do not contain linings or retaining walls (Hribar 2010). Thus, manure can leach into groundwater or leak from piles and lagoons, especially during precipitation events. Regarding the latter, when manure is overspread on surrounding farmland the ground is unable to fully absorb the excessive nutrients present in manure (EPA 2001; Kellogg et

al. 2014). When this scenario occurs, precipitation events and melting snow carry the manure to surface waterbodies. This form of water pollution is referred to as non-point source pollution and is largely exempt from CWA regulation.[11] There is considerable heterogeneity in the regulation of non-point source pollution at the state level. In Wisconsin for example, all agricultural operators are subjected to broad non-point source policies, which include the adoption of a nutrient management plan. However, non-point source pollution remains difficult to regulate and most states focus on the adoption of Best Management Practices (BMP) or Total Maximum Daily Loads for nutrient loadings into surface waterbodies.

In 2003, EPA updated the permitting program of the CWA, the National Pollutant Discharge Elimination System (NPDES), to require CAFOs to obtain permits to operate and to develop nutrient management plans to control animal waste (Sneeringer and Key 2011; Chen et al. 2019).[12] Permitted CAFOs are then considered point source dischargers and are subjected to discharge limits (EPA 2004). The implementation of the NPDES program and the issuance of permits are administered by individual states that have been granted primary authority to operate the NPDES program, which include 46 of the 50 states.[13] In Wisconsin, the state legislature implemented two administrative codes during our sample period that govern agricultural water quality performance standards. Administrative code NR151 sets agricultural performance standards for the state and administrative code ATCP50 guides how livestock operators meet the performance standards, e.g., nutrient management plans, which all CAFOs must submit.

Regardless of this regulatory environment, the control of wastewater discharges and non-point source pollution from CAFOs remains insufficient. First, almost all CAFOs are considered "minor" dischargers (EPA 2020), which are not required to systematically report discharges to EPA or authorized state agencies (Raff and Earnhart 2019). Indeed, direct discharge data from

CAFOs are virtually non-existent in EPA's point source discharge data system, the DMR Pollutant Loading Tool (EPA 2020). It is therefore difficult to determine the NPDES compliance status of CAFOs. Second, there is considerable heterogeneity between states in the regulation of discharges from CAFOs. Many states, e.g., North Carolina, Arkansas, have large numbers of unpermitted CAFOs, which is likely the result of the differences in NPDES implementation and the definition of potential dischargers among states (GAO 2003) and regulatory avoidance by large AFOs (Sneeringer and Key 2011). Third, NPDES permits do not regulate the amount of animal waste from CAFOs that is spread onto fields. Non-point source pollution is still possible even if the CAFOs secure the proper NPDES permits. And finally, nutrient management plans are often insufficient for the control of animal waste produced at CAFOs. The nutrient management requirements of EPA's 2003 CAFO rule did not significantly affect surface water quality in Iowa after the rule's implementation (Chen et al. 2019). Collectively, there exist shortcomings in the regulation of point source and non-point source pollution from CAFOs in the United States. Water pollution from these sources remains a large concern for policymakers.

*Selection of research site*

A nationwide study of the impact of CAFOs on surface water quality is not feasible because data on the numbers and locations of CAFOs are sparse. EPA has reported national summaries of the total number of estimated CAFOs and permitted facilities by state since 2011. The EPA summaries show that many states have issued permits to only a small percentage of CAFOs, which results in little information available about CAFO location and size for these states. For example, the 2017 summary document indicates that Idaho has issued permits for 0 of the state's 365 CAFOs, Illinois has issued permits for 32 of its 297 CAFOs, and New York has issued permits for just 21 of its

571 CAFOs.[14] Earlier data on CAFOs are more difficult to locate. A 2008 GAO report states that "no federal agency collects accurate and consistent data on the number, size, and location of CAFOs" (GAO 2008).

As a result, we focus our analysis on one state: Wisconsin. Our focus on a single state is like other studies that examine non-point source pollution (e.g., Meyer 2018; Chen et al. 2019) and is typical of policy analyses in other literatures. Wisconsin represents a useful state to study for several reasons. First, Wisconsin collects much better data on CAFO size and location over time than other states. EPA's 2017 CAFO summary document shows that Wisconsin has issued permits to 295 of its 315 CAFOs. More historically, EPA estimated in 2000 the number of CAFOs in each state as part of its proposed NPDES permitting rule (EPA 2000b). The report estimated that there were 141 CAFOs located in Wisconsin at that time. In the same year, the Wisconsin Department of Natural Resources (WDNR) had issued permits to 110 CAFO locations. Thus, we are confident that the historical data that we have acquired represent the majority of the CAFO universe in Wisconsin during our sample period.

In addition to implementing the NPDES rule at nearly all CAFOs, Wisconsin is also a leader in the control of wastewater from CAFOs in other ways. The NR151 and ATCP50 rules in the state place further restrictions on CAFOs, which do not exist in many states. Additionally, data on CWA Section 319 grants administered to the state indicate that most non-point source pollution grants from the program target BMPs for confined livestock waste control (GRTS 2020). As a result, the estimated effects of CAFOs on surface water quality in Wisconsin can be reasonably considered a conservative baseline comparison for states with little CAFO regulation.

Second, Wisconsin contains a tremendous amount of fresh, surface water (WDNR 2011). Importantly, the state also monitors its large amounts of surface water at levels well above the

national average, which increases the statistical power of our analysis. We use the Water Quality Portal to determine the count of surface waterbody locations that each state sampled in 2017.[15] Wisconsin sampled 6.72 surface waterbody locations per 100 square miles in 2017. For the rest of the country, the average state sampled 1.35 surface waterbody locations per 100 square miles.[16]

Third, Wisconsin is typical of other corn belt states that contain considerable CAFO presence (EPA reports that the corn belt contains 8,083 of the country's 19,961 CAFOs, or over 40%), which lends support to the external validity of this study to states in this region.[17] The assimilative capacity for manure phosphorus and nitrogen of Wisconsin's soil is like that of other corn belt states. According to the U.S. Department of Agriculture (USDA), nearly all counties in Wisconsin and other states in this region have cropland and pastureland with the assimilative capacity to process manure phosphorus and nitrogen of 5 to 10 million pounds per county and 25 to 40 million pounds per county, respectively (USDA 2000). The types of livestock at Wisconsin CAFOs are also like those of other corn belt states. In Wisconsin, 97% of CAFOs contain cattle or swine, while only 3% are poultry operations. The rest of the corn belt mirrors this composition, as it contains only 7% of the poultry CAFOs in the country (USDA 2002). The similarity in operations is important because poultry waste contains three to five times more phosphate per pound than waste from dairy and beef cows and swine (Madison et al. 1995). Collectively, livestock manure that is discharged from CAFOs or spread onto fields in Wisconsin creates point source and non-point source pollution similarly to water pollution from CAFOs in other corn belt states.

The preceding points notwithstanding, we acknowledge that a study of one state may not generalize to other regions of the country. However, our results can be useful for policymakers in other states outside of the corn belt. For example, the southeastern portion of the United States contains a large share of poultry CAFOs and has soil with poor assimilative capacity for manure

phosphorus and nitrogen (USDA 2000). Counties in this section of the country are therefore more conducive to non-point source pollution from CAFOs that spread waste with higher phosphate concentrations than Wisconsin and other corn belt states.

## Data

This section describes the data used in our empirical analysis. We first provide an overview of the process we use to geographically and temporally match CAFO location and size with water quality measures. We then describe our data sources in detail, discuss summary statistics, and examine trends in CAFOs and surface water nutrient concentrations over the study period.

### *Construction of panel*

To construct our panel, we must first spatially link water quality data and CAFO intensity data. For our purposes, there exists a tradeoff when deciding which HUC level to use in the analysis. We want to use a small enough geographic area so that CAFOs could feasibly affect water quality readings. But using too small of a geographic area can be problematic because animal waste could potentially be spread outside of the small area.

We use HUC8 level spatial delimiters in our analysis for two reasons. First, it is imperative for our analysis that the manure from CAFOs remains in the same HUC region as the CAFO itself; thus, using too small a HUC region is problematic for our identification. When CAFOs store waste onsite, geolocating the manure from a CAFO to a HUC8 region is straightforward. However, manure from CAFOs is often spread onto nearby fields. It is therefore important to understand whether this manure spreading likely occurs in the same HUC8 region as the CAFO itself. We do not have data on where the CAFOs in our sample spread their manure. Existing literature suggests that manure is typically transported short distances. Ali et al. (2012) provide evidence that farmers

in two similar states to Wisconsin (Iowa and Missouri) transport manure short distances. The authors find that the average maximum manure transport distance is 2.35 miles for dairy cows, 2.78 miles for beef cows, 2.95 miles for swine less than 55 pounds, 4.25 miles for swine greater than 55 pounds, 14.78 miles for broilers, and 13.66 miles for turkeys.[18] The vast majority of CAFOs in our sample (greater than 90%) are dairy cows so the manure for most sample CAFOs is likely spread within a few miles of their location. Thus, we are confident that the manure from CAFOs in our sample is spread inside the geolocated HUC8 region. Second, the growing literature on non-point source pollution primarily uses HUC8 level measures (Grant and Grooms 2017; Grant and Langpap 2019; Paudel and Crago 2020). We follow the lead of these studies and use HUC8 level measures as well.[19]

We construct our panel by matching the HUC8 region of each water quality monitoring location to HUC8 level CAFO intensity measures. As a specific example, we pair the nutrient readings from a monitoring location of a surface waterbody in a specific HUC8 on April 15 of 2001 with the CAFO intensity measures from that same HUC8 region on April 15 of 2001. Thus, the CAFO variable measures the level of CAFO exposure for a given water quality monitor on a given day.[20]

We then merge time-varying controls to our panel. For several controls, the data are only available at the county-year level. We spatially transform these county level measures to HUC8 level measures because multiple HUC8s can span a single county. To do so, we weight each county level measure by the area of the county that lies within each HUC8 region. We then aggregate the oftentimes multiple weighted measures to create a single HUC8 level measure for these control variables. Finally, we standardize these values to account for HUC8 regions that span into other states. We temporally convert the control measures to daily or monthly observations via linear

extrapolation and interpolation.[21]

*CAFO intensity*

We obtain CAFO data through an open records request with the WDNR, which results in information on the historical population of CAFO permits in Wisconsin dating back to 1990. Figure 2 shows the variation over time in the number of CAFOs in Wisconsin from the beginning of 1995 to 2017. Prior to 1995, there were only six permitted CAFOs throughout the state.[22] CAFO presence in Wisconsin has steadily increased during our sample period; there were 274 permitted CAFOs in Wisconsin by 2017.[23] Figure 2 also shows that the number of total animal units at CAFOs has closely trended with the number of CAFOs in the state, suggesting that the average CAFO size has not changed significantly since 1995.[24] However, CAFO growth has not been uniform throughout Wisconsin. Figure 1 shows the geographic distribution of CAFOs in Wisconsin in 1994 and 2017, respectively. Some HUC8 regions have seen significant CAFO growth over time, some experienced little CAFO growth, and some HUC8s have zero CAFOs for the entirety of our panel.

*Water quality*

We collect water quality data from three primary sources: U.S. Geological Survey (USGS) National Water Information System (NWIS), EPA Storage and Retrieval (STORET), and USGS Biodata.[25] In our sample, roughly 70% of observations come from STORET. We are interested in non-point source pollution of excess nutrients from CAFOs for our primary analysis, so we collect data for total phosphorus and ammonia readings. We also collect data on pesticide concentrations for placebo tests and total suspended solids (TSS) for our valuation exercise. We focus on total

phosphorus and ammonia because these are two of the most common pollutants resulting from animal waste, as highlighted by EPA and WDNR (EPA 2000b; EPA 2002a,b; EPA 2004; WDNR 2011). Finally, we only collect water quality readings for surface waterbodies, e.g., lakes, rivers; we do not collect data for non-surface waterbodies, e.g., wells, ditches.

Each water quality observation represents a given monitoring location, in a HUC8 region, on a specific day. We retain all observations on nutrient levels that match to a HUC8 region in Wisconsin. There exist some extremely high readings on nutrient levels for a very small subset of the water quality data. Therefore, we follow Keiser and Shapiro (2019) and winsorize data at the 99% level.[26]

Previous studies in the in the ecological literature have used these water quality measures to examine individual surface waterbodies (e.g., Landon et al. 2014). We are aware of only three published studies in the economics literature that use these water quality measures to answer economic questions over a large spatial scale (Grant and Langpap 2019; Keiser and Shapiro 2019; Paudel and Crago 2020). However, there exist several working papers that use these data in this way (Chen et al. 2019; Behrer et al. 2019). Thus, our study and use of these data serve as a contribution to the growing economic literature on non-point source water pollution.

*Control variables*

Various time-varying factors may affect the concentrations of nutrients in surface waterbodies. Adding these factors to our regression specifications can help reduce residual variation and increase the precision of our estimated coefficients of interest. First, we use weather and precipitation data from Schlenker and Roberts (2009; 2020). The authors use PRISM climate data and weather monitoring stations to develop daily weather and precipitation data for 2.5 by 2.5-mile grids

throughout the contiguous United States from 1900-2020.[27] We geocode each water quality monitoring location to the nearest PRISM grid centroid. In our analysis, we use daily measures of total precipitation and maximum temperature at each monitoring location.

We also control for variation in macroeconomic conditions during our sample period. We control for median household income in our analysis, which the U.S. Census Bureau's American Community Survey (ACS) provides at the county-year level. We also use unemployment statistics from the Bureau of Labor Statistics Local Area Unemployment Statistics (LAUS) program. The LAUS provides unemployment percentages for each county in the United States every month.

Finally, we control for land usage patterns and land cover. We gather these measures from two data sources. First, we collect our measures of land usage – the county level number of farm acres and the number of acres with commercial fertilizer applied – from the USDA Census of Agriculture, which USDA administers every five years.[28] Second, the National Land Cover Database (NLCD) provides the most comprehensive and most frequently updated land cover maps for the United States. The NLCD classifies land cover at 30m resolution for the years of 2001, 2003, 2006, 2008, 2011, 2013, and 2016. For each of these years, we overlay the NLCD maps on HUC8 regions and calculate the percentage of each HUC8 that is developed, forested, and planted.

*Sample summary statistics and trends*

We construct two final analysis samples: one for total phosphorus readings and another for ammonia readings. Table 1 provides summary statistics for each sample. The mean total phosphorus and ammonia readings at the monitoring location level are 0.258 mg/L and 0.227 mg/L, respectively.[29] Online Appendix Table A1 presents the heterogeneity in these measures by CAFO exposure. Table 1 also presents summary statistics for our measure of CAFO intensity, which represents

treatment. The mean monitoring location (for each nutrient type) is exposed to over six HUC8 level CAFOs between 1995 and 2017. During our sample period, some HUC8s do not contain any CAFOs, while the maximum CAFO count is 41.[30]

We next discuss the trends in our CAFO and nutrient concentration measures. Figure 3 shows how average nutrient concentrations and exposure to CAFOs have trended over time across all monitoring locations. Panel A shows trends for total phosphorus and Panel B shows trends for ammonia. Both panels reveal increases in the number of CAFOs to which an average water quality monitor was exposed coincident with the increases in nutrient concentrations, especially during the latter portion of our study period. Online Appendix figures A1 and A2 present analogous figures where we separate water quality monitors into three groups based on ultimate levels of CAFO intensity. These figures reinforce the correlation between CAFO intensity and nutrient concentrations in surface waterbodies. Although these patterns are suggestive, time-invariant factors across HUC8 regions or secular trends throughout Wisconsin in CAFO intensity and nutrient levels could drive the correlations seen in these figures. We address both potential confounders in the empirical analysis of the following section.

**Empirical analysis**

In this section, we describe our empirics. We first discuss the baseline estimation and describe our identification strategy. Next, we provide the primary DD estimation results and evidence supporting our identifying assumptions. We then monetize the water pollution externality associated with CAFOs using EPA methodology to value non-market water quality benefits. Third, we address several threats to our identification. Finally, we examine the sensitivity of our results to the choice of regression specification and analysis sample.

*Estimating equation and identification*

We are interested in the effect of CAFO presence and expansion on a HUC8 region's surface water quality. We begin by estimating the following DD specification:

$$Y_{ijdmt} = \beta_1 CAFO_{jdmt} + \beta_2 M_{ijdmt} + \beta_3 X_{jdmt} + \gamma_j + \psi_m + \lambda_t + \varepsilon_{ijdmt}, \quad (1)$$

where $Y_{ijdmt}$ is the concentration (in mg/L) of total phosphorus or ammonia at monitoring location $i$ in HUC8 region $j$ on day $d$ in month $m$ of year $t$. We code $CAFO_{jdmt}$ as the count of CAFOs present in the HUC8 region each day. In the online appendix, we also estimate specifications where $CAFO_{jdmt}$ represents the number of animal units at CAFOs within each HUC8.

$M_{ijdmt}$ and $X_{jdmt}$ are vectors of time-varying controls at the monitoring location- and HUC8-day levels, respectively. The components of each vector likely impact the nutrient concentrations in surface waterbodies and thus, inclusion of these control factors can reduce the variance of the error term and improve the precision of our estimates. $M_{ijdmt}$ is a vector of climatological controls at the monitoring location-day level: precipitation (including its square) and maximum temperature. We control for daily total precipitation at each monitoring location because snow and rain events affect how much non-point source runoff occurs, which impacts surface water quality. We include the square of total precipitation because dilution occurs during the runoff process, which may affect the level of pollutants that reach each waterbody. We also include a measure for the maximum daily temperature because temperature affects nutrient concentrations, primarily through the release of legacy pollutants from sediment (Jensen and Andersen 1992; Genkai-Kato and Carpenter 2005).

Next, $X_{jdmt}$ includes several time-varying controls at the HUC8-day level. First, $X_{jdmt}$ contains macroeconomic indicators. We include as controls median household income and unemployment rates, both of which may affect surface water quality for several reasons. If

macroeconomic conditions signify a poor economy, then citizen pressure for environmental protection is likely to fall as the concern most important to policymakers is not the environment but the economy (Earnhart 2004). Also, unemployment can affect pro-environmental behavior and attitudes (Meyer 2016). Moreover, these measures could proxy other economic activities that plausibly affect surface water quality. If changes in non-CAFO economic activity are correlated with CAFO expansion, we could misattribute general economic effects to CAFOs. Third, $X_{jdmt}$ contains land usage and land cover measures. For land usage, we include in our specification controls for the number of farm acres and the number of farm acres that are applied with commercial fertilizer; each of these measures contributes to non-point source pollution. Finally, $X_{jdmt}$ includes NLCD land cover classifications in percentages, e.g., percent of the HUC8 that is forested. These land cover measures control for the likelihood that other, non-CAFO, non-point source pollution occurs in each HUC8. In addition, these measures capture the likelihood of urban runoff, which can affect nutrient concentrations in surface waterbodies. Alternatively, more developed land may be associated with decreased non-point source nutrient runoff if there is no agriculture in the region. $X_{jdmt}$ contains measures for the percentage of land in each HUC8 that is forested, planted, or developed; our omitted category is the percentage of wetlands.

Equation (1) also contains a series of fixed effects. $\gamma_j$ are HUC8 fixed effects that control for time-invariant HUC8 level characteristics that affect surface water quality. HUC8 fixed effects control for factors such as the size of the HUC8, land slope, and soil type, the last of which affects the ability of non-point source pollution to absorb into the ground before reaching surface waterbodies. $\psi_m$ are month fixed effects that control for the seasonality of the pollutant concentration readings and $\lambda_t$ are year fixed effects that control for unobservable factors common to the entire state that change over time, e.g., state or national policies to control non-point source pollution.[31,32]

Finally, we cluster standard errors in the baseline regressions at the HUC8 level, which is our level of identifying variation, to allow for within hydrological region correlation in the error term (spatial and serial correlation). We also show results for our main specification where we use two-way clustering of standard errors (HUC8 and year).

Identification of our coefficient of interest, $\beta_1$, comes from changes in water quality within HUC8 regions coincident with plausibly exogenous variation in CAFO presence and intensity. Our identifying assumption, which is standard in DD models, is the parallel trends assumption; in the absence of any CAFO expansion, average nutrient concentrations in surface waterbodies would have trended the same way in HUC8 regions with varying intensities of treatment from CAFOs. Although this assumption is fundamentally untestable, examining the trends in nutrient concentrations in the years leading up to CAFO treatment can be informative. We examine pre-treatment trends in Online Appendix C and address other threats to identification in a subsequent sub-section.

*Primary estimation results*

Table 2 tabulates results for the estimation of equation (1).[33] The first three columns show results for total phosphorus and the second three columns present results for ammonia. Columns 1 and 4 show regression results with no time-varying controls and all other columns present results including monitor level and HUC8 level time-varying covariates. Including these controls does not change the qualitative results but does add precision to our estimates. Exclusion of these factors could incorrectly attribute all changes in nutrient readings solely to CAFO exposure if the controls are correlated with CAFO exposure. The addition of one CAFO to a HUC8 region leads to an increase in average total phosphorus levels of 0.00436 mg/L. Relative to the mean total phosphorus reading from our sample of 0.258 mg/L, this change represents a 1.7% increase in total phosphorus

levels in surface waterbodies. And relative to the median of 0.07 mg/L, an additional CAFO increases total phosphorus levels by 6.2%. The magnitude of the effect is somewhat larger for ammonia. One additional CAFO in a HUC8 region increases average ammonia concentrations by 0.00614 mg/L. This change represents a 2.7% increase relative to the mean and a 12.1% increase relative to the median. To further put the magnitude of the results into context, note that the average water quality reading in our total phosphorus (ammonia) sample is exposed to 6.4 (6.1) CAFOs. Thus, the average total phosphorus (ammonia) reading in Wisconsin is approximately 10.9% (16.5%) higher than it would be in a counterfactual world without any CAFOs. As seen in columns 3 and 6, our main coefficients of interest remain statistically significant at conventional levels when we implement two-way clustering at the HUC8 and year levels.[34]

*Economic impacts*

It is important to quantify the increased nutrient concentrations associated with CAFO expansion to understand the implications for ecological health and compliance with environmental regulations, e.g., surface water quality standards. However, society loses non-market benefits when surface water pollution increases so we also need an estimate of the extent of economic damage. Our model predicts improved surface water quality in the counterfactual world where Wisconsin experiences less CAFO expansion. Improved water quality enhances aquatic ecosystems and provides benefits to humans outside of market transactions; we provide four examples. First, enhanced fish populations increase the scope and quality of recreational fishing opportunities. Second, decreased eutrophication enhances swimming and boating experiences. Third, outings that occur near surface waters such as hiking, biking, or wildlife viewing benefit from improved water quality. Finally, non-use benefits occur when individuals place value on water quality improvements even

without directly using the water resource themselves (Griffiths et al. 2012).

We estimate the non-market benefits from improved water quality using a benefit transfer approach (Carson and Mitchell 1993; EPA 2009). Benefit transfer is a standard non-market valuation method used to estimate the welfare change from a policy or scenario using value estimates from past research. Broadly, we 1) use our econometric model to simulate changes in nutrient concentrations in a counterfactual world with one fewer CAFO per HUC8 per year (for years in which the HUC8 had at least one CAFO),[35] 2) translate changes in nutrient concentrations to changes in water quality indices, and 3) monetize changes in water quality indices using two benefit transfer functions from the literature (Carson and Mitchell 1993; EPA 2009). We highlight three important details of our benefit calculations:

1. The analysis uses a water quality index (WQI) approach, which translates water quality measurements into a single index. The WQI is measured on a scale of 0 to 100, with 0 representing the worst water quality and 100 representing the best water quality. We adapt the EPA (2009) WQI, which generalizes McClelland's (1974) and Cude's (2001) WQI methods to the national scale. This method takes water quality measurements from each parameter included in the WQI, transforms the measurements into sub-index values on a 0-100 scale, and aggregates the sub-indices into the overall WQI. Walsh and Wheeler (2013) provide justification that WQI aggregation is appropriate for estimating benefits and costs of water quality changes.
2. The EPA (2009) WQI has six sub-indices based on six water quality parameters [total phosphorus (P), total nitrogen (N), dissolved oxygen (DO), biochemical oxygen demand (BOD), fecal coliform (FC), and TSS]. Of these six water quality parameters, we estimate the effects of CAFOs on P and N (from ammonia). Of the remaining four

pollutants, only TSS is largely uncorrelated with P and N (Griffin et al. 2020). We therefore set TSS at its HUC8 median to create an unbiased baseline water quality index.[36] The EPA (2009) WQI weights P at 14%, N at 14%, and TSS at 11%, so we create a WQI that maintains these relative weights and sums to 100% (35.9% P, 35.9% N, 28.2% TSS).[37] As in EPA (2009), we use the weighted geometric mean of the subindices.

3. We present WTP estimates using the benefit transfer functions from both Carson and Mitchell (1993) and EPA (2009). In both cases, the benefit transfer functions translate predicted changes in the WQI into household level WTP. Some studies in the extant benefit transfer literature (e.g., Walsh and Wheeler 2013) use the benefit transfer function from Carson and Mitchell (1993) while others (e.g., EPA 2009) instead use a function fitted from a meta-analysis of surface water valuation studies. There are arguments for and against using a meta-analysis benefit transfer function in lieu of a single study such as Carson and Mitchell (1993) to monetize benefits.[38] EPA (2009) provides the complete technical details of the meta-analysis estimation used to generate their benefit transfer function.[39]

To illustrate methods, we focus on the example of total phosphorus. We first apply the equation from EPA (2009) that links total phosphorus concentrations to the total phosphorus subindex (0-100 scale) for each water quality monitor in our sample. We then use equation (1), with all controls included (column 2 of Table 2), to simulate total phosphorus levels at each water quality monitoring location in the counterfactual scenario of one fewer CAFO in each HUC8 for each year in which the HUC8 had at least one CAFO. We again use EPA (2009) equations to link total phosphorus levels to a WQI subindex. Finally, we calculate the HUC8-year mean of the total

phosphorus subindex for the observed and counterfactual phosphorus concentrations.

We complete the same process for nitrogen[40] and TSS (holding the latter subindex constant between the two scenarios) and then generate the composite HUC8-year WQI. The mean HUC8-year WQI for the observed historical CAFO expansion is 69.18. For reference, Vaughan (1986) classifies water quality based on its suitability for potential uses: 25=boating, 45=rough fishing, 50=game fishing, 70=swimming, and 95=drinking without treatment. Therefore, Wisconsin's average water quality during our sample period is suitable for game fishing. However, there is substantial heterogeneity amongst surface waterbodies; 20% of sample locations have phosphorus subindices below 25 and 28% of sample locations have phosphorus subindices below 50.[41] We input the change in the HUC8-year WQI into the two benefit transfer functions, which generate mean annual household WTP values in 1983 dollars for the Carson and Mitchell (1993) function and in 2008 dollars for the EPA (2009) function. We next create statewide average household WTP values for each sample year by averaging over HUC8 regions and weighting by HUC8 population. We then apply an annual discount rate of 3% to the mean annual WTP values and sum the WTP into a present value of the 23 years of benefits that were lost associated with CAFO expansion. Finally, we convert the present value of the 23 years of benefits to 2017 dollars.

The WTP value is a nonlinear function of the WQI, which is a nonlinear function of the coefficient on CAFOs from our regression model. These nonlinearities imply that procedures to estimate WTP, such as the delta method, are inappropriate because they produce symmetric confidence intervals. Therefore, we empirically bootstrap the distribution of WTP using a Krinsky and Robb (1986) procedure. We assume that the coefficient on CAFOs is normally distributed with mean and standard deviation taken from our baseline estimation of equation (1). We randomly draw from the coefficient distribution to calculate new parameter estimates for the effects of

CAFOs on total phosphorus and ammonia concentrations. Then, we follow the above-described procedure for calculating the simulated WQI and associated WTP for the counterfactual of one fewer CAFO per HUC8 per year. We repeat this process for 10,000 random draws and then obtain a 95% confidence interval around the median by dropping the highest and lowest 2.5% of the simulated WTP values.

We find that the average household would be willing to pay a present discounted value of $67.93 using the Carson and Mitchell (1993) WTP function and $274.18 using the EPA (2009) meta-analysis WTP function (both in 2017 dollars) for the non-market benefits from improved surface water quality they would have experienced with one fewer CAFO per HUC8 each year (1995-2017).[42] Averaged over the 23 years of our sample, these values imply that the average household would be willing to pay $2.95 to $11.92 per year to avoid a marginal CAFO in their HUC8. According to the U.S. Census Bureau ACS 5-year estimates, there were 2,343,129 households in Wisconsin from 2014-2018 (U.S. Census Bureau 2019). Therefore, on an annual basis, a policy that would have prevented a marginal CAFO in each HUC8 region of Wisconsin from 1995-2017 would be valued between $6.9 million and $27.9 million (in 2017 dollars).

To put these values into context, Wisconsin's GDP was approximately $322 billion in 2017. And for further context, a 2019 University of Wisconsin report estimates that dairy farming directly contributed around $5.5 billion of sales revenue to Wisconsin's economy in 2017 (Deller 2019). In 2019, "CAFOs represent about 3.5% of the state's dairy farms but are home to nearly 25% of its dairy cows" (Cushman 2019). Therefore, dairy CAFOs in Wisconsin produce around $1.375 billion in annual sales revenue. There were 282 permitted dairy CAFOs in 2017, which implies approximately $4.9 million in annual sales revenue per dairy CAFO. In 2017, 34 HUC8 regions in Wisconsin had at least one CAFO. Thus, further assuming that all CAFOs in all regions

have equal impact on water quality and willingness to pay, Wisconsin's total annual willingness to pay to avoid the water quality damages from one CAFO (including dairy CAFOs) would be $203,541-$821,533.[43] Our exercise therefore implies that the water quality damages from a marginal dairy CAFO in Wisconsin compare to approximately 4-17% of their associated sales revenue.

To gauge the relative level of our estimated values, we place them into the context of several studies that estimate non-market benefits of surface water quality. First, EPA (2000a) estimates the benefits of the CWA (using the Carson and Mitchell (1993) benefit transfer function) at $177 per household per year. Thus, our valuation estimates from the same benefit transfer function represent roughly 1.5% of those from the CWA. Second, Walsh and Wheeler (2013) also use the Carson and Mitchell (1993) benefit transfer function to examine the benefits of EPA's 2003 CAFO rule. The authors estimate nationwide benefits of the rule at over $287 million per year. Our aggregated annual value of $6.9 million, which represents the non-market benefits from one fewer CAFO per HUC8 in Wisconsin, compares to roughly 2% of Walsh and Wheeler's (2013) estimate, which represents the non-market benefits from the improved manure management at all CAFOs in the country. Third, Egan et. al (2009) develop a recreational demand WTP model using several water quality parameters. Using their preferred specification, the authors estimate that improving water quality at 128 lakes in Iowa to the level of one of the cleanest lakes in the state (West Okoboji Lake) is worth nearly $189 in annual benefits for the average household. The estimated benefits from one fewer CAFO in each HUC8 in Wisconsin represents over 1.5% of Egan et al.'s (2009) value. The ratio of our estimated benefits to those of Egan et al. (2009) coincides with the ratio of water quality changes in each scenario.[44] Fourth, we provide a final comparison with the hedonics literature. As mentioned, nutrients (the outcome of interest in this study) contribute to algal growth in surface waterbodies, which can lead to harmful algal blooms (Dodds et al. 2009). Therefore,

CAFOs in Wisconsin likely contribute to the prevalence of algal blooms in the state. We compare the surface water damages from CAFO's nutrient increases in Wisconsin to the property value damages from algal blooms in Ohio (Wolf and Klaiber 2017). The authors find that harmful algal levels at one lake in Ohio are responsible for $51 million in decreased property values surrounding the lake. As before, our estimated damages ($6.9 million) are smaller than Wolf and Klaiber's (2017) damages because we estimate relatively small changes in water quality, while their estimates stem from a discrete jump to extremely poor water quality. Collectively, the estimated damage values of our study parallel those from different contexts within the literature.

*Threats to identification*

In this sub-section, we lend support to our identification and to the validity of our results in several ways. We consider the possibility of endogenous water quality sampling, examine the effects of CAFOs on placebo outcomes, and address possible endogenous CAFO location and expansion decisions.

The water quality readings that we use as outcomes are measured intermittently in time and space. It is therefore possible that government agencies and citizen volunteers are more likely to sample water quality under certain conditions. And these conditions could be correlated with nutrient concentrations. For example, it is well known that precipitation events create runoff from agricultural fields, which increases nutrient concentrations in sampled waterbodies. It is plausible that concerned agencies or individuals would then be more likely to sample water quality on the day of the precipitation event, or possibly a day or two after the event. Because we control for precipitation in our DD framework, the biggest threat to our identification would be if the probability of sampling increased differentially in watersheds with more CAFOs. Empirically, a

statistically significant and economically meaningful positive effect of the interaction between total precipitation and CAFO intensity would likely indicate that sampling is endogenous and our estimates biased. We test for differential sampling by examining the effects of CAFOs and total precipitation levels on the probability of sampling at a given water quality monitoring location on a given day.[45] To perform this exercise, we create a daily panel dataset of all the monitoring locations in each analysis sample. We then link these panels with daily CAFO exposure (at the HUC8 level) and daily total precipitation. Using these panels, we estimate the following specification using ordinary least squares:

$$Sampled_{ijdmt} = \beta_1 CAFO_{jdmt} + \beta_2 Prec_{ijdmt} + \beta_3 CAFO_{jdmt} \times Prec_{ijdmt} + \gamma_j + \psi_m + \lambda_t + \varepsilon_{ijdmt}, \qquad (2)$$

where the notation is identical to that used in equation (1). There exist three key differences between equations (1) and (2). First, the outcome, $Sampled_{ijdmt}$, in equation (2) is now a dummy that indicates the presence of a water quality sample (including non-detects) of total phosphorus or ammonia, at a given monitoring location, on a given day. Second, we include as a stand-alone regressor $Prec_{ijdmt}$, which measures: 1) daily precipitation at each monitoring location or 2) a three-day average of precipitation at each monitoring location that includes precipitation of the day itself and the two days prior. Third, we include the interaction of our CAFO measure and the precipitation measures to examine the differential probability of sampling based on precipitation amounts and treatment intensity. As before, we cluster standard errors at the HUC8 level.

Table 3 presents results for the estimation of equation (2). We highlight two conclusions from these results. First, the interaction between $CAFO_{jdmt}$ and $Prec_{ijdmt}$ is statistically insignificant and very close to zero. Second, the precipitation coefficients are also statistically insignificant and practically zero. Collectively then, precipitation events (and those of the recent past) do not

affect the probability that a given water quality monitoring location is sampled when there is zero CAFO exposure at that location. Importantly, as CAFO intensity increases, the probability of a sample on these days with precipitation events does not change. Thus, we are confident that the intermittent nature of water quality readings does not bias our primary estimates.

Although we provide evidence that sampling does not appear to be endogenous, doubts about the timing of sampling surrounding precipitation events could remain. To assuage this concern, we also estimate equation (1) at the aggregated HUC8-month level. The estimating equation then becomes:

$$Y_{jmt} = \beta_1 CAFO_{jmt} + \beta_2 X_{jmt} + \gamma_j + \psi_m + \lambda_t + \varepsilon_{jmt}, \quad (3)$$

where $Y_{jmt}$ is the monthly average nutrient reading of all samples taken from all monitoring locations in HUC8 region $j$. Some HUC8 regions have many more underlying water quality readings available in a month than others and thus provide a more reliable measure of true water quality. We therefore use the number of dependent variable readings in a HUC8 region in a month as analytic weights in all regressions at the HUC8 level.[46] As shown in Table 4, estimates are like our baseline results.

Next, we lend support to our identification by estimating our primary regression specification with placebo outcomes. If CAFO presence and intensity significantly affect surface waterbody concentrations of pollutants unrelated to CAFOs, then our model is likely misspecified. We therefore aim to identify surface waterbody pollutants that are ubiquitous throughout the United States and Wisconsin but should not be associated with CAFOs. Non-point source pollution can consist of pesticides used on planted acres; concentrations of these chemicals in surface waterbodies likely indicate agricultural activity within the watershed. However, CAFOs do not grow crops on any part of the lot/facility and operators therefore do not apply pesticides at CAFOs. We estimate

equation (1) using two separate measures of pesticide presence in surface waterbodies as outcomes. We first use the concentration of pesticides present at surface waterbodies as our outcome. We collect pesticide measurements from the water quality portal's "Organics, Pesticide" pollutant group and operationalize this measure in a way identical to that of our primary nutrient readings.[47] We also estimate a specification where the placebo outcome is a dummy variable that indicates a reading that does not detect a positive concentration of pesticides, i.e., non-detect. We examine non-detected measurements because pesticide concentrations during our sample period are low and thus, results for pesticide concentrations may be sensitive to small changes in the outcome. In fact, 85% of pesticide measurements in our sample are coded as non-detects. Table 5 presents results for the estimation of these falsification tests. As expected, the number of CAFOs at the HUC8 level does not affect pesticide concentrations or the detection of non-zero pesticide levels at surface waterbodies in Wisconsin; coefficient estimates for the CAFO measures in these falsified specifications are statistically insignificant and close to zero. These falsification tests lend support to our identification and to the validity of our primary estimates.

Finally, we address the possibility of endogenous CAFO location and expansion. For our estimates to be unbiased, the CAFO intensity measures should be uncorrelated with the error term in equation (1) conditional on time-invariant unobservables, aggregate time effects, and time-varying controls. It is possible that CAFO permittees endogenously choose to locate in a HUC8 region concurrently with changes in other unobserved determinants of nutrient concentrations. Another potential endogeneity issue is raised by Sneeringer and Kay (2011), who find bunching in the distribution of operations just below the size threshold for CAFOs. Operations that cross the threshold to become CAFOs may be different in ways that are systematically related to water pollution and correlated with location selection and expansion timing, creating a selection into

treatment on unobservables problem. The ideal solution to these sources of potential omitted variable bias is to find a time-varying instrument for CAFO expansion within a HUC8. The instrument must influence location decisions of CAFOs without directly affecting surface water quality; such an instrument is difficult to find in the real world. In Online Appendix D, we show an alternative long-difference estimation strategy where we instrument for CAFO expansion with freeway access. Here, we show how Oster (2019) bounding methods can inform on the direction and magnitude of bias from potential omitted variables.

Oster (2019) shows that one can establish the robustness of an estimate to omitted variable bias by examining coefficient and R-squared stability. Under two assumptions, one can estimate a consistent bias-adjusted treatment effect (Oster 2019). The first necessary assumption is the relative selection of unobservable and observable factors that are related to treatment; Oster denotes this ratio as $\delta$. Oster argues that $\delta = 1$ is an appropriate assumption for generating a bounding set on the treatment effect. Intuitively, this assumes that treatment selection is equally driven by observable and unobservable characteristics. The second assumption involves the theoretical maximum value of R-squared, denoted $R^2_{max}$. Here, Oster suggests that $R^2_{max}$ will typically be smaller than one because of measurement error in *Y* or because other factors affect *Y* after the assignment of treatment. In our context, there is reason to suspect measurement error in surface waterbody nutrient concentrations. Furthermore, decisions to expand CAFOs are made well before the realization of many other factors that affect pollution levels. Oster's analysis of randomized data suggests setting $R^2_{max}$ to 1.3 times the r-squared of the regression including controls. In the notation of Oster (2019), this means that $R^2_{max} = 1.3\tilde{R}$. Then, under these assumptions, the bounding set for the treatment effect of interest is $\Delta = \left[\tilde{\beta}, \beta^*\left(1.3\tilde{R}, 1\right)\right]$. For each estimated treatment effect, we report the bounding set and indicate whether this set includes zero.

Table 6 summarizes the bounding sets for our treatment effects of interest. For both dependent variables, we show the bounding sets for the coefficient on the number of CAFOs. In each case, the bounding set excludes zero; correcting for selection on unobservables moves the estimated treatment effect further away from zero. Table 6 suggests that CAFOs may affect water quality more than we report in our baseline estimates on the monitor-level sample. The bias-corrected effect sizes are especially large for total phosphorus when we set $R^2_{max} = 1.3\tilde{R}$. For comparison purposes, we also show bias-corrected effects when assuming the theoretically max R-squared is 1.15 times the R-squared of the regressions including controls. Bias-corrected effects are still larger than the baseline estimates but get closer to the magnitude of the baseline estimates. Table 6 also shows a similar pattern in bounding sets for estimates from our HUC8 aggregated analysis. The takeaway message from this robustness exercise is that our baseline estimates likely represent a lower bound on the true treatment effects of interest. Selection on unobservables may lead us to understate the extent that CAFOs affect surface water quality.

*Sensitivity analysis*

In this sub-section, we test the sensitivity of our primary estimation results to changes in the regression specification and the analysis sample. First, we add to equation (1) a HUC8 level measure for the number of animals at livestock operations net of those on our sampled CAFOs. We include this control to better account for the non-point source pollution that can occur from livestock operations, regardless of their size.[48] We gather the livestock count within each county-year from the USDA National Agricultural Statistics Service June Area Surveys. We then transform this measure to the HUC8-day level like our other county-year measures and subtract the HUC8 level count of CAFO animal units. The first and second columns of Table 7 present estimation results for this

exercise and show that our treatment measure is robust to the inclusion of non-CAFO animals as a control.

To further disentangle the effects of CAFO and non-CAFO agriculture on surface water quality, we examine the relationship between the livestock inventory at CAFOs and the livestock inventory at operations that are not CAFOs, in each HUC8-year. Because neither measure necessarily represents a dependent or independent variable, we examine this relationship using a partial correlation. We calculate the partial correlation between these measures by first separately estimating each measure as a function of year and HUC8 fixed effects and the control factors of equation (1). Then, we calculate the correlation between the two sets of residuals. For our two measures of livestock inventory, the partial correlation is -0.295 (p=0.000). We interpret this negative and significant correlation to represent two separate, but possibly related, scenarios. First, this relationship provides evidence of the consolidation of the livestock industry. According to USDA (2020) data, the number of dairy cows in Wisconsin has declined over the past half century. The negative relationship that we witness may suggest that when CAFOs enter an area, they consolidate smaller livestock operations that previously existed. Second, the relationship may signify the "crowding out" of smaller operations. Because CAFOs can produce more output at a lower cost than smaller livestock operations, they may push smaller operations out of business or to different areas. Importantly, the negative relationship between CAFO and non-CAFO animals supports the argument that the extreme concentration of manure at CAFOs is likely the cause of the estimated water quality effects (Kellogg et al. 2014; Jones et al. 2018), independent of non-CAFO animals, because of the spatial separation of the different types of livestock operations.[49] However, we acknowledge that we only identify a correlational relationship between these two measures. Future research should better disentangle these effects and identify a possible mechanism.

Second, we add to our primary regression specification controls for the legacy concentrations of nutrients. Past period CAFO exposure or other agricultural or urban activity can lead to the build-up of nutrients in surface waterbodies. Thus, inclusion of these controls allows us to model both the stock and flow of nutrients in surface waterbodies and to isolate the effects on nutrient concentrations from current CAFO exposure. Because nutrient concentrations are measured intermittently, we construct two variables, one for each outcome, that measure the average nutrient concentrations of each HUC8 over the past five years to account for the stock of pollutants in surface waterbodies. The third and fourth columns of Table 7 present results for the estimation of equation (1) with these measures included. Inclusion of these controls does not substantially alter the estimates of our primary specification presented in Table 2. We also provide several sets of estimation results in the online appendix that further model the stock and flow of pollutants. Online Appendix Table A8 tabulates estimation results where we control for the previous nutrient concentration level at each monitoring location (Paudel and Crago 2020) and where we control for CAFO exposure one through five years before each pollutant concentration reading. Our primary estimation results are robust both qualitatively and quantitatively to the inclusion of these additional controls.

We also test the robustness of the results to several alternative specifications. We provide complete details, justification, and results for these robustness tests in Online Appendix G. Additional specifications include: 1) adding year-by-HUC2 fixed effects (rather than only year fixed effects), 2) using HUC10 level fixed effects (rather than HUC8 fixed effects), 3) adding a variable to control for HUC8 level land in the Conservation Reserve Program (CRP), 4) examining the effects of upstream CAFO presence and intensity, and 5) allowing for differential trends before and after CAFOs enter a HUC8 watershed. For changes to the analysis sample, we: 1) limit the

sample to HUC8 regions with a CAFO presence throughout our sample period, 2) limit the sample to the growing season months of April-October, and 3) limit the sample to exclude observations from HUC8 regions with poultry CAFO presence. Collectively, empirical results are robust to changes in model specification and analysis sample.

**Conclusion**

Previous studies suggest a negative correlation between CAFO presence and surface water quality in nearby areas, but prior work lacks evidence using longitudinal data on a large spatial scale. To address this gap in the literature, we assemble a panel dataset that links essentially the entire history of CAFO expansion in Wisconsin with readings on total phosphorus and ammonia levels at surface water monitoring locations in corresponding HUC8 regions. Our identification strategy leverages plausibly exogenous within-HUC8 level variation in the timing and extent of CAFO expansion to identify an effect. We find that adding an additional CAFO to a HUC8 region increases average total phosphorus (ammonia) levels in surface waterbodies by approximately 1.7% (2.7%), relative to sample means.

We then use water quality index and benefit transfer methodologies to convert increased nutrient concentrations associated with CAFO expansion to losses in non-market surface water quality benefits, i.e., damages. This methodology estimates changes in a water quality index and monetizes the changes with a benefit transfer function. We estimate an annualized WTP of $3-$12 per Wisconsin household for the improved water quality that would exist in the counterfactual world with one fewer CAFO in each HUC8 region. Aggregated to the entire state, Wisconsin households would be willing to pay between $6.9 million and $27.9 million annually for one fewer CAFO in each HUC8 region.

Our findings present important policy implications. We find that the surface water quality

damages from the marginal CAFO in Wisconsin are at least $203,541 per year. These damages do not represent the only external cost from CAFOs. However, most of the literature simply presents associations between CAFOs and various outcomes, making our valuation of the water quality damages from CAFOs an important contribution. We are aware of only two monetized estimates of CAFO related external costs in dimensions other than surface water quality. First, Herriges et al. (2005) estimate that, in Iowa, upwind CAFOs decrease home prices by three percent, which is roughly $2,500 for the average home in their sample. And second, Sneeringer (2010) estimates that the annual air pollution related damages from hogs in the United States are $31 per animal. Applying this value, a CAFO with 2,500 hogs (equivalent to 1,000 animal units) produces $77,500 in air pollution damages per year. In addition to these monetized damages, previous studies find that CAFO exposure is associated with damages to individual health (Sigurdarson and Kline 2005; Radon et al. 2007)[50] and occupational safety (Ramos et al. 2018). Other potential, but not yet studied, damages include commercial fishing impacts, increased contamination of private wells, and increased water treatment costs (EPA 2002b).[51] To achieve efficiency in CAFO product markets, policymakers and regulatory agencies can use our results, combined with those of previous studies, to develop policy that factors the external costs of CAFOs into their production decisions.

Also important for CAFO policy, we present evidence that the extreme concentration of manure in fewer locations likely drives the surface water quality damages, which prior work also suggests (Kellogg et al. 2014; Jones et al. 2018). Thus, CAFO control policy should require better manure management, e.g., subsidies to transport the manure off-site, to better control the surface water damages from these operations. However, future work should examine the specific avenue by which nutrients enter surface waterbodies, e.g., leaching, spreading.

Finally, we note that there is still much to learn concerning the effects of CAFOs on various

outcomes. We have documented the first large scale link that CAFO growth leads to worsening water quality in Wisconsin. This link establishes one mechanism for other potential impacts such as changes in recreational fishing patterns, changes in property values, or changes in human health. The next step should be analyses to discern whether these outcomes of interest are affected by CAFOs, and if so, to what (monetized) extent. We also acknowledge the difficulties in separating the effects of CAFO animals and non-CAFO animals on surface water quality. Although our results are robust to the inclusion of a measure of non-CAFO animals in each HUC8, future research should use experimental and quasi-experimental techniques to further disentangle this potential confounder. Lastly, our study focuses exclusively on one midwestern U.S. state. Although Wisconsin has significant CAFO presence, there are nine other states with higher numbers of total animal units and almost every state has some CAFO presence. Additional studies to identify water quality impacts in other states would help create a more comprehensive picture.

**References**

Ali, S., L. McCann, and J. Allspach, 2012, "Manure transfers in the Midwest and factors affecting adoption of manure testing", *Journal of Agricultural and Applied Economics*, 44(4), 533-548.

Anderson, D.M., P. Hoagland, Y. Kaoru, and A.W. White, 2000, "Estimated annual impacts from harmful algal blooms in the United States", Woodshole Oceanographic Institute Technical Report, WHOI-2000-11.

Barker, J.C. and J.P. Zublena, 1995, "Livestock manure nutrient assessment in North Carolina", Raleigh, NC: North Carolina Agricultural Extension Service, NC State University.

Bateman, I.J. and A.P. Jones, 2003, "Contrasting conventional with multi-level modeling

approaches to meta-analysis: expectation consistency in U.K. woodland recreation values", *Land Economics*, 79(2), 235-258.

Behrer, A.P., E.L. Glaeser, G.A.M. Ponzetto, and A. Shleifer, 2019, "Securing property rights", working paper, available at: https://scholar.harvard.edu/files/shleifer/files/securing_property_rights_and_appendices_august19.pdf.

Boxall, A.B.A., D.W. Kolpin, B. Halling-Sorenson, and J. Tolls, 2003, "Are veterinary medicines causing environmental health risks?", *Environmental Science & Technology*, 37, 286-294.

Burkholder, J., B. Libra, P. Weyer, S. Heathcote, D. Kolpin, P.S. Thorne, and M. Wichman, 2006, "Impacts of waste from Animal Feeding Operations on water quality", *Environmental Health Perspectives*, 115(2), 308-312.

Cameron, A. C., Gelbach, J.B., and D.L. Miller, 2008, "Bootstrap-based improvements for inference with clustered errors", *The Review of Economics and Statistics*, 90(3), 414-427.

Cameron, A. C., Gelbach, J.B., and D.L. Miller, 2011, "Robust inference with multiway clustering", *Journal of Business & Economic Statistics*, 29(2), 238-249.

Carson, R.T. and R.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.

Chen, C., G.E. Lade, J.M. Crespi, D.A. Keiser, 2019, "Size-based regulations and environmental quality: Evidence from the U.S. livestock industry", working paper, available at: https://ageconsearch.umn.edu/record/291262/files/Abstracts_19_05_15_21_01_17_38__129_186_248_69_0.pdf.

Copeland C., 2010, "Animal waste and water quality: EPA regulation of Concentrated Animal

Feeding Operations (CAFOs)", Congressional Research Service Report for Congress Order Code RL31851.

Corona, J., Doley, T., Griffiths, C., Massey, M., Moore, C., Muela, S., Rashleigh, B., Wheeler, W., Whitlock, S., and J. Hewitt. 2020. An integrated assessment model for valuing water quality changes in the United States. *Land Economics*. 96(4). 478-492.

Cude, C.G., 2001, "Oregon water quality index: A tool for evaluating water quality management effectiveness", *Journal of the American Water Resources Association*, 37(1), 125-137.

Cushman, W., 2019, "CAFO oversight in Wisconsin and who pays for it?" WisCONTEXT, Jun 13, 2019. https://www.wiscontext.org/what-does-cafo-oversight-look-wisconsin-and-who-pays-it.

Deller, S.C., 2019, "The contribution of agriculture to the Wisconsin economy: An Update for 2017", Department of Agricultural and Applied Economics Staff Paper, University of Wisconsin-Madison/Extension.

DeRouchey, J.M., 2014, "Manure management", In: Dikeman, M. and C. Devine (eds.), Encyclopedia of Meat Sciences, 2nd ed. Elsevier Academic Press, 152-156.

Dodds, W.K., W.W. Bouska, J.L. Eitzmann, T.J. Pilger, K.L. Pitts, A.J. Riley, J.T. Schloesser, and D.J. Thornbrugh, 2009, "Eutrophication of U.S. freshwaters: Analysis of potential economic damages", *Environmental Science & Technology*, 43(1), 12-19.

Earnhart, D., 2004, "The effects of community characteristics on polluter compliance levels", *Land Economics*, 80(3), 408-432.

Egan, K.J., J.A. Herriges, C.L. Kling, and J.A. Downing, 2009, "Valuing water quality as a function of water quality measures", *American Journal of Agricultural Economics*, 91(1), 106-123.

EPA, 2000a, “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”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2000b, “Pre-publication version: National Pollutant Discharge Elimination System permit regulation and effluent limitations guidelines and standards for Concentrated Animal Feeding Operations”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2001, “Environmental assessment of proposed revisions to the National Pollutant Discharge Elimination System regulation and the effluent guidelines for Concentrated Animal Feeding Operations”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2002a, “Economic analysis of the final revisions to the National Pollutant Discharge Elimination System regulation and effluent guidelines for Concentrated Animal Feeding Operations”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2002b, “Environmental and economic benefit analysis of final revisions to the national pollutant discharge elimination system regulation and the effluent guidelines for Concentrated Animal Feeding Operations”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2004, “Risk assessment evaluation for Concentrated Animal Feeding Operations”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2009, “Environmental impact and benefits assessment for final effluent guidelines and standards for the construction and development category”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2011, “The benefits and cost of the Clean Air Act from 1990 to 2020”, Washington, DC: U.S. Environmental Protection Agency.

EPA, 2020, “Water pollutant loading tool”, https://echo.epa.gov/trends/loading-tool/water-

pollution-search.

GAO, 2003, “Livestock agriculture: Increased EPA oversight will improve environmental program for Concentrated Animal Feeding Operations”, Washington, DC: U.S. Government Accountability Office.

GAO, 2008, “Concentrated Animal Feeding Operations: EPA needs more information and a clearly defined strategy to protect air and water from pollutants of concern”, Washington, DC: U.S. Government Accountability Office.

Genkai-Kato, M. and S.R. Carpenter, 2005, “Eutrophication due to phosphorus recycling in relation to lake morphometry, temperature, and macrophytes”, *Ecology*, 86(1), 210-219.

Grant, L. E. and K.K. Grooms, 2017, "Do nonprofits encourage environmental compliance?", *Journal of the Association of Environmental and Resource Economists* 4(S1), S261-S288.

Grant, L. and C. Langpap, 2019, “Private provision of public goods by environmental groups”, *PNAS*, 116(12), 5334-5340.

Gray, W.B. and J.P. Shimshack, 2011, “The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence”, *Review of Environmental Economics and Policy*, 5(1), 3-24.

Green, T.R., H. Kipka, O. David, and G.S. McMaster, 2018, “Where is the USA Corn Belt, and how is it changing”, *Science of the Total Environment*, 618, 1613-1618.

Greger, M. and G. Koneswaran, 2010, “The public health impacts of Concentrated Animal Feeding Operations on local communities”, *Family and Community Health*, 33(1), 373-382.

Grenestam, E. and M. Nordin, 2018, “Estimating the impact of agri-environmental payments on nutrient runoff using a unique combination of data”, *Land Use Policy*, 75, 388-398.

Griffin, R., Vogl, A., Wolny, S., Covino, S., Monroy, E., Ricci, H., Sharp, R., Schmidt, C., and

Uchida, E., 2020, "Including additional pollutants into an integrated assessment model for estimating nonmarket benefits from water quality", *Land Economics*, 96(4). 457-477.

Griffiths, C., H. Klemick, M. Massey, C. Moore, S. Newbold, D. Simpson, P. Walsh, and W. Wheeler, 2012, "U.S. Environmental Protection Agency valuation of surface water quality improvements", *Review of Environmental Economics and Policy*, 6(1), 130-146.

GRTS, 2020, "Grants reporting and tracking system", https://iaspub.epa.gov/apex/grts/f?p=109:9118.

Heaney, C.D., K. Myers, S. Wing, D. Hall, D. Baron, and J.R. Stewart, 2015, "Source tracking swine fecal waste in surface water proximal to swine Concentrated Animal Feeding Operations", *Science of the Total Environment*, 511, 676-683.

Heiskary, S. and H. Markus. 2001. Establishing Relationships among Nutrient Concentrations, Phytoplankton Abundance, and Biochemical Oxygen Demand in Minnesota, USA, Rivers. *Lake and Reservoir Management*. 17 (4). 251–62.

Herriges, J.A., S. Secchi, and B.A. Babcock, 2005, "Living with hogs in Iowa: The impact of livestock facilities on rural residential property values", *Land Economics*, 81(4), 530-545.

Hooda, P.S., A.C.Edwards, H.A. Anderson, and A. Miller, 2000, "A review of water quality concerns in livestock farming areas", *Science of the Total Environment*, 250, 143-167.

Hu, Y., H. Cheng, and S. Tao, 2017, "Environmental and human health challenges of industrial livestock and poultry farming in China and their mitigation", *Environment International*, 107, 111-130.

Hribar, C., 2010, "Understanding Concentrated Animal Feeding Operations and their impact on communities", Bowling Green, OH: National Association of Local Boards of Health.

Isakson, H.R. and M.D. Ecker, 2008, "An analysis of the impact of swine CAFOs on the value of

nearby houses", *Agricultural Economics*, 39, 365-372.

Jensen, H.S. and F. Andersen, 1992, "Importance of temperature, nitrate, and pH for phosphate release from aerobic sediments of four shallow, eutrophic lakes", *Limnology and Oceanography*, 37(3), 577-589.

Johnston, R.J, Besedin, E.Y, Iovanna, R., Miller, C.J., Wardwell, R.F., and M.H. Ranson, 2005, "Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: a meta-analysis", *Canadian Journal of Agricultural Economics*, 53, 221-248.

Johnston, R.J, Ranson, M.H., Besedin, E.Y., and E.C. Helm, 2006, "What determines willingness to pay per fish? A meta-analysis of recreational fishing values", *Marine Resource Economics*, 21(1), 1-32.

Johnston, R.J., Boyle, K.J, Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., and C.A. Vossler, 2017, "Contemporary Guidance for Stated Preference Studies", *Journal of the Association of Environmental and Resource Economists*, 4(2), 319-405.

Jones, C.S., C.W. Drake, C.E. Hruby, K.E. Schilling, and C.F. Wolter, 2018, "Livestock manure driving stream nitrate", *Ambio*, 48, 1143-1153.

Kellogg, R.L., D.C. Moffitt, and N.G. Gollehon, 2014, "Estimates of recoverable and non-recoverable manure nutrients based on the Census of Agriculture", Washington, DC: U.S. Department of Agriculture, Natural Resources Conservation Service.

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

Krinsky, I. and A.L. Robb, 1986, "On approximating the statistical properties of elasticities", *The*

*Review of Economics and Statistics*, 68(4), 715-719.

Landon, M.K., C.A. Burton, T.A. Davis, K. Belitz, and T.D. Johnson, 2014, "Statistical evaluation of variables affecting occurrence of hydrocarbons in aquifers used for public supply, California", *Journal of the American Water Resources Association*, 50(1), 179-195.

MacDonald, J.M and D. Newton, 2014, "Milk production continues shifting to large-scale farms", *Amber Waves*, USDA, Economic Research Service, Dec. 1, 2014. https://www.ers.usda.gov/amber-waves/2014/december/milk-production-continues-shifting-to-large-scale-farms.

MacKinnon, J.G, Nielsen, M.∅., and M.D. Webb, 2019, "Wild bootstrap and asymptotic inference with multiway clustering", *Journal of Business & Economic Statistics*, DOI: 10.1080/07350015.2019.1677473.

Madison, F.K., K. Kelling, L. Massie, and L.W. Good, 1995, "Guidelines for applying manure to cropland and pasture in Wisconsin", University of Wisconsin Extension Publication A3392, University of Wisconsin, Madison, WI.

Mallin, M.A., M.R. McIver, A.R. Robuck, and A.K. Dickens, 2015, "Industrial swine and poultry production causes chronic nutrient and fecal microbial stream pollution", *Water, Air, & Soil Pollution*, 226, 1-13.

McClelland, N.I., 1974, "Water quality index application in the Kansas River Basin", Kansas City, Missouri: U.S. Environmental Protection Agency Region 7.

Meyer, A., 2016, "Is unemployment good for the environment?", *Resource and Energy Economics*, 45, 18-30.

Meyer, D., P.L. Price, H.A. Rossow, N. Silva-del-Rio, B.M. Karle, P.H. Robinson, and E.J. DePeters, 2011, "Survey of dairy housing and manure management practices in

California", *Journal of Dairy Science*, 94, 4744-4750.

Meyer, K., 2018, "The impact of agricultural land use change on lake water quality: Evidence from Iowa", *Studies in Agricultural Economics*, 120, 105-111.

Moore, L.C., 2011, "Fifth Circuit Rules EPA Cannot Require Clean Water Permits For Livestock Farmers Unless Discharging into U.S. Waters", *Foster Swift Agricultural Law Update*, May 2011, available at: https://www.fosterswift.com/communications-EPA-Clean-Water-Permits-Livestock-Farmers.html.

Nickell, S., 1981, "Biases in dynamic models with fixed effects", *Econometrica*, 49(6), 1417-1426.

O'Connor, A.M., B. Auvermann, D. Bickett-Weddle, S. Kirkhorn, J.M. Sargent, A. Ramirez, and S.G. Von Essen, 2010, "The association between proximity to animal feeding operations and community health: A systematic review", *PLOS One*, 5(3), 1-13.

Olmstead, S.M., 2009, "The economics of water quality", *Review of Environmental Economics and Policy*, 4(1), 44-62.

Oster, E., 2019, "Unobservable selection and coefficient stability: Theory and evidence", *Journal of Business & Economic Statistics*, 37(2), 187-204.

Osterberg, D. and D. Wallinga, 2004, "Addressing externalities from swine production to reduce public health and environmental impacts", *American Journal of Public Health*, 94(10), 1703-1708.

Paudel, J. and C.L. Crago, 2020, "Environmental externalities from agriculture: Evidence from water quality in the United States", *American Journal of Agricultural Economics*, 103(1), 185-210.

Petersen, M.A., 2009, "Estimating standard errors in finance panel data sets: Comparing approaches", *The Review of Financial Studies*, 22(1), 435-480.

Prasad, M. B. K., Long, W., Zhang, X, Wood, R.J., and R. Murtugudde. 2011. Predicting Dissolved Oxygen in the Chesapeake Bay: Applications and Implications. *Aquatic Sciences.* 73 (3) 437–51.

Radon, K., A. Shulze, V. Ehrenstein, R.T. van Strien, G. Praml, and D. Nowak, 2007, "Environmental exposure to Confined Animal Feeding Operations and respiratory health of neighboring residents", *Epidemiology*, 18(3), 300-308.

Raff, Z. and D. Earnhart, 2019, "The effects of Clean Water Act enforcement on environmental employment", *Resource and Energy Economics*, 57, 1-17.

Ramos, A.K., A. Fuentes, and M. Carvajal-Suarez, 2018, "Self-reported occupational injuries and perceived occupational health problems among Latino immigrant swine confinement workers in Missouri", *Journal of Environmental and Public Health*.

Ready, R.C. and C.W. Abdalla, 2005, "The amenity and disamenity impacts of agriculture: Estimates from a hedonic pricing model", *American Journal of Agricultural Economics*, 87(2), 314-326.

Roodman, D., 2015. "BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed", Statistical Software Components S458121, Boston College Department of Economics, revised 26 Mar 2020.

Roodman, D., MacKinnon, J.G., Nielsen, M.∅., and M.D. Webb, "Fast and wild: Bootstrap inference in Stata using boottest", *The Stata Journal*, 19(1), 4-60.

Rosenberger, R., and T. Phipps, 2007, "Correspondence and convergence in benefit transfer accuracy: Meta-analytic review of the literature", In: Navrud, S. and R. Ready (eds.), Environmental Value Transfer: Issues and Methods. Dordrecht, The Netherlands: Springer, 23-43.

Seaber, P.R., F.P. Kapinos, and G.L. Knapp, 1987, Hydrologic Unit Maps: U.S. Geological Survey

Water Supply Paper 2294, available at: https://doi.org/10.3133/wsp2294.

Schlenker, W. and M. Roberts, 2009, "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change", *PNAS*, 106 (37), 15594-15598.

Schlenker, W. and M. Roberts, 2020, "Fine-scaled weather data", http://www.columbia.edu/~ws2162/links.html.

Shrestha, R., Rosenberger, R., and J. Loomis, 2007, "Benefit transfer using meta-analysis in recreation economic valuation", In: Navrud, S. and R. Ready (eds.), Environmental Value Transfer: Issues and Methods. Dordrecht, The Netherlands: Springer, 161-177.

Sigurdarson, S.T., and J.N. Kline, 2005, "School proximity to Concentrated Animal Feeding Operations and prevalence of asthma in students", *Chest*, 129, 1486-1491.

Sneeringer, S., 2009, "Does Animal Feeding Operation pollution hurt public health? A national longitudinal study of health externalities identified by geographic shifts in livestock production", *American Journal of Agricultural Economics*, 91(1), 124-137.

Sneeringer, S. 2010, "A national longitudinal study of the effects of concentrated hog production on ambient air pollution", *American Journal of Agricultural Economics*, 92(3), 821-835.

Sneeringer, S. and N. Key, 2011, "Effects of size-based environmental regulations: Evidence of regulatory avoidance", *American Journal of Agricultural Economics*, 93(4), 1189-1211.

Spellman, F.R. and N.E. Whiting, 2007, "Environmental management of Concentrated Animal Feeding Operations, Boca Raton, FL: CRC Press.

Townsend, A.R., R.W. Howarth, F.A. Bazzaz, M.S. Booth, C.C. Cleveland, S.K. Collinge, A.P. Dobson, P.R. Epstein, E.A. Holland, D.R. Keeney, M.A. Mallin, C.A. Rogers, P. Wayne, and A.H. Wolfe, 2003, "Human health effects of a changing global nitrogen cycle", *Frontiers in Ecology and the Environment*, 1(5), 240-246.

U.S. Census Bureau, 2019, QuickFacts, Wisconsin, https://www.census.gov/quickfacts/WI.

USDA, 2000, "Manure nutrients relative to the capacity of cropland and pastureland to assimilate nutrient: Spatial and temporal trends for the United States", Washington, DC: U.S. Department of Agriculture, Natural Resources Conservation Service.

USDA, 2002, "Profile of farms with livestock in the United States: A statistical summary", Washington, DC: U.S. Department of Agriculture, Natural Resources Conservation Service.

USDA, 2020, "Number of milk cows 1933 to current year", https://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Dairy/Historical_Data_Series/milkcowno.pdf.

Vaughan, W.J., 1986, "The RFF Water Quality Ladder." In: Mitchell, R.C. and R.T. Carson, The Use of Contingent Valuation Data for Benefit/Cost Analysis in Water Pollution Control, Final Report, Washington: Resources for the Future, Appendix B.

Walsh, P.J. and W.J. Wheeler, 2013, "Water quality indices and benefit-cost analysis", *Journal of Benefit Cost Analysis*, 4(1), 81-105.

Wells, M.L., V.L. Trainer, T.J. Smayda, B.S.O. Karlson, C.G. Trick, R.M. Kudela, A. Ishikawa, S. Bernard, A. Wulff, D.M. Anderson, W.P. Cochlan, 2015, "Harmful algal blooms and climate change: Learning from the past and present to forecast the future", *Harmful Algae*, 49, 68-93.

WDNR, 2011, "Environmental assessment comment summary and response for the large dairy Concentrated Animal Feeding Operation WPDES general permit", Madison, WI: Wisconsin Department of Natural Resources.

WDNR, 2019, "Wisconsin 2020 consolidated assessment and listing methodology (WisCALM)", Madison, WI: Wisconsin Department of Natural Resources.

Weldon, M.B. and K.C. Hornbuckle, 2006, "Concentrated Animal Feeding Operations, row crops, and their relationship to nitrate in Eastern Iowa Rivers", *Environmental Science &*

*Technology*, 40(10), 3168-3173.

Wolf, D. and H.A. Klaiber, 2017, "Bloom and bust: Toxic Algae's impact on nearby property values", *Ecological Economics*, 135, 209-221.

Zhang, W. and B. Sohngen, 2018, "Do US anglers care about harmful algal blooms? A discrete choice experiment of Lake Erie recreational anglers", *American Journal of Agricultural Economics*, 100(3), 868-888.

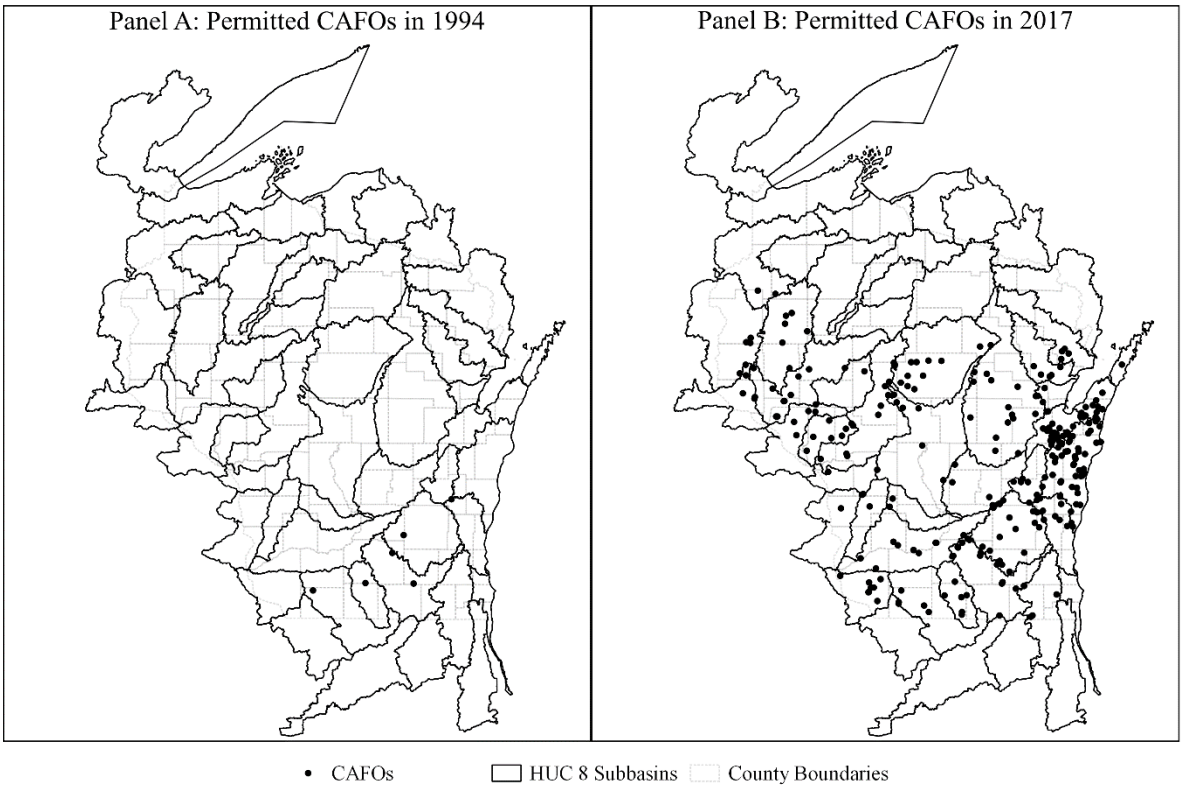



**Figure 1. Map of Wisconsin CAFOs in 1994 and 2017**

Notes: The panels of this figure show the HUC8 regions of Wisconsin (dark lines) overlaying maps of county borders (faded lines). Each black dot represents a permitted CAFO. Panel A shows permitted CAFOs that were operating in 1994 (the year prior to our sample period) and Panel B shows permitted CAFOs that were operating in 2017.

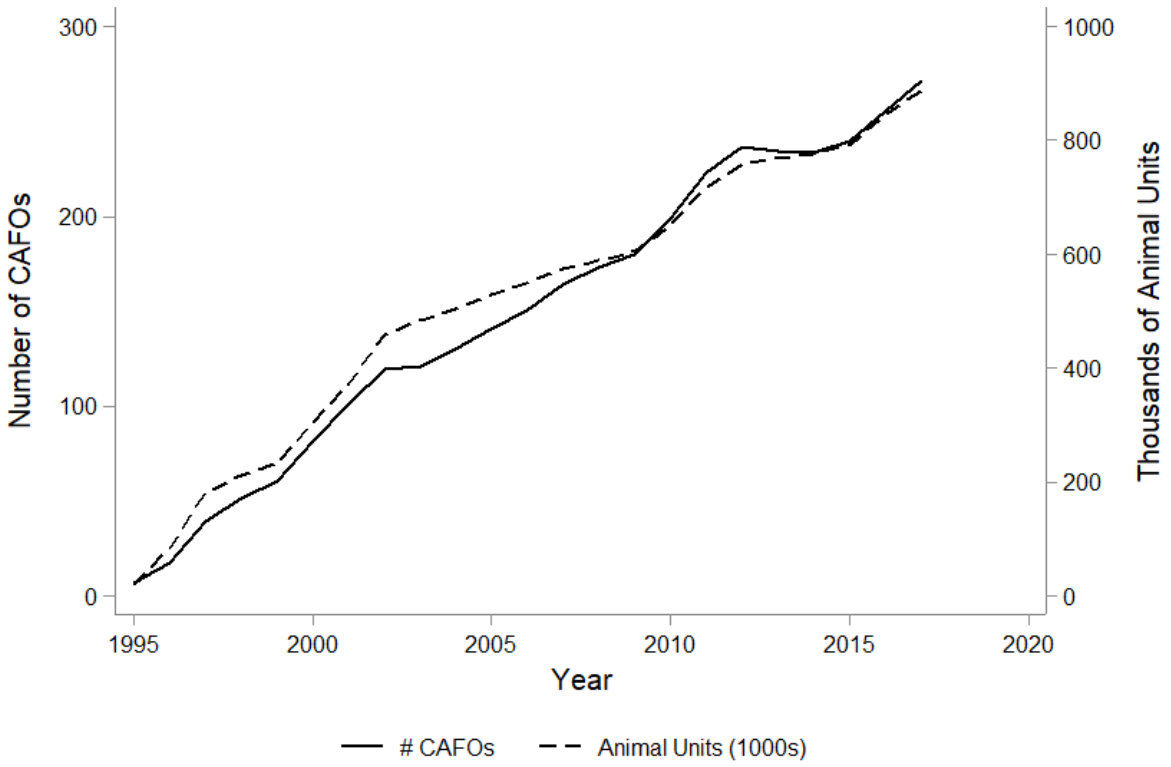



**Figure 2. CAFO expansion in Wisconsin**

Notes: This figure shows the total number of CAFOs and the total number of animal units in CAFOs for the state of Wisconsin from 1995 to 2017.

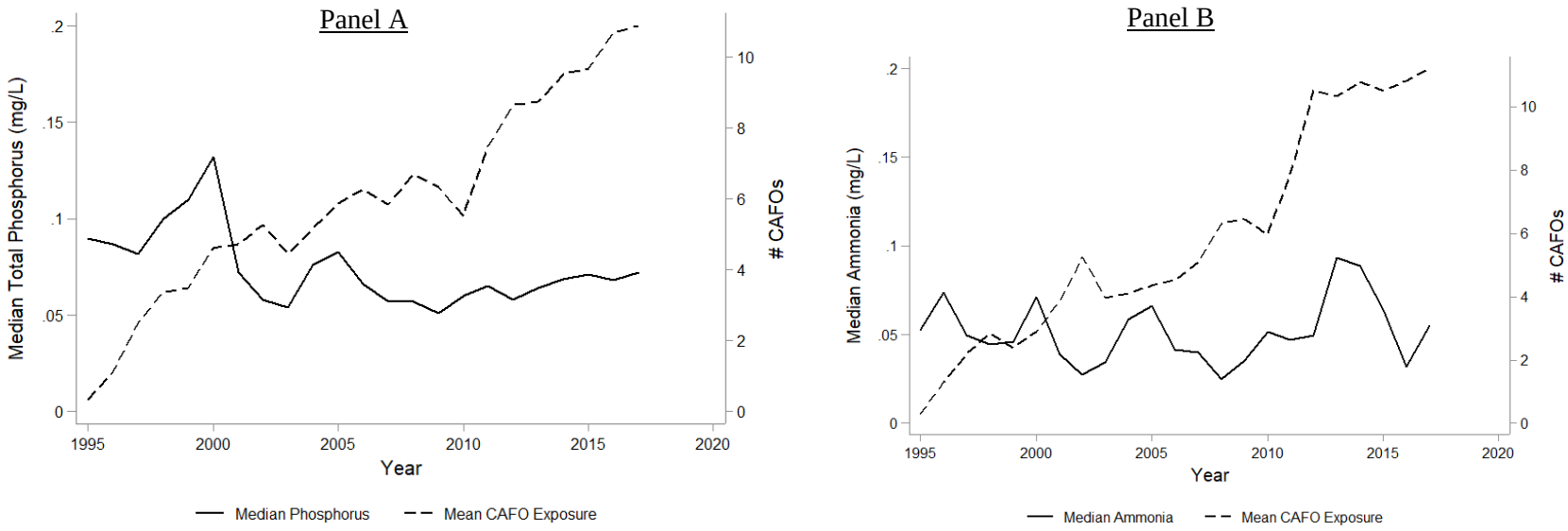



**Figure 3. Average total phosphorus and ammonia vs. CAFO exposure, 1995-2017**

Notes: The two panels of this figure show median annual total phosphorus (Panel A) and ammonia (Panel B) readings and the average number of CAFOs in the HUC8 region of a water quality monitor in our sample for the years 1995-2017.

**Table 1. Summary statistics for analysis samples**

| | Total phosphorus | | Ammonia | |
|---|---|---|---|---|
| Variable | Mean | SD | Mean | SD |
| Pollutant concentration (mg/L) | 0.258 | 0.720 | 0.227 | 0.591 |
| CAFOs | 6.387 | 6.942 | 6.075 | 6.745 |
| Total precipitation (cm) | 0.548 | 1.061 | 0.609 | 1.123 |
| Maximum daily temperature (°C) | 19.17 | 9.334 | 17.30 | 10.08 |
| Median income ($) | 48,197 | 8,806 | 48,075 | 8,835 |
| Unemployment rate | 5.538 | 1.932 | 5.361 | 1.891 |
| Farm acres | 221,002 | 110,020 | 234,096 | 105,038 |
| Acres spread with commercial fertilizer | 114,050 | 67,256 | 116,766 | 60,195 |
| Developed land (%) | 12.48 | 10.51 | 13.47 | 11.17 |
| Forested land (%) | 23.37 | 16.88 | 22.38 | 16.59 |
| Planted land (%) | 42.29 | 20.60 | 44.53 | 17.96 |
| Observations | 237,528 | | 108,577 | |

**Notes:** Summary statistics for the pollutant concentration, total precipitation, and temperature measures are at the individual reading level, i.e., monitor-day level. All other measures are at the HUC8-day level. HUC8 level measures are standardized to account for multi-state HUC8s.

**Table 2. Effects of CAFOs on surface water nutrient levels**

| | Panel A: Total phosphorus | | | Panel B: Ammonia | | |
|---|---|---|---|---|---|---|
| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
| CAFOs | 0.00523** | 0.00436*** | 0.00436*** | 0.00637** | 0.00614** | 0.00614** |
| | (0.00232) | (0.00135) | (0.0097)[+] | (0.00261) | (0.00246) | (0.0131)[+] |
| Time varying controls | | X | X | | X | X |
| Observations | 237,528 | 237,528 | 237,528 | 108,577 | 108,577 | 108,577 |
| Clustering | HUC8 | HUC8 | HUC8 and year | HUC8 | HUC8 | HUC8 and year |
| R-squared | 0.111 | 0.141 | 0.141 | 0.145 | 0.146 | 0.146 |

**Notes:** Each column presents regression results from a separate specification of equation (1). Each specification includes year, month, and HUC8 fixed effects. CAFOs is the treatment variable and measures the number of CAFOs within each HUC8. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level for columns 1, 2, 4, and 5. [+] Wild cluster bootstrapped p-value in parentheses are for two-way clustering on HUC8 and year in columns 3 and 6. *** p<0.01, ** p<0.05, * p<0.1.

**Table 3. Effect of CAFOs and precipitation on water quality sampling**

| Variable | Total phos-phorus (1) | Ammonia (2) | Total phos-phorus (3) | Ammonia (4) |
|---|---|---|---|---|
| CAFOs | -0.000019 | -0.000045 | -0.000036 | -0.000052 |
| | (0.000031) | (0.000047) | (0.000032) | (0.000045) |
| Total day of sample precipitation (cm) | 0.000210 | 0.000211 | | |
| | (0.000145) | (0.000138) | | |
| Total precipitation, 3-day average (cm) | | | 0.000029 | 0.000244 |
| | | | (0.000021) | (0.000185) |
| CAFOs × total day of sample precipitation (cm) | 0.000076 | 0.000031 | | |
| | (0.000049) | (0.000024) | | |
| CAFOs × total precipitation, 3-day average (cm) | | | 0.000144 | 0.000058 |
| | | | (0.000086) | (0.000043) |
| Observations | 90,596,384 | 57,933,296 | 90,596,384 | 57,933,296 |

**Notes:** Each column presents OLS regression results from a separate specification of equation (2). Column headings represent the type of pollutant sample examined. The dependent variable is a dummy indicating the presence of a water quality sample of total phosphorus or ammonia (including non-detects) at a given monitoring location on a given day. Each specification includes year, month, and HUC8 fixed effects. CAFOs is the treatment variable and measures the number of CAFOs within each HUC8. The total precipitation, three-day average variable measures the mean of precipitation of the day itself and the two preceding days at each monitoring location. Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Table 4. Data aggregated to HUC8-month level**

| | Panel A: Total phosphorus | | Panel B: Ammonia | |
|---|---|---|---|---|
| Variable | (1) | (2) | (3) | (4) |
| CAFOs | 0.00488** | 0.00433*** | 0.00638** | 0.00627** |
| | (0.00187) | (0.00128) | (0.00262) | (0.00243) |
| Time varying controls | | X | | X |
| Observations | 10,199 | 10,199 | 8,673 | 8,673 |

**Notes:** Each column presents regression results from a separate specification of equation (3) where data are aggregated to the HUC8-month level and the number of water quality readings are used as analytic weights. Each specification includes year, month, and HUC8 fixed effects. CAFOs is the treatment variable and measures the number of CAFOs within each HUC8. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Table 5. Placebo test: Pesticides**

| Variable | Concentrations (1) | Non-detect (2) |
|---|---|---|
| CAFOs | 0.00423 | -0.00179 |
| | (0.00617) | (0.00468) |
| Time varying controls | X | X |
| Observations | 131,831 | 131,831 |

**Notes:** Each column presents regression results from a separate specification of equation (1) with a placebo outcome. The outcome in column 1 is the monitor-day surface water concentration of pesticides (measured in mg/L). The outcome in column 2 is a dummy indicating a sample that does not measure a positive concentration of pesticides. Each specification includes year, month, and HUC8 fixed effects. CAFOs is the treatment variable and measures the number of CAFOs within each HUC8. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Table 6. Bounding set for estimates**

| Dep. variable | Analysis level | $\tilde{\beta}$ | $(R^2_{max} = 1.3\tilde{R})$ Bias-adjusted $\beta^*(1.3\tilde{R}, 1)$ | Robust to excluding 0 | $(R^2_{max} = 1.15\tilde{R})$ Bias-adjusted $\beta^*(1.15\tilde{R}, 1)$ | Robust to excluding 0 |
|---|---|---|---|---|---|---|
| Total phosphorus | Monitor reading | 0.00436*** | 0.403 | X | 0.0131 | X |
| Total phosphorus | HUC8-month | 0.00439*** | 0.0592 | X | 0.00858 | X |
| Ammonia | Monitor reading | 0.00614** | 0.147 | X | 0.0150 | X |
| Ammonia | HUC8-month | 0.00627** | 0.0187 | X | 0.00921 | X |

**Notes:** This table presents bounding sets for the coefficients on CAFOs from the estimation of equations (1) and (3). The column denoted $\tilde{\beta}$ shows the original estimates from the full models (columns 2 and 4 from Tables 2 and 5). The bias-adjusted estimates adjust for selection on unobservables using the methods of Oster (2019). We indicate whether the bounding set excludes 0 for each bias-adjusted estimate.

**Table 7. Primary sensitivity analyses**

| Variable | Total Phos-phorus (1) | Ammonia (2) | Total phos-phorus (3) | Ammonia (4) |
|---|---|---|---|---|
| CAFOs | 0.00620*** | 0.00682** | 0.00358** | 0.00601** |
| | (0.00134) | (0.00234) | (0.00148) | (0.00266) |
| Time varying controls | X | X | X | X |
| Non-CAFO animals | X | X | | |
| Legacy total phosphorus levels | | | X | |
| Legacy ammonia levels | | | | X |
| Observations | 237,528 | 108,577 | 237,503 | 108,531 |

**Notes:** Each column presents regression results from a separate specification of equation (1) with additional controls included. Each specification includes year, month, and HUC8 fixed effects. CAFOs is the treatment variable and measures the number of CAFOs within each HUC8. The non-CAFO animals measure represents the count of all animals at livestock operations net of those located on the CAFOs in our sample at the HUC8-day level. The legacy total phosphorus and ammonia level measures are the average HUC8-year concentrations of each respective pollutant over the previous five years. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

[1] An "animal unit" is the equivalent of 1,000 pounds live weight of animals.

[2] Phosphorus and ammonia (nitrogen) are two of the most prevalent pollutants found in animal waste (EPA 2000b; EPA 2002a,b; EPA 2004; WDNR 2011). Total phosphorus is a measure of all forms of phosphorus in surface water-bodies (e.g., orthophosphate, organic phosphate) and the Wisconsin Department of Natural Resources reports it as such. Ammonia concentrations, however, are reported in two forms: (1) ammonia (NH3) and (2) ammonia as nitrogen (NH3-N). Like other studies (e.g., Chen et al. 2019), we convert ammonia to ammonia as nitrogen using the following weight conversion: NH3=NH3-N*1.12589.

[3] Surface waterbody concentrations of nutrients such as phosphorus and ammonia can be linked to physical and economic impacts in several ways. We illustrate these physical linkages in the "Economic impacts" sub-section, using a water quality index approach, and in Online Appendix F, using a trophic state index.

[4] O'Connor et al. (2010) systematically review this literature.

[5] Most like our study, Chen et al. (2019) examine water pollution from CAFOs before and after the implementation of an EPA regulation that requires CAFO permitting under the CWA. Our study differs from Chen et al. (2019) in several ways. First, we estimate the effect on water quality of CAFOs themselves, specifically the increase in the number of CAFOs (a trend that is likely to continue), while Chen et al. (2019) perform a policy analysis of an EPA rule implemented to curb water pollution from CAFOs. This difference is critical because CAFO intensity has increased substantially since the 2003 rule. Thus, our study reveals the effects of increasing CAFO intensity even in the presence of permitting requirements. Second, we monetize the changes in water quality that result from CAFO exposure, whereas Chen et al. (2019) look exclusively at the water quality impacts of the policy. Finally, Chen et al. (2019) examine ammonia as the outcome of interest and we examine ammonia and total phosphorus as outcomes.

[6] See Gray and Shimshack (2011) for a thorough review of the environmental compliance literature, which studies point source water pollution in depth.

[7] Algal blooms lead to reduced recreation (Zhang and Sohngen 2018), low fish yields (Anderson et al. 2000), decreased property values (Wolf and Klaiber 2017), and potential adverse health effects (Townsend et al. 2003).

[8] Several press outlets have covered the prevalence of algal blooms in the presence of climate change in detail, see, e.g., https://www.nytimes.com/2018/07/09/us/algae-blooms-florida-nyt.html.

[9] There are 2,264 HUC8 regions in the United States (50 in Wisconsin).

[10] EPA (2001) refers to manure collectively as manure, litter, and process wastewater.

[11] As one exception, Section 319 of the Clean Water Act provides grants for non-point source pollution protection of watersheds.

[12] The permit requirements of the 2003 EPA rule consider spreading manure on fields as discharge events and extend to all "large" CAFOs, which are those with over 1,000 animal units. Prior to 2008, these requirements could be waived only if the CAFO could prove that it did not have the potential to discharge. The 2003 rule was challenged in court so EPA revised regulations in 2008 to only require a NPDES permit if the CAFO discharged or proposed to discharge pollutants. Liability still existed for failing to apply for a permit (Moore 2011). A March 15, 2011 ruling further reduced the permit requirements for non-discharging CAFOs, vacating the 2008 EPA rule that required proposed dischargers to apply for a NPDES permit. The 2011 ruling also vacated the earlier liability provisions for failure to apply for a NPDES permit. "However, the court upheld the requirement that CAFOs discharging manure to waters of the United States are required to apply for a NPDES permit, and upheld the ability to regulate CAFO land application of manure." (Moore 2011). Wisconsin Rule NR243 continues to require all Wisconsin CAFOs to obtain a permit, regardless of discharge status. Additionally, AFOs with confined livestock can be considered "small" or "medium" CAFOs if they contain less than 1,000 animal units and have the potential to discharge waste to surface waterbodies. These operations must then obtain NPDES permits.

[13] Only Massachusetts, New Hampshire, New Mexico, and Idaho are not authorized to implement the NPDES program.

[14] The summary document is available at https://www.epa.gov/sites/production/files/2018-05/documents/track-sum_endyear_2017.pdf. NPDES CAFO regulations require all discharging CAFOs to obtain permits. However, these regulations are size based. The summary document provides numbers for those AFOs that exceed the 1,000-animal unit threshold necessary for CAFO consideration. See above for further details.

[15] We discuss this data source in greater detail below.

[16] For nutrient readings per 100 square miles (rather than readings for all pollutants), these values are 3.63 for Wisconsin and 0.75 for other states.

[17] We define the corn belt as those states identified by Green et al. (2018): South Dakota, Nebraska, Minnesota, Iowa, Wisconsin, Illinois, Indiana, and Ohio.

[18] Livestock operations typically store cow and swine manure in liquified form, so it is more expensive and difficult to transport relative to poultry manure, which livestock operations typically store in dried form.

[19] We also estimate our primary specification (equation (1)) using HUC10 level variation and fixed effects. Point estimates on our main coefficients of interest (marginal effects of CAFOs on nutrient levels) resemble those reported in this article (in sign and magnitude) and are statistically significant at the 10% level. There are 357 HUC10 regions in Wisconsin so we have less water quality variation within HUC10 regions. It is also possible that HUC10 areas are small enough that "treated" areas spill into neighboring "control" areas because of manure spreading practices.

[20] We use daily water quality readings matched with daily CAFO intensity (from CAFO permit data) as our baseline. However, we also show that our results are robust to aggregating to monthly observations.

[21] For each eventually interpolated measure, we create a balanced daily panel for the entirety of our sample period. Then, we fill the last day of each year with the yearly (or bi-yearly, etc.) value for that variable, which is measured. We then linearly interpolate the missing values between the two known values. For this linear interpolation, we assume a constant rate of change between the two known values. We acknowledge that the assumption of a constant rate of change for each day between the two known end points is unlikely to perfectly hold, but that these approximations are reasonable in our case.

[22] We begin our analysis in 1995 due to the paucity of data available between 1990 and 1994. There are only six permitted CAFOs in Wisconsin for these years and the CAFO data do not contain information on several key measures like location and animal units. Furthermore, water quality sampling was apparently not as widespread prior to 1995 in Wisconsin. For example, there are only around 5,000 total phosphorus readings from Wisconsin in 1990 as compared to the annual sample average of around 10,000 (1995-2017).

[23] The discrepancy between the number of CAFOs in our dataset and the number of permitted CAFOs in EPA's summary document is because of a single case where the WDNR considers more than 20 CAFOs under a single "general" permit while EPA considers these individual CAFOs.

[24] Figure 1 shows that CAFOs and animal units almost perfectly correlate over time. The pairwise correlation coefficient is 0.830 for the total phosphorus sample and 0.861 for the ammonia sample. We therefore use the number of CAFOs as our primary independent variable for our analysis. Online Appendix Table A3 presents results using animal units in place of CAFOs; results are consistent with those presented for CAFOs.

[25] These sources are aggregated within the publicly available Water Quality Portal; see https://www.waterqualitydata.us/portal/ for details. NWIS contains historical water quality data on surface water and groundwater from over 1.5 million sites throughout the United States. STORET contains similar water quality information dating back to the

1960s. Biodata contains bioassessment data collected by local, regional, and national USGS projects.

[26] For pollutant concentration readings on the left tail of the distribution, we follow the literature and transform measurements of zero and non-detects as positive values (e.g., Keiser and Shapiro 2018; Chen et al. 2019). Here, we replace zero readings and non-detects with ½ of the smallest positive value in each sample.

[27] Schlenker and Roberts (2009) use the original data and the lead author maintains updated records on his personal website (Schlenker and Roberts 2020). For a more detailed description of these data, see http://www.wolframschlenker.com/dailyData/dataDescription.pdf.

[28] Paudel and Crago (2020) estimate the effect of fertilizer usage on water quality. The authors use a measure of fertilizer usage purchased from the Association of American Plant Food Control Officials. Because we only include fertilizer usage as a control and are not interested in the causal effect of fertilizer usage on surface water quality, we use the data measure that is publicly available.

[29] We examine these concentrations in terms of water quality indices in the following sub-sections.

[30] The average monitoring location for each analysis sample is exposed to roughly 26,000 CAFO animal units. This value suggests that most CAFOs contain more than the 1,000 animal units that permits require.

[31] During our sample period, there were three policy changes that affect surface water quality and CAFO management. First, at the state level, NR151 sets agricultural performance standards and ATCP50 guides how farmers meet the performance standards. Second, the EPA CAFO rule implemented in 2003 required that more CAFOs receive NPDES permits and for CAFOs to submit nutrient management plans; this rule was updated in 2008 (Sneeringer et al. 2011; Chen et al. 2019). Changes to these regulations that cause changes in nutrient concentrations and that all HUC8s in our sample experience are absorbed by the year fixed effects.

[32] We do not include in our primary regression specification controls for two measures that could affect the magnitude of our estimates because they could be endogenous. First, it is possible that the legacy nutrients present in surface waterbodies affect nutrient concentrations. Thus, one could model both the stock and the flow of pollutants in hydrological networks. However, inclusion of a lagged dependent variable or a series of lags of CAFO exposure likely introduces bias into our specification (Nickell 1981; Paudel and Crago 2020). Second, the purpose of our primary regression specification is to estimate the effect of CAFOs on surface water quality, independent of other agricultural and livestock operations. As such, we include control factors for agricultural land use and land cover. One could also argue that we should control for the total number of non-CAFO animals in a HUC8 region (animals at AFOs other

than CAFOs) because this measure could be correlated with CAFO intensity and affect nutrient levels. However, the number of non-CAFO animals could also be an outcome of the CAFO treatment, in which case we would not want to control for it in our regression. Nonetheless, we estimate equation (1) with these additional controls as robustness checks in the "Sensitivity analysis" sub-section. The inclusion of these controls does not meaningfully change our estimates of treatment effects.

[33] We tabulate estimation results for the control factors used in our primary empirical specification in Online Appendix Table A2. We include these controls to more precisely estimate the primary regressor coefficients and to mitigate omitted variable bias. Given this purpose, we do not claim that the estimated coefficients on controls represent causal effects on our outcomes. Nevertheless, the control factor coefficients are all reasonable in magnitude.

[34] Equation (1) includes year fixed effects, so one-way clustering at the HUC8 level should be sufficient for common time shocks. However, two-way clustering could be appropriate if there are heterogeneous effects of state level time shocks (Petersen 2009; Cameron et al. 2011). Furthermore, we have a relatively small number of clusters in the year dimension (23), so we use the wild cluster bootstrap (Cameron et al. 2008). We use Roodman's boottest module within Stata (Roodman 2015; Roodman et al. 2019). As recommended by MacKinnon et al. (2019), we cluster on both dimensions and bootstrap along the dimension with the smallest number of clusters (year).

[35] We also calculate WTP for a counterfactual scenario with zero CAFOs in Wisconsin in Online Appendix E.

[36] TSS subindex water quality calculations are EPA ecoregion specific. We use GIS data from EPA (https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states) to determine the appropriate EPA Level 3 ecoregion for each HUC8 (often a weighted average of multiple ecoregions because most HUC8 regions fall into multiple ecoregions). We then apply the ecoregion specific TSS subindex curve parameters given in Appendix F of EPA (2009) to calculate the TSS subindex for each HUC8.

[37] One could set other water quality measures at some historical or average level as in Corona et al. (2020). However, these other water quality measures are likely highly correlated with changes in nitrogen and phosphorus, so we would be understating water quality changes if we were to hold them constant. For example, Griffin et al. (2020) note that "dissolved oxygen and biochemical oxygen demand are positively correlated with nitrogen and phosphorus through their role in eutrophication, respiration, and decomposition (Heiskary and Markus 2001; Prasad et al. 2011) and leaving these at their sample means may bias results toward higher water quality under increased nitrogen and phosphorus scenarios."

[38] For example, EPA (2009, page G-7) states "there is, however, some controversy regarding whether regression-based meta-analyses should be used for direct benefit transfer." Studies such as Shrestha et al. (2007) show that meta-analyses do not necessarily produce more reliable estimates. Also, regression-based meta-analysis transfer functions such as EPA (2009) require researchers to assign a value to each model variable and the estimated WTP is often sensitive to these variable assignment choices. For example, Johnston et al. (2006) find "specification of levels for study methodology variables have significant implications for WTP over the entire range of quality change."

[39] Appendix G of EPA (2009) provides complete details. We assign independent variables in the benefit transfer function as described in EPA (2009), with one exception. We set "mail" to 1 because contemporary literature suggests that mail surveys are likely more reliable than telephone surveys (which is the default category when "mail" is set to 0). As discussed in Johnston et al. (2017, pg. 340), "There are advantages and disadvantages to all survey administration modes. However, given the inability to effectively convey complex valuation materials in a telephone interview, this survey mode should be used with caution." Other references for this approach include Bateman and Jones (2003), Johnston et al. (2005; 2006), Shrestha et al. (2007) and Rosenberger and Phipps (2007).

[40] The WQI depends on total nitrogen levels, which includes ammonia. We do not have total nitrogen measures for each HUC8-year, so we estimate total nitrogen as a function of ammonia using historical USGS data on total nitrogen. $\widehat{N} = 0.865 + 7.094 * ammonia$ (p-value on ammonia=0.00015). Our approach follows that of Griffin et al. (2020) for estimating levels of correlated pollutants when data are missing. An alternative approach that holds non-ammonia nitrogen constant results in only slightly smaller WTP values.

[41] These surface water quality measures are consistent with Wisconsin's most recent water quality report to Congress. In that report, WDNR reports that 82% of assessed waterbodies are healthy. A waterbody is considered healthy if it meets at least one designated use (recreation, aquatic life, or fish consumption) and is not impaired for any use (WDNR 2019).

[42] The 95% confidence interval for the Carson and Mitchell (1993) WTP function is \$35.42-\$102.23 and the 95% confidence interval for the EPA (2009) meta-analysis WTP function is \$210.12-\$323.23.

[43] This calculation takes the \$6.9 million – \$27.9 million WTP to prevent a marginal CAFO in each HUC8 and divides that value by the number of HUC8 regions with CAFOs (34).

[44] Improving the water quality of the 128 lakes in Iowa to the level of West Okoboji Lake requires an average decrease of total phosphorus concentrations of over 0.085 mg/L at each lake. In our counterfactual scenario, the difference in average total phosphorus concentrations is 0.004 mg/L.

[45] We also examine if water quality monitoring locations are endogenously determined with respect to CAFO locations. Online Appendix Table A4 shows that the number of CAFOs within a HUC8 does not significantly affect the number of water quality monitoring locations in that HUC8, for each analysis sample. Additionally, Online Appendix Figure A6 maps the surface waterbody monitoring locations in Wisconsin with at least one reading during our sample period. The figure shows that there are no systematic spatial patterns in water quality monitoring locations during our sample period. We are therefore confident that surface water quality monitoring locations and CAFO locations are not endogenously determined.

[46] Keiser and Shapiro (2019) similarly aggregate and weight by the number of underlying pollution readings. Results without using analytic weights are qualitatively similar and we tabulate them in Online Appendix Table A6.

[47] We winsorize readings at the 99% level, replace zero readings and non-detects (for the specification where concentration is the outcome) with ½ the smallest positive value, and measure pesticide concentrations in mg/L.

[48] As one example, there exists bunching at the CAFO size threshold (Sneeringer and Key 2011) so AFOs with hundreds of animal units may affect surface water nutrient concentrations in ways like CAFOs.

[49] Differences in how CAFOs and non-CAFOs manage and store their manure can also have water quality implications. For example, if CAFOs store their manure in unlined lagoons and non-CAFOs ship their manure offsite, then we could better identify the cause of the non-point source pollution. However, there are no systematic differences in the ways that differently sized livestock operations manage or store their manure (although CAFOs produce much more manure than non-CAFOs). Both types of operations store manure onsite, either in lagoons, pits, or piles, and spread excess manure onto agricultural fields (Meyer et al. 2011; DeRouchey 2014). CAFOs and non-CAFOs alike rarely transport their manure far offsite because of the prohibitive costs (Ali et al. 2012). As a result, we are unable to exploit heterogeneity in manure management practices to help disentangle the effects of CAFOs and non-CAFOs on surface water quality.

[50] The damage values provided from air pollution most likely include these costs because nearly all the damages of air pollution are health related (EPA 2011).

[51] EPA (2002b) estimates that non-market surface water quality benefits comprise around 75-95% of the benefits of revised CAFO regulations, so our estimate likely captures most external costs of CAFOs related to surface water.

# Appendices (for online publication only)

## A. Additional data description

In this online appendix, we provide additional descriptive details about our analysis samples. We discuss descriptive differences in nutrient concentrations by CAFO exposure and provide descriptive graphical evidence of correlations between CAFO exposure and nutrient concentrations.

Online Appendix Table A1 shows the heterogeneity in descriptive statistics by CAFO exposure. Mean total phosphorus readings for monitoring locations in HUC8 regions without any CAFO exposure is 0.172 mg/L, while the mean level for monitoring locations with at least one CAFO in the same HUC8 is 0.283 mg/L; this univariate difference of means is statistically significant (p=0.000). For ammonia readings with and without any CAFO exposure, the mean values are 0.212 mg/L and 0.230 mg/L, respectively; also a significant difference (p=0.00).

In Online Appendix Figure A1, we split our total phosphorus sample into three groups based upon the highest number of CAFOs that ultimately locate in the HUC8 region by the end of the panel. Group 1 represents the first quartile with the lowest eventual CAFO intensity, group 2 includes water quality monitoring locations from HUC8 regions having eventual CAFO intensity between the 25$^{th}$ and 75$^{th}$ percentile, and group 3 includes water quality monitoring locations from HUC8 regions with the highest eventual CAFO intensity (above the 75$^{th}$ percentile). Online Appendix Figure A2 does the same for ammonia. To show how these water quality trends correlate with treatment intensity within each treatment group, we superimpose average CAFO exposure by treatment group. Collectively, these figures provide evidence that CAFO exposure correlates with nutrient concentrations in surface waterbodies during our sample period.

## B. CAFO animal units as treatment

In this online appendix, we re-estimate equation (1) replacing the number of CAFOs with an alternative treatment variable of the number of CAFO animal units, also measured at the HUC8 level. All controls and specifications are identical to those discussed for Table 2 in the main text. Appendix Table A3 presents the results for this estimation, which are consistent with those in Table 2 of the main text; an increase of 1,000 CAFO animal units per HUC8 leads to an increase in total phosphorus levels of approximately 0.0013 mg/L. There are 2.99 thousand animal units on an average CAFO in our total phosphorus sample. Therefore, the addition of average CAFO exposure to a HUC8 region would result in an approximate increase in total phosphorus levels in surface waterbodies of 0.00389 mg/L, about 89% of the effect size found in Table 2 of the main text. For ammonia, we similarly calculate an effect that is about 140% of the effect found in Table 2 of the main text (0.003*2.85=0.00855 increase in mg/L from an increase in animal units on an average CAFO). Therefore, results from specifications with CAFO animal units broadly agree with our baseline results that use the number of CAFOs as the primary regressor.

## C. Evidence to support identifying assumptions

In this online appendix, we provide evidence to support our key identifying assumption of parallel trends in potential outcomes. This assumption is fundamentally untestable but finding parallel trends in nutrient concentrations in the years leading up to various levels of CAFO treatment provides some empirical support to the assumption. We first present a specification that, like our

primary specification in the main text, allows for varying intensity of CAFO treatment but tests for pre-trends in the outcome variables. We then conduct event studies where we define a dichotomous "treatment" at various CAFO intensity thresholds.

Modifying our primary specification (1) to include pre-treatment indicators in the spirit of an event study, the estimating equation becomes:

$$Y_{ijdmt} = \sum_{k=-4}^{-1} \tau_k I_j\left(t - c_j = k\right) + \beta_1 CAFO_{jdmt} + \beta_2 M_{ijdmt} + \beta_3 X_{jdmt} + \gamma_j + \psi_m + \lambda_t + \varepsilon_{ijdmt}, \quad \text{(A1)}$$

where notation is the same as described for equation (1). In this specification, $c_j$ denotes the year that the first CAFO was issued a permit in monitor *i*'s HUC8 region *j*. Thus, $I_j\left(t - c_j = k\right)$ are event indicators, equal to 1 when the year of observation is *k* years before the first permitted CAFO begins operation in HUC8 region *j*. We group years more than four years before the first permitted CAFO into the *k*=-4 indicator.[1] The coefficients, $\tau_k$, represent the evolution in nutrient concentrations at eventually treated monitors before the first CAFO arrived relative to changes in untreated monitors net adjustments in model covariates.

Online Appendix Table A5 shows that the pre-treatment indicators are never statistically significant for either outcome. Furthermore, the point estimates on these pre-treatment indicators do not reveal systematic patterns in pre-treatment trends and are close to zero. The point estimate for the coefficient on CAFO is also like the baseline estimate in Table 2 and remains statistically significant. This specification suggests that there are no noticeable pre-trends in nutrient concentrations before the arrival of CAFOs to surface waterbody monitoring locations.

We next specify event studies where treatment is successively defined at the following CAFO levels: one (first CAFO), median CAFO intensity (three CAFOs for both total phosphorus and ammonia), and 75$^{th}$ percentile CAFO intensity (nine CAFOs for total phosphorus and eight CAFOs for ammonia). For each treatment level, the estimation equation is:

$$Y_{ijdmt} = \sum_{k=-4}^{5} \tau_k I_j\left(t - c_j = k\right) + \beta_1 M_{ijdmt} + \beta_2 X_{jdmt} + \gamma_j + \psi_m + \lambda_t + \varepsilon_{ijdmt}, \quad \text{(A2)}$$

where notation is the same as described above for equation (A1), except $I_j\left(t - c_j = k\right)$ are now event indicators, equal to 1 when the year of observation is *k* years before or after the year in which CAFO intensity crosses a certain threshold in HUC8 region *j*. We again group years more than four years before the first year of dichotomous treatment into the *k*=-4 indicator and we group years more than five years after the first year of dichotomous treatment into the *k*=5 indicator.

In Appendix Figures A3 (total phosphorus) and A4 (ammonia), we present event studies for each dichotomous CAFO treatment level. We do not see any evidence of pre-trends for any of the treatment levels in Figure A3. Figure A3 suggests that the effects of CAFOs on phosphorus levels are concentrated in the top quartile of treatment intensity. Likewise, we see little evidence of pre-trends for the outcome of ammonia in Figure A4. Again, the effects of CAFOs on ammonia appear to be driven by higher levels of CAFO treatment.

[1] We use four years of pre-treatment indicators because CAFO permits significantly expanded in 1999 (WDNR 2011), which is four years after the start of our panel. Thus, including pre-treatment indicators greater than four years in this specification significantly reduces the statistical power for those years.

## D. Long-difference instrumental variables estimation

In this online appendix, we provide an instrumental variables (IV) estimation strategy to account for the potentiality of endogenous CAFO siting. We can estimate equation (1) at the aggregated HUC8-year level rather than at the individual monitoring location level. The estimating equation then becomes:

$$P_{jt} = \beta_1 CAFO_{jt} + \beta_2 X_{jt} + \gamma_j + \psi_t + \varepsilon_{jt}, \tag{A3}$$

where $P_{jt}$ is the annual average nutrient (total phosphorus or ammonia) reading of all samples taken from all monitoring locations in a given HUC8 region. Some HUC8 regions have many more underlying water quality samples available in a given year than others and thus provide a more reliable measure of true water quality. We therefore use the number of nutrient samples in a HUC8 region in a given year as analytic weights in all regressions at the HUC8 level. Equation (A3) is useful in motivating our second estimation strategy, IV.

Online Appendix Figure A5 presents a map of CAFO locations in Wisconsin. This map shows that CAFOs tend to locate in areas with U.S. highway and interstate (hereafter "freeway") access. This location preference is intuitive because CAFOs need to transport large quantities of a marketable product, such as the animals themselves (for meat producers) or milk (for dairy operations, which are especially prevalent in Wisconsin). Additionally, CAFOs receive frequent deliveries of food for the livestock. This observation suggests that CAFO growth could be explained at least partially by freeway access. Such freeway access in Wisconsin does not measurably change within our sample period. Relatedly, it is not possible to use a time-invariant instrument in equation (A3) and include HUC8 fixed effects to control for time-invariant unobservables. However, it is possible to use a time-invariant instrument in a first difference model with two time periods.

In our long-difference approach, we create two time periods that correspond to the beginning and end of our sample period. Rather than simply taking 1995 and 2017 (the first and last years of our panel), we average over the first five years and last five years of the panel. Thus, the first time period contains averages for the five years centered around 1997 and the second time period contains averages for the five years centered around 2015. Our cross-sectional regression specifications then become:

$$P_{j1997} = \beta_0 + \beta_1 CAFO_{j1997} + \beta_2 X_{j1997} + \gamma_j + \psi_{1997} + \varepsilon_{j1997} \tag{A4}$$

and

$$P_{j2015} = \beta_0 + \beta_1 CAFO_{j2015} + \beta_2 X_{j2015} + \gamma_j + \psi_{2015} + \varepsilon_{j2015}. \tag{A5}$$

First-differencing absorbs the HUC8 specific fixed effect and we are left with

$$\Delta P_j = \beta_1 \Delta CAFO_j + \beta_2 \Delta X_j + (\psi_{2015-}\psi_{1997}) + (\varepsilon_{j2015} - \varepsilon_{j1997}), \tag{A6}$$

where all notation is identical to that of equations (1) and (A3). OLS estimation of equation (A6) is equivalent to the estimation of equation (A3) with two time periods. Our estimate of $\beta_1$ is biased if $\Delta CAFO_j$ and $\Delta\varepsilon_j$ are correlated. That is, if CAFO expansion is correlated with other unobserved factors that affect changes in nutrient readings, our estimate of the causal effect of CAFO

expansion is biased. We therefore instrument for changes in the number of CAFOs with a measure of freeway access within a HUC8 region. Specifically, we calculate the number of freeway miles within each HUC8 region using the U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER)/Line shapefiles.[2] This estimation strategy – instrumenting for a long-difference variable of interest – is similar in spirit to the estimation strategy of Chay and Greenstone (2005). The authors instrument for changes in county level air pollution changes with mid-decade nonattainment status, which is time invariant.

Our instrument is valid if it is correlated with the change in CAFOs but has no direct effect on changes in surface water quality. It is apparent from Online Appendix Figure A5 that CAFOs substantially expanded in areas that are close to freeways, suggesting that the instrument is relevant; this is confirmed empirically by first stage F-stats presented in Online Appendix Table A7. Although the exclusion restriction is fundamentally untestable, we argue that it is unlikely that freeway miles have a direct effect on changes in nutrient readings outside of the way that freeway miles facilitate the siting of CAFOs, which in turn affects nutrient levels. It is possible that other types of agriculture outside of CAFOs, e.g., crop farming, could also benefit from access to freeways and this non-CAFO agricultural growth could also affect surface waterbody nutrient levels. However, we include in our analysis several controls for other non-CAFO agriculture. Specifically, equation (A6) contains controls for farm acreage and the number of acres with commercial fertilizer spread, both measured at the HUC8 level. Likewise, freeway access may contribute to population growth and associated increases in nutrient discharges from wastewater treatment plants or other types of urban runoff. We control for land cover (percentage of the HUC8 that is planted, forested, or developed) to alleviate this concern. Regardless, it is possible that this instrument may not be valid if these controls do not sufficiently alleviate the correlation between non-CAFO agriculture (or other factors) and freeway access. As such, we only provide these IV estimation results as support of our primary estimation results.

Online Appendix Table A7 presents estimation results from the long difference specification of equation (A6). Panel A shows the effect of a change in the number of CAFOs on changes in nutrient levels and Panel B shows the effect of a change in the number of CAFO animal units on changes in nutrient levels. Columns 1 and 3 show OLS results for this long difference specification; estimated coefficients for these OLS regressions are highly significant and approximately twice the magnitude of their counterparts in columns 2 and 4 of Table 2 and Online Appendix Table A3.

Column 2 of Online Appendix Table A7 shows estimation results when we instrument for the change in the number of CAFOs with kilometers of freeway miles in the HUC8 region. The IV estimates are again qualitatively consistent with the baseline difference-in-differences estimates, although larger in magnitude.[3] The first stage effect of freeway miles on CAFO growth is strong in each case; the F-stat on the excluded instrument is always well above 10, which shows that the probability of CAFOs locating in a HUC8 increases with better access to freeways. Similarly, column 4 shows that there is a strong first stage effect of freeway miles on CAFO animal unit growth and the estimated second stage effect qualitatively agrees with baseline estimates.

### E. Welfare estimates for a counterfactual of zero CAFOs

[2] The Tiger shapefiles group together U.S. highways and interstate highways. We refer to these as freeways.

[3] Under certain assumptions, the IV estimate with a continuous instrument such as freeway corresponds to a weighted average of local average treatment effects (Heckman and Vytlacil 2005; Imbens and Wooldridge 2009; Cornelissen et al. 2016).

In the main text, we estimate the willingness to pay for improved water quality associated with one fewer CAFO in each HUC8 for each year in which the HUC8 had at least one CAFO. In this online appendix, we conduct the analogous exercise for a counterfactual of Wisconsin never having any CAFOs. This exercise should be cautiously interpreted because we are working in a partial equilibrium framework without considering any possible general equilibrium effects. This analysis also assumes that the marginal impact of a CAFO on nutrient concentration remains constant over the sample period of 1995-2017. We use the same methodology as described in the main text with one exception; we use our empirical model to simulate total phosphorus and nitrogen levels at each water quality monitor in the counterfactual world of zero CAFOs (rather than a marginal CAFO in each HUC8 as in the main text).

Using the Carson and Mitchell (1993) benefit transfer function, we find the average household in Wisconsin would be willing to pay a present discounted value of $ 299.40 (in 2017 dollars) for the improved water quality from a counterfactual world with no CAFOs (95% confidence interval of $170.15-$403.87). For the EPA (2009) meta-analysis benefit transfer function, we find an analogous average WTP of $461.05 (95% confidence interval of $361.77-$525.40).

**F. Total phosphorus results and the Trophic State Index (Carlson 1977)**

Within the main text, we present the effects of CAFOs on surface water quality in Wisconsin in terms of mg/L, water quality indices (Cude 1974; Vaughan 1986; EPA 2009), and lost non-market water quality benefits. In this online appendix, we provide greater context for the total phosphorus results by using another index, the Trophic State Index (TSI) (Carlson 1977). The TSI is one of three primary measures that WDNR uses when performing its surface water quality "General Condition Assessment" (WDNR 2019), i.e., its required biennial Clean Water Act 305(b) assessment report. The TSI is a scale from 0-100+ and measures the level of "productivity", i.e., the generation of algal biomass, in surface waterbodies using secchi depth, chlorophyll levels, or total phosphorus levels. The scale of the TSI is logarithmic and each level of 10 (e.g., from 10 to 20) represents a doubling of algal biomass in that surface waterbody (Carlson 1977). Although some level of algal biomass is necessary for plant and aquatic life (Egan et al. 2009; WDNR 2019), high levels of algal biomass lead to reduced visibility and toxic algal blooms. Within the TSI, the categories of lake productivity range from <30 (oligotrophic), which are very clear waterbodies with little algal biomass, to 70+ (hypereutrophic), which are dark waterbodies with much algal growth and the propensity for harmful algal blooms. Finally, in its assessment, WDNR (2019) classifies shallow surface water (which is the majority of Wisconsin surface waterbodies) quality as the following, based on TSI: >70=poor, 62-70=fair, 53-61=good, <53=excellent. "Poor" waterbodies are those that do not meet the CWA's designated uses (fishable, swimmable, boatable) and are listed as impaired on the state's Clean Water Act Section 303(d) list.

We use total phosphorus concentrations and the coefficients from our primary regression specification to calculate TSI values for each monitoring location in Wisconsin for the current scenario with CAFO expansion and for the counterfactual world with no CAFOs. We then take these TSI values and aggregate them to calculate an average TSI value for the state in each scenario The current average TSI throughout the state is 65.1, which indicates eutrophic waters in "fair" condition. (This value matches the most recent assessment of the state's waterbodies (WDNR 2019), which mentions that "very few" Wisconsin waterbodies would fall in the oligotrophic range.) In the counterfactual world with no CAFOs, this value drops to 59.6, which is still

considered eutrophic, but waterbodies with this TSI are considered "good" by WDNR.[4] Additionally, a waterbody with a TSI value of 65.1 contains 1.5 times the amount of algal biomass as a waterbody with a TSI of 59.6, which considerably increases the potential of harmful algal blooms. In terms of water quality standards, neither value indicates that the mean waterbody in the state is impaired. However, "fair" waterbodies are flagged as approaching impaired status (WDNR 2019). Thus, the change from "fair" to "good" is economically meaningful considering the state's water quality standards and the amount of algal biomass in the average waterbody.

**G. Sensitivity analysis**

In this online appendix, we report additional estimation results from alternative specifications and sub-samples. We first discuss the alternative regression specifications. These additional specifications include: 1) adding year-by-HUC2 fixed effects (rather than only year fixed effects), 2) using HUC10 level fixed effects (rather than HUC8 fixed effects), 3) adding a variable to control for HUC8 level land in the Conservation Reserve Program (CRP), 4) examining the effects of upstream CAFO presence and intensity, and 5) allowing for differential trends before and after CAFOs enter a HUC8 watershed. We examine each in turn.

First, HUC2 regions represent the largest watershed categorization and span on average 175,000 square miles. In Wisconsin, there are only two HUC2s: 1) the Upper Mississippi Region and 2) the Great Lakes Region. Year-by-HUC2 fixed effects control for differential time effects in each HUC2, e.g., a cleanup initiative in the Upper Mississippi Region during a portion of the sample period. Second, HUC10 regions are smaller hydrological regions than HUC8s and span on average 225 square miles. A given HUC10 is a sub-region of some HUC8 so inclusion of these fixed effects provides a finer grain control for time-invariant unobservable characteristics than HUC8 fixed effects. The first and sixth and second and seventh columns of Online Appendix Table A9 show that the estimation of equation (1) with year-by-HUC2 and HUC10 fixed effects, respectively, produces results that are nearly identical to those of our primary empirical specification; the point estimates differ only slightly from their baseline counterparts in Table 2 and the estimates remain statistically significant. Third, we account for the possible effects of the Conservation Reserve Program (CRP) on surface water quality in Wisconsin during our sample period. The program, which USDA administers, aims to decrease non-point source runoff by making payments to farmers for taking cropland out of use. Thus, we re-estimate our primary regression specification of equation (1) with a HUC8-level measure of the number of farmland acres that receive rents from the CRP; we gather these data directly from USDA.[5] The third and eighth columns of Online Appendix Table A9 provide evidence that the results for CAFO treatment are not affected by the inclusion of CRP acres as a control. Fourth, we use Wisconsin's hydrological network to include in equation (1) an additional control for the number of CAFOs located upstream of each HUC region. It is possible that the number of upstream CAFOs affects downstream surface water quality; if this is the case, and upstream and downstream CAFO exposure are correlated, then our estimation would suffer from omitted variable bias. For this sensitivity check, we perform the analysis at the HUC10 level, because HUC8 regions do not necessarily have up- or downstream

[4] Like the Vaughan (1986) WQI, there is heterogeneity in the TSI amongst waterbodies throughout the state. For example, over 30% of monitoring locations are considered "poor" based on the TSI in the counterfactual scenario of no CAFO expansion.

[5] CRP acreage data from USDA are reported at the county level. We transform these measures to HUC8-year measures using the same technique described in the main text for control variables also collected at the county level.

HUC8s (WDNR 2017). However, WDNR provides data on the downstream HUC10 for each of the 357 HUC10s in the state. Additionally, we wish to continue to perform the analysis at the HUC level, rather than the CAFO level, because our identification relies on the spreading of manure within an area that affects surface waterbody monitoring locations. As described in the main text, we choose the HUC8 level to ensure that the manure from each CAFO is spread within the hydrological region of each CAFO. For this exercise however, we believe that the HUC10 level serves as a reasonable area. Columns 4 and 9 of Online Appendix Table A9 provide the results for this estimation. Estimation results for HUC10 level CAFO treatment are qualitatively robust to the inclusion of the upstream HUC10 CAFO control. However, estimation results for the ammonia sample are more imprecisely estimated than our primary HUC8 level analysis. We hypothesize that the imprecision is the result of two factors: 1) a lack of statistical power, with some HUC10s containing little water quality data and 2) introducing noise into the estimation because CAFOs may spread manure outside of their HUC10. Fifth, we may be interested in relaxing equation (1) to allow differential water quality trends before and after CAFOs enter a HUC8 watershed. To allow for these possible differential trends, we add to equation (1) by interacting an indicator for "CAFO treatment" with the year fixed effects. As shown in the fifth and tenth columns of Online Appendix A9, the inclusion of these trends does not qualitatively change the estimates presented in Table 2 of the main text.

Next, we examine the robustness of our primary estimation results to the change in analysis sample. For these changes to the analysis sample, we: 1) limit the sample to HUC8 regions with a CAFO presence throughout our sample period, 2) limit the sample to the growing season months of April-October, and 3) limit the sample to exclude observations from HUC8 regions with poultry CAFO presence. We again examine in each turn.

First, we limit the sample to HUC8 regions with a CAFO presence throughout our sample period. By removing HUC8 regions that never have any CAFO presence, we rely only on the timing of CAFO expansions to identify the effect of interest. An argument in favor of limiting the sample this way is that areas that have never experienced any CAFO presence may be a poor counterfactual for areas that have significant CAFO presence. Panel A of Online Appendix Table A10 provides results for the estimation of equation (1) on this alternative sample. The coefficient estimates in columns 1 and 2 are larger than their baseline counterparts in Table 2, but the coefficient for the ammonia sample is measured more imprecisely. Second, we examine only those surface water quality readings taken during the months of April to October. We perform this analysis for several reasons. First, surface waterbodies in our sample are often frozen over during the winter months because our sample contains observations in the state of Wisconsin. Second, sunlight and warmer weather can increase nutrient concentrations in surface waterbodies through the release of legacy pollutants from sediment. Thus, November-March surface water readings contain fewer observations and nutrient concentrations are likely lower than those from April-October. Finally, a primary mechanism for the effects of CAFOs on surface water quality is the overspreading of manure on agricultural fields. Rainfall during the growing season then carries excess manure to surface water as non-point source pollution. This runoff pollution generally does not occur during the winter months in Wisconsin, so there exists less variation in water quality measures during these months. As seen in columns 3 and 4 of Online Appendix Table A10, the coefficient estimates are like their counterparts in Table 2 of the main text and remain statistically significant at conventional levels. Finally, we eliminate from the analysis sample those monitoring locations exposed to poultry CAFOs. The motivation for this robustness test is that poultry waste contains a much higher phosphorus content than the waste from other animals. According to a University of

Wisconsin extension publication, poultry waste contains three to five times more phosphate per pound as compared to waste from dairy and beef cows and swine (Madison et al. 1995). We present the results of this estimation in Panel C of Online Appendix Table A10. These results are nearly identical both qualitatively and quantitatively to those of our primary specification. Thus, empirical results are robust to modifications of the analysis sample.

## H. Online appendix tables

**Online Appendix Table A1. Summary statistics of the dependent variables by CAFO presence**

| | Panel A: Total phosphorus | | Panel B: Ammonia | |
|---|---|---|---|---|
| | None | >1 | None | >1 |
| Variable | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Concentration (mg/L) | 0.172 (0.724) | 0.283 (0.720) | 0.212 (0.696) | 0.230 (0.566) |
| Difference of means p-value | 0.000 | | 0.000 | |

**Notes:** Summary statistics for the nutrient measures are at the individual reading level. P-values represent univariate difference tests of mean nutrient concentrations at monitoring locations based on CAFO exposure. Panel A shows average total phosphorus readings for HUC8s with zero CAFOs vs. those with at least one CAFO and panel B shows average ammonia readings for HUC8s with zero CAFOs vs. those with at least one CAFO.

**Online Appendix Table A2. Effects of CAFOs on surface water nutrient levels: Control factors**

| Variable | Total phosphorus (1) | Ammonia (2) |
|---|---|---|
| Maximum temperature (˚C) | 0.0076*** | 0.0010 |
| | (0.0016) | (0.0020) |
| Total precipitation (cm) | 0.1377*** | 0.0131 |
| | (0.0306) | (0.0149) |
| Total precipitation$^2$ (cm) | -0.0057 | -0.0026 |
| | (0.0040) | (0.0019) |
| Median income ($000) | -0.0204** | 0.0034 |
| | (0.0100) | (0.0137) |
| Unemployment rate | 0.0250 | 0.0194 |
| | (0.0179) | (0.0379) |
| Farm acres (000s) | -0.0006 | -0.0015 |
| | (0.0024) | (0.0031) |
| Acres spread with fertilizer (000s) | -0.0059** | -0.0005 |
| | (0.0029) | (0.0033) |
| Developed land (%) | -0.1424** | -0.0782 |
| | (0.0624) | (0.0641) |
| Forested land (%) | -0.0519 | -0.0442 |
| | (0.0391) | (0.0582) |
| Planted land (%) | -0.1757** | -0.1084 |
| | (0.0754) | (0.0807) |
| CAFOs treatment | X | X |
| Observations | 237,528 | 108,577 |
| R-squared | 0.141 | 0.146 |

**Notes**: Each column presents regression results from the estimation of equation (1). Each specification includes year, month, and HUC8 fixed effects. CAFOs treatment is the HUC8 level count of CAFOs on each day. Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A3. CAFO animal units as treatment.**

| | Panel A: Total phosphorus | | | Panel B: Ammonia | | |
|---|---|---|---|---|---|---|
| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
| CAFO animal units (000s) | 0.00193*** | 0.00126*** | 0.00126* | 0.00307*** | 0.00300*** | 0.00300*** |
| | (0.000642) | (0.000452) | (0.0986) + | (0.000651) | (0.000748) | (0.0060) + |
| Time varying controls | | X | X | | X | X |
| Observations | 237,528 | 237,528 | 237,528 | 108,577 | 108,577 | 108,577 |
| Clustering | HUC8 | HUC8 | HUC8 and year | HUC8 | HUC8 | HUC8 and year |

**Notes**: Each column presents regression results from a separate specification of equation (1), where treatment is defined as the HUC8 level count of CAFO animal units. Each specification includes year, month, and HUC8 fixed effects. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level for columns 1, 2, 4, and 5. + Wild cluster bootstrapped p-value in parentheses are for two-way clustering on HUC8 and year in columns 3 and 6. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A4. Effect of CAFOs on water quality monitoring locations**

| | Total phosphorus | Ammonia |
|---|---|---|
| Variable | (1) | (2) |
| CAFOs | 2.754 | -7.682 |
| | (7.672) | (7.048) |
| Observations | 1,109 | 999 |

**Notes:** Each column presents regression results from a separate specification where the outcome is the HUC8-year count of individual surface water quality monitoring locations. Each specification includes year and HUC8 fixed effects. Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A5. Testing for pre-trends**

| | Total phosphorus | | Ammonia | |
|---|---|---|---|---|
| Variable | (1) | (2) | (3) | (4) |
| CAFOs | 0.00488** | 0.00457*** | 0.00618** | 0.00636** |
| | (0.00227) | (0.00152) | (0.00274) | (0.00267) |
| $\tau_{-4}$ | -0.0409 | -0.0313 | 0.0311 | -0.0271 |
| | (0.0625) | (0.0591) | (0.0823) | (0.0851) |
| $\tau_{-3}$ | -0.0340 | -0.0309 | -0.0505 | -0.0970 |
| | (0.0627) | (0.0626) | (0.0611) | (0.0653) |
| $\tau_{-2}$ | 0.0347 | 0.0402 | -0.0309 | -0.0657 |
| | (0.0886) | (0.0876) | (0.0428) | (0.0566) |
| $\tau_{-1}$ | -0.0277 | -0.0204 | 0.0135 | -0.000483 |
| | (0.0529) | (0.0566) | (0.0620) | (0.0637) |
| Time varying controls | | X | | X |
| Observations | 237,528 | 237,528 | 108,577 | 108,577 |

**Notes:** Each column presents regression results from a separate specification of equation (A1). Each specification includes year, month, and HUC8 fixed effects. CAFOs is the treatment variable and measures the number of CAFOs within each HUC8. Tau variables represent indicators for 1 through 4+ years before each HUC8 experienced any CAFO expansion. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A6. Data aggregated to HUC8-month level: unweighted.**

| | Total phosphorus | | Ammonia | |
|---|---|---|---|---|
| Variable | (1) | (2) | (3) | (4) |
| CAFOs | 0.00318* | 0.0265* | 0.00319 | 0.00254 |
| | (0.00173) | (0.00145) | (0.00258) | (0.00196) |
| Time varying controls | | X | | X |
| Observations | 10,199 | 10,199 | 8,673 | 8,673 |

**Notes**: Each column presents regression results from a separate specification of equation (3), where observations are aggregated to the HUC8-month level and are not weighted. Each specification includes year, month, and HUC8 fixed effects. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A7. Long-difference IV results**

| Panel A: Total phosphorus | CAFOs | | CAFO animal units (000s) | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| Variable | (1) | (2) | (3) | (4) |
| Δ CAFOs | 0.00945*** | 0.0239* | | |
| | (0.00324) | (0.0135) | | |
| Δ CAFO animal units (000s) | | | 0.00293** | 0.00343*** |
| | | | (0.00113) | (0.00163) |
| *First Stage* | | | | |
| Freeway kilometers | | 0.00607*** | | 0.0422*** |
| | | (0.00155) | | (0.00480) |
| First stage F-stat on excluded instrument | | 15.40 | | 77.44 |
| Observations | 48 | 48 | 48 | 48 |

| Panel B: Ammonia | CAFOs | | CAFO animal units (000s) | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| Variable | (1) | (2) | (3) | (4) |
| Δ CAFOs | 0.0107*** | 0.0586*** | | |
| | (0.00501) | (0.0178) | | |
| Δ CAFO Animal Units (000s) | | | 0.00462*** | 0.00773*** |
| | | | (0.00150) | (0.00159) |
| *First Stage* | | | | |
| Freeway kilometers | | 0.00687*** | | 0.0521*** |
| | | (0.00144) | | (0.00627) |
| First stage F-stat on excluded instrument | | 22.66 | | 68.95 |
| Observations | 47 | 47 | 47 | 47 |

**Notes:** Each column presents regression results from a separate specification of equation (A6). All specifications include controls for the long-differences in median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Heteroskedasticity robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A8. Sensitivity analysis: Additional stock and flow pollutant controls**

| Variable | Total phos-phorus (1) | Ammonia (2) | Total phos-phorus (3) | Ammonia (4) |
|---|---|---|---|---|
| CAFOs | 0.00221*** | 0.00207* | 0.00477*** | 0.00582** |
| | (0.00082) | (0.00108) | (0.00177) | (0.00239) |
| Time varying controls | X | X | X | X |
| Lagged total phosphorus levels | X | | | |
| Lagged ammonia levels | | X | | |
| CAFO exposure 1-5 years prior | | | X | X |
| Observations | 224,320 | 100,986 | 237,528 | 108,577 |

**Notes:** Each column presents regression results from a separate specification of equation (1) with additional controls included. The lagged total phosphorus and ammonia level measures are the one "period" lagged concentration of each respective pollutant at each monitoring location. For days with multiple readings, the readings are averaged into a single daily measure. These measures therefore represent the previous reading at each monitoring location, regardless of the amount of time between each reading. CAFO exposure 1-5 years prior is a dummy that indicates if the HUC8 of each monitoring location's reading had CAFO exposure. For example, $CAFO_{t-5}$ indicates the HUC8 level CAFO exposure April 15, 1996 for a nutrient reading on April 15, 2001. Each specification includes year, month, and HUC8 fixed effects. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and are clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A9. Sensitivity analysis: Alternative specifications**

| | Total phosphorus | | | | | Ammonia | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| CAFOs | 0.00642*** | 0.00362*** | 0.00384*** | 0.02916** | 0.00366** | 0.0111*** | 0.00356* | 0.00522** | 0.01285 | 0.00418 |
| | (0.00197) | (0.00131) | (0.00120) | (0.01227) | (0.00157) | (0.00310) | (0.00202) | (0.00235) | (0.01062) | (0.00300) |
| HUC 8 FE | X | | X | | X | X | | X | | X |
| Time varying controls | X | X | X | | X | X | X | X | | X |
| Year-by-HUC2 FE | X | | | | | X | | | | |
| HUC10 FE | | X | | X | | | X | | X | |
| CRP acres | | | X | | | | | X | | |
| Upstream CAFOs | | | | X | | | | | X | |
| Differential trends | | | | | X | | | | | X |
| Observations | 237,528 | 237,124 | 237,528 | 237,124 | 237,528 | 108,577 | 107,353 | 108,577 | 107,353 | 108,577 |

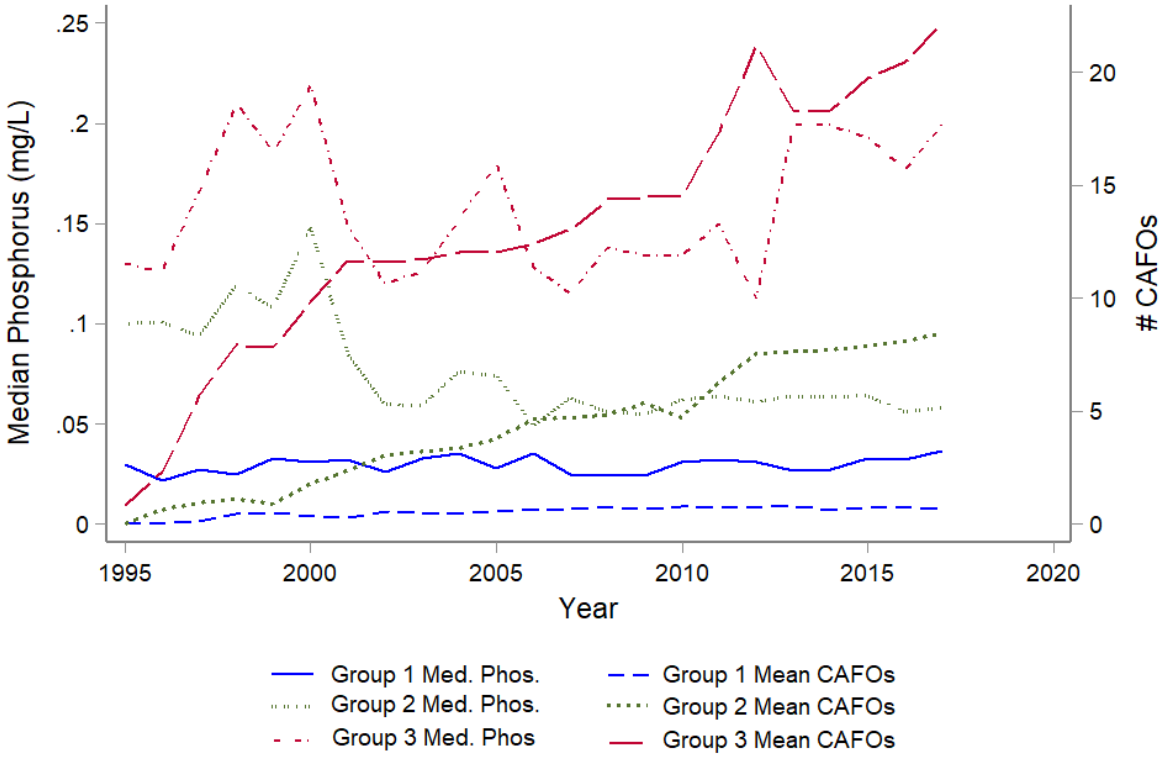
**Notes:** Each column presents regression results from a separate specification of equation (1) with additional controls included. CAFOs represents the count of CAFOs at the HUC8 level (columns 1-3, 5-8, and 10) and at the HUC10 level (columns 4 and 9). Each specification includes year and month fixed effects. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Year-by-HUC2 FE account for differential time trends in each HUC2 region. CRP acres are the HUC8 level number of acres enrolled in the Conservation Reserve Program. Upstream CAFOs represents the count of CAFOs in each HUC10's upstream HUC10, as identified by WDNR. Differential trends represent the interactions of a treatment indicator with the year fixed effects to allow trends to differ after treatment. Robust standard errors are in parentheses and are clustered at the HUC8. *** p<0.01, ** p<0.05, * p<0.1.

**Online Appendix Table A10. Sensitivity analyses: Alternative samples**

| | Panel A: CAFO regions only | | Panel B: April-October only | | Panel C: Exclude HUC8s with poultry CAFOs | |
|---|---|---|---|---|---|---|
| Variable | Total phosphorus (1) | Ammonia (2) | Total phosphorus (3) | Ammonia (4) | Total phosphorus (5) | Ammonia (6) |
| CAFOs | 0.00404*** | 0.00367 | 0.00313* | 0.00615*** | 0.00486*** | 0.00669** |
| | (0.00130) | (0.00289) | (0.00180) | (0.00181) | (0.00147) | (0.00285) |
| Time varying controls | X | X | X | X | X | X |
| Observations | 207,377 | 100,108 | 193,646 | 80,939 | 129,924 | 61,882 |

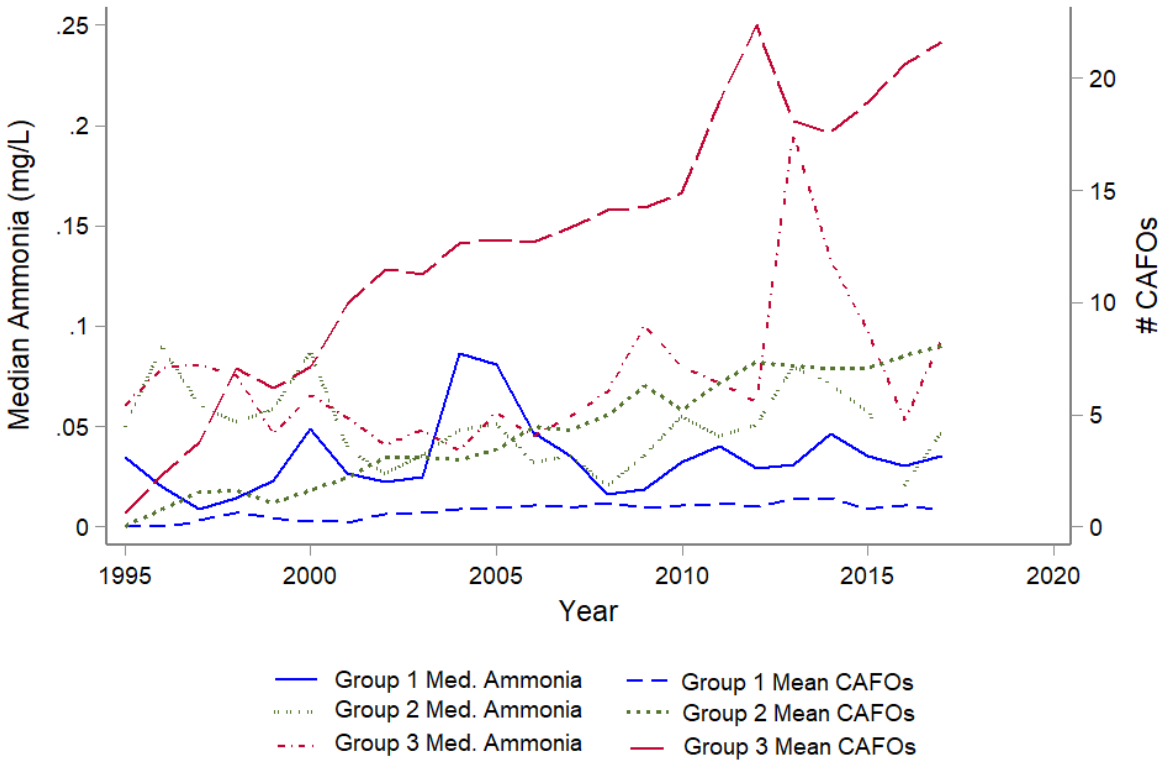
**Notes:** Each column presents regression results from a separate specification of equation (1). Panel A limits the sample to monitoring locations in HUC8 regions that have at least one CAFO at some point in the panel. Panel B limits the sample to water quality results taken in the months of April-October. Panel C excludes observations from any HUC8 region containing poultry CAFOs (chickens, ducks, or turkeys). Each specification includes year, month, and HUC8 fixed effects. Time varying controls include monitor-day level maximum temperature, total precipitation, and total precipitation squared, and HUC8-day level median income, unemployment rate, farm acres, acres of commercial fertilizer application, and NCLD land cover (percent planted, developed, forested). Robust standard errors are in parentheses and clustered at the HUC8 level. *** p<0.01, ** p<0.05, * p<0.1.

## I. Online appendix figures



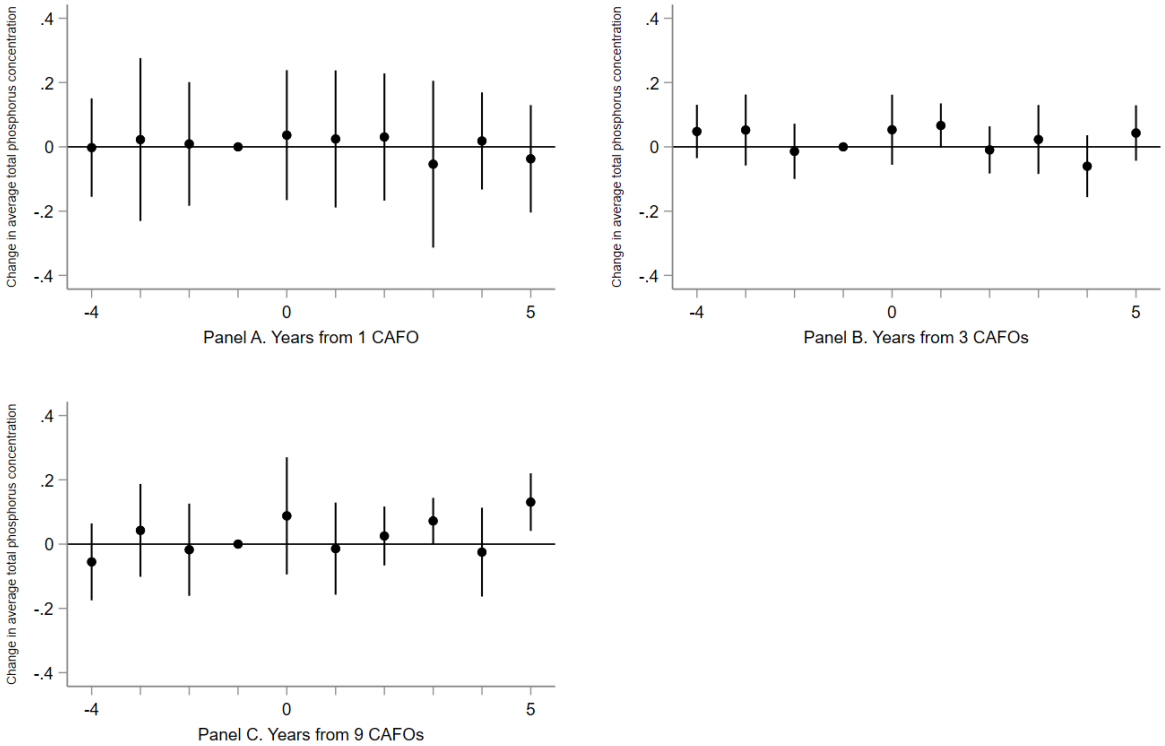

**Online Appendix Figure A1. Total phosphorus vs CAFO exposure by intensity of eventual treatment, 1995-2017**

Notes: This graph shows median total phosphorus and mean CAFOs over time for three groups. The groups are divided by intensity of eventual CAFO exposure. Group 1 is the no/low treatment group and includes monitors in HUC8 watersheds that are exposed to at most three CAFOs during the sample (lowest quartile). Group 2 is the medium treatment group and includes monitors in HUC8 watersheds that are exposed to between four and 16 CAFOs during the sample period (second and third quartiles). Group 3 is the high treatment group and includes monitors in HUC8 watersheds that are exposed to more than 16 CAFOs during the sample period (highest quartile).

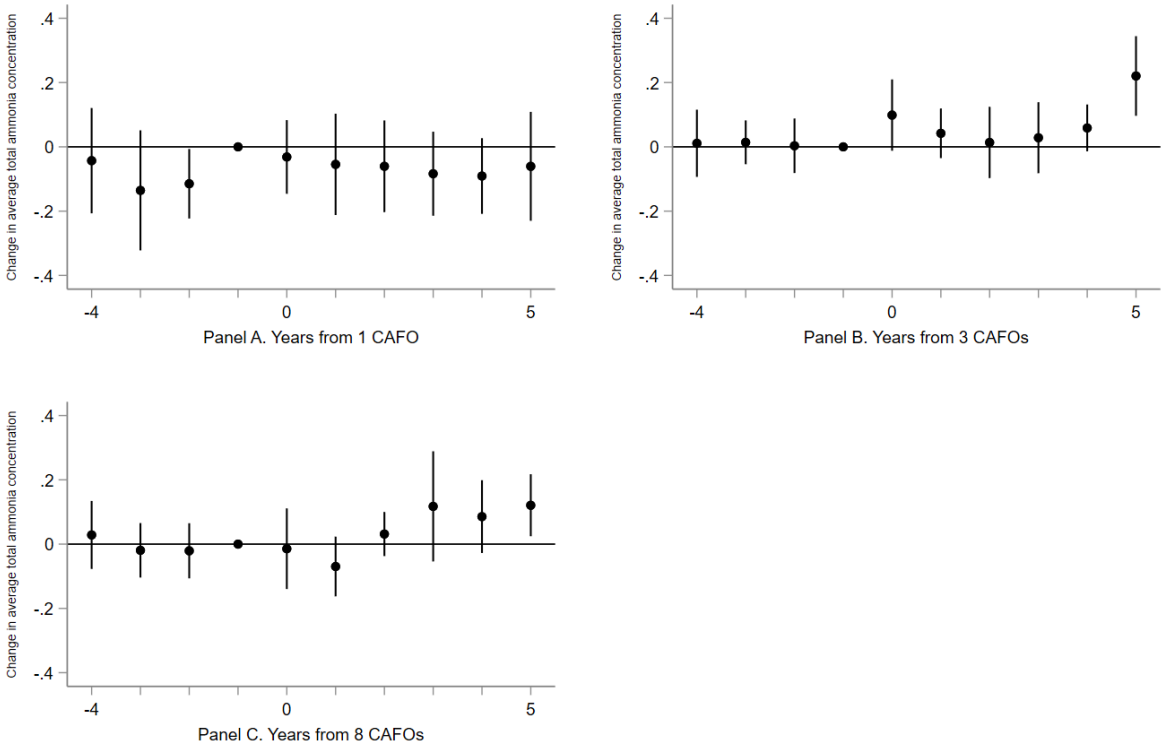



**Online Appendix Figure A2. Ammonia vs CAFO exposure by intensity of eventual treatment, 1995-2017**

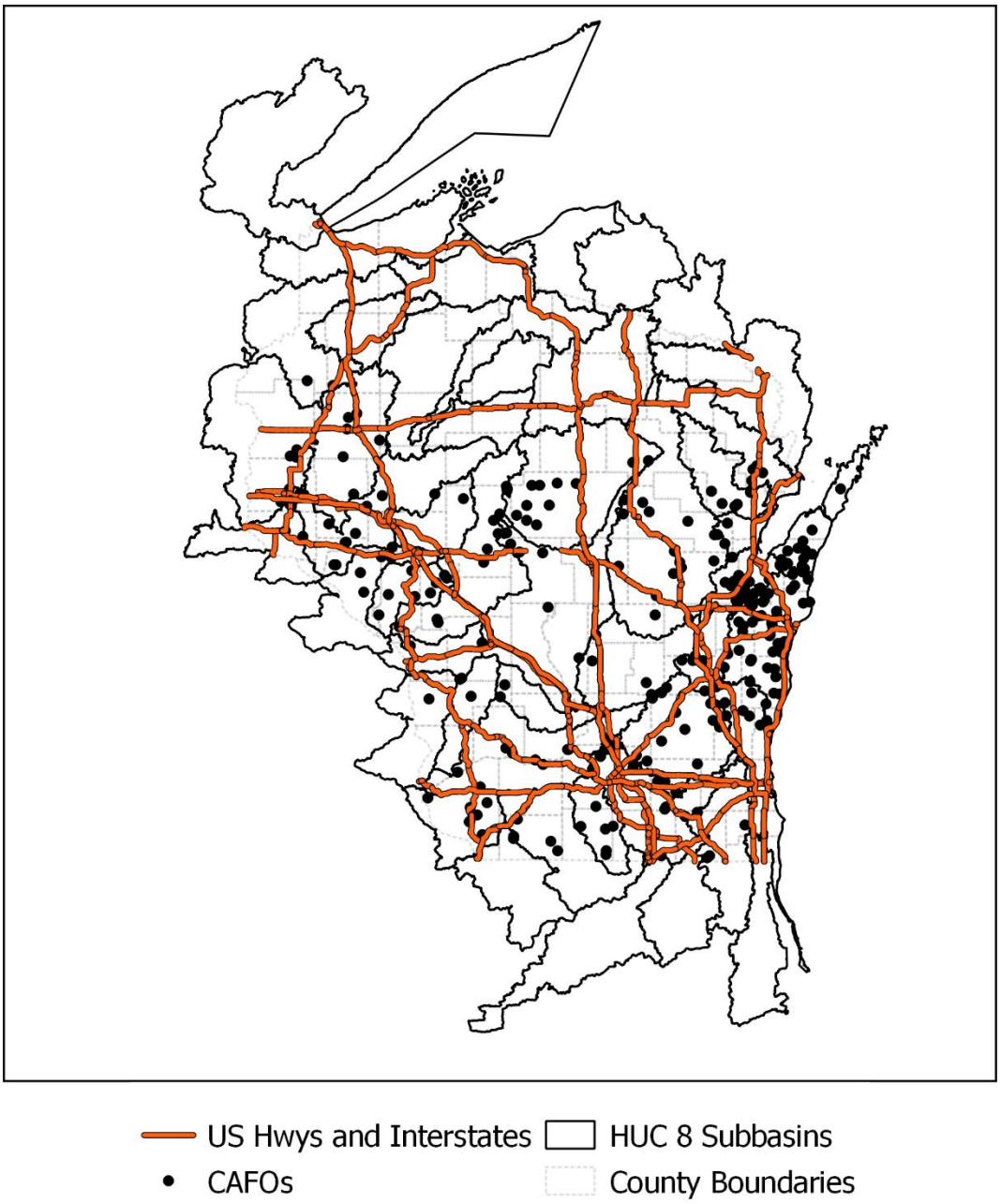
Notes: This graph shows median ammonia and mean CAFOs over time for three groups. The groups are divided by intensity of eventual CAFO exposure. Group 1 is the no/low treatment group and includes monitors in HUC8 watersheds that are exposed to at most three CAFOs during the sample (lowest quartile). Group 2 is the medium treatment group and includes monitors in HUC8 watersheds that are exposed to between four and 16 CAFOs during the sample period (second and third quartiles). Group 3 is the high treatment group and includes monitors in HUC8 watersheds that are exposed to more than 16 CAFOs during the sample period (highest quartile).




**Online Appendix Figure A3. Event studies for total phosphorus.**
Notes: This figure plots coefficients with 95 percent confidence intervals from separate estimations of equation (A2) where total phosphorus is the dependent variable. Panel A defines the event as the year in which the first permitted CAFO began operating in the HUC8 region, Panel B defines the event as the year in which the third permitted CAFO (median level of treatment) began operating in the HUC8 region, and Panel C defines the event as the year in which the ninth permitted CAFO (75$^{th}$ percentile of treatment) began operating in the HUC8 region.

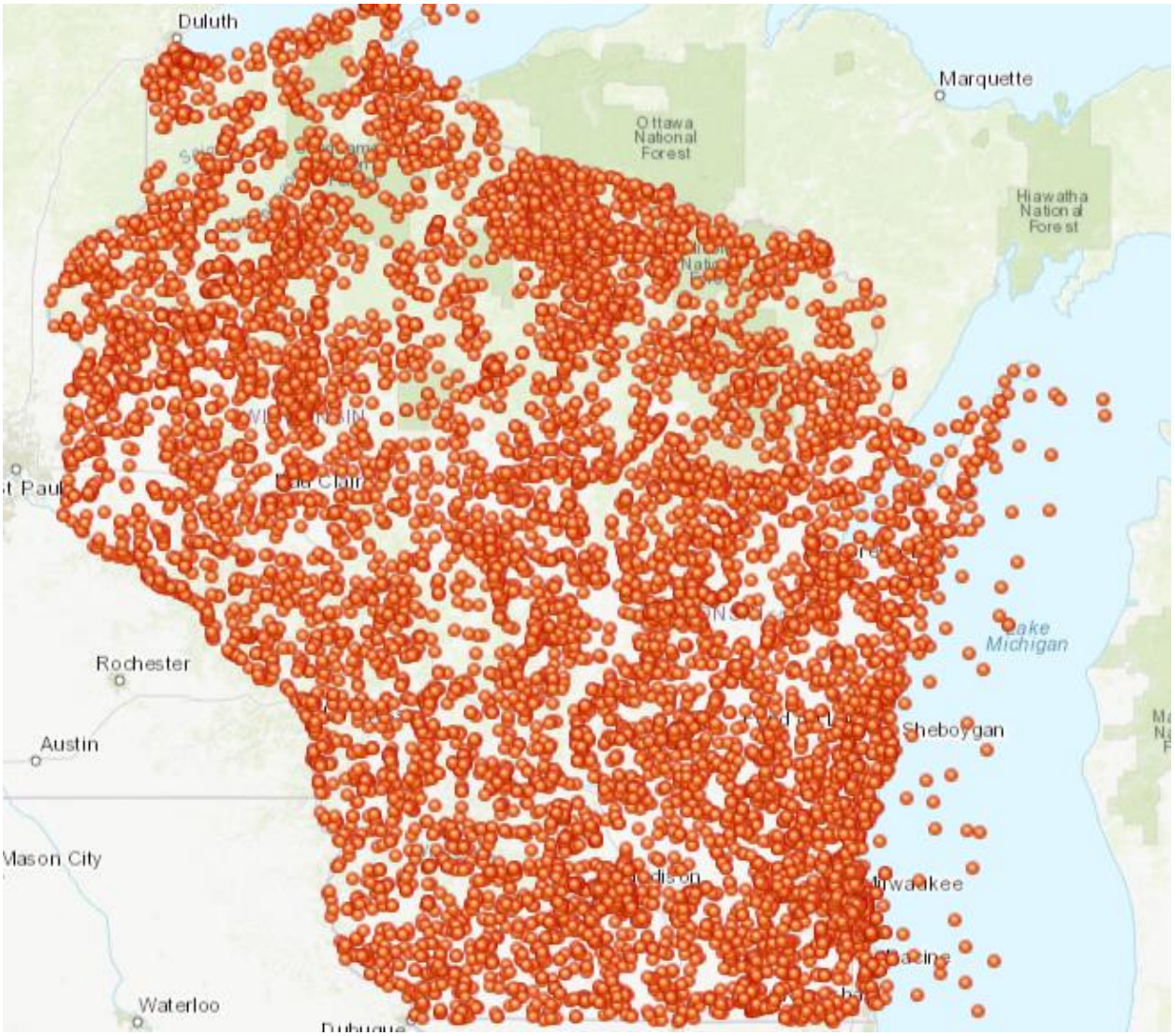



**Online Appendix Figure A4. Event studies for ammonia.**
Notes: This figure plots coefficients with 95 percent confidence intervals from separate estimations of equation (A2) where ammonia is the dependent variable. Panel A defines the event as the year in which the first permitted CAFO began operating in the HUC8 region, Panel B defines the event as the year in which the third permitted CAFO (median level of treatment) began operating in the HUC8 region, and Panel C defines the event as the year in which the eighth permitted CAFO (75$^{th}$ percentile of treatment) began operating in the HUC8 region.




**Online Appendix Figure A5. Location of major freeways and CAFOs in Wisconsin, 2017**

Notes: Red lines on this map show U.S. highways and interstate highways, i.e., freeways. Black dots represent CAFOs in 2017 and borders are shown for HUC8 regions (black, solid lines) and counties (faded, dashed lines).




**Online Appendix Figure A6. Location of water quality monitoring locations**

Notes: Dots represent the locations of water quality monitoring locations for total phosphorus and ammonia in Wisconsin. The water quality monitoring locations in the Great Lakes are not used in the final analysis samples.

## References


Carlson, R.E., 1977, "A trophic state index for lakes", *Limnology and oceanography*, 22(2), 361-369.

Chay, K.Y. and M. Greenstone, 2005, "Does air quality matter? Evidence from the housing market", *Journal of Political Economy*, 115(2), 376-423.

Cornelissen, T., D. Dustmann, A. Raute, and U. Schönberg, 2016, "From LATE to MTE: Alternative methods for the evaluation of policy interventions", *Labour Economics*, 41, 47-60.

Heckman, J.J and E. Vytlacil, 2005, "Structural equations, treatment effects, and econometric policy evaluation", *Econometrica*, 73(3), 669-738.

Imbens, G.W. and J.M. Wooldridge, 2009, "Recent developments in the econometrics of program evaluation", *Journal of Economic Literature*, 47(1), 5-86.

WDNR, 2017, Hydrologic units – 12 digit (subwatersheds), https://data-wi-dnr.opendata.arcgis.com/datasets/5a5f60f2f51945a5a98c1b470c5a42af_0.