2018

Three essays on the economics of U.S. water policy

Xianjun Qiu
Iowa State University

Three essays on the economics of U.S. water policy

by

Xianjun Qiu

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
David A. Keiser, Major Professor
Gabriel Lade
Ivan Rudik
Catherine Kling
Wendong Zhang

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

DEDICATION

To most Venerable Master KuanRu, who taught me how to be a better person and have a Bodhi Heart.

To my parents, Yanhui Qiu and Huiyan Meng, who have always loved and supported me unconditionally. Their love gives me the courage and confidence to overcome all the challenges in life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I thank all my professors for their guidance, wisdom and patience. I am deeply grateful to my major professor, Dr. David A. Keiser, for his guidance, trust and support throughout this research and the writing of this thesis. His insights, encouragement and utmost patience helped me achieve what I have today. I am also indebted to Dr. Gabriel Lade and Dr. Ivan Rudik for their guidance and constructive comments during the completion of this thesis. I would also like to thank Dr. Catherine Kling and Dr. Wendong Zhang for their efforts and contributions to this work. I thank Dr. Keri Jacobs for her valuable comments on the CRP project. I am also grateful to all the comments and inspiring ideas from all the participants at the graduate student seminars. I also thank my dear friend Jiajia Gu for all the encouragement and joy. I would additionally like to thank my friends, the department faculty and staff for making my time at Iowa State University a wonderful experience.

ABSTRACT

This dissertation focuses on the relationship between human activities and water resources in the United States. The first chapter studies how the Conservation Reserve Program (CRP), a national land conservation program, affects nutrient concentrations in groundwater. The second chapter utilizes a comprehensive and unique historical dataset on drinking water facilities and studies how water supply sources affect communities' resilience during droughts. The third chapter studies the effectiveness of a series of incomplete phosphorus lawn fertilizer bans in Florida.

1. GENERAL INTRODUCTION

The complicated relationship between human activities and water quality/quantity has interested many researchers. My primary research interest lies in environmental economics and water resource policy. This dissertation studies water quality and water quantity in the United States, especially groundwater. In particular, the first chapter, The Conservation Reserve Program and Nutrient Pollution in Groundwater, studies how the Conservation Reserve Program (CRP) affects groundwater quality in the United States. CRP is a land conservation program administrated by the Farm Service Agency. It was first implemented in 1986 and one of its many goals to address environmental concerns is to improve water quality. I am particularly interested in the CRPs impact on nutrient concentration in groundwater because: first, groundwater quality is essential. Groundwater provides 25 to 40 of the worlds drinking water, provides over 30% of the United States drinking water and around 15 million rural households in the United States rely on private wells for drinking water. Second, nutrients in groundwater, especially nitrate, has been proven to be harmful to human health. Excess nitrate in drinking water is fatal to infants. The source of nutrients in groundwater mostly comes from agricultural activities. Third, previous studies on CRPs impact on water quality focus on surface water, and I find very few studies specifically focus on groundwater quality. I use fixed effects model and difference-in-differences method in this research. Results from estimating the fixed effects model indicate that for a 1% increase in CRP acres / (CRP + cropland acres) ratio in a county, nitrate concentration in groundwater drop by 7.9%, and phosphorus concentration in groundwater drop by 27%. With the difference-in-differences method, I am able to test the robustness of my previous results, and I show that there is a significant difference in nutrient concentration level in groundwater between counties with CRP enrollment and counties without. My finding is markedly different from existing literature because its focuses on groundwater quality, its application of fixed effects model and its robustness in results. The second chapter of my dissertation, Reliable Drinking Water Supply and Cities Resilience to Drought (joint with Dr. David Keiser, Dr. Gabriel Lade, and Dr. Ivan Rudik), studies how differences in drinking water supply sources affect how cities adapt to and respond to extreme weather events. We compiled a unique historical panel dataset of drinking water treatment facilities across the U.S. for this research. Our findings suggest that a groundwater source for drinking water prevents migration during severe droughts. This finding shows that it is crucial for city planners to consider a variety of water supply sources and invest in groundwater supply source, and this conclusion is especially meaningful to developing countries that are still building or expanding modern drinking water infrastructure. The third chapter of my dissertation, The Effectiveness of Phosphorus Lawn Fertilizer Bans (joint with Dr. David Keiser), studies the effectiveness of incomplete phosphorus lawn fertilizer bans in Florida by utilizing a restricted consumer scanner data and fixed effects model. We find this series of bans on phosphorus fertilizer application results in a 21.7% drop in fertilizer purchase in ban counties. We also find there is a spillover effect on fertilizer sales in ban counties during before-ban seasons, and we show there is no consistent evidence that the bans result in spatial spillover to non-ban, border counties. Our study contributes to the literature by studying a ban that restricts fertilizer use but not fertilizer sale, and we explore the effectiveness of this ban by looking into changes in consumer behavior. We also take advantage of the spatial and temporal variation of a series of bans, compared to previous studies on similar household phosphorus restrictions which mostly focus on a single ban.

2. THE CONSERVATION RESERVE PROGRAM AND NUTRIENT POLLUTION IN GROUNDWATER

This paper examines how the Conservation Reserve Program (CRP) affects nutrient concentrations in groundwater. Using a fixed effects model on data from the National Resources Inventory (NRI) and the United States Geological Survey (USGS), I find evidence that CRP enrollment reduces nitrate and phosphorus concentrations in groundwater. A 1% increase in the five-year moving average of CRP/(CRP+cropland) ratio leads to a 7.9% reduction in nitrate concentration and a 27% reduction in phosphorus concentration in groundwater. However, the finding is sensitive to model specification. I also find evidence of substantial heterogeneity in the program's impact both over time and across space.

2.1 Introduction

Nutrient pollution from agriculture is identified as a leading cause of water quality impairment worldwide. Nutrient runoff increases surface water pollution that contributes to algal blooms and death of aquatic life. Nutrient leaching to groundwater, though mostly invisible, is no less dangerous. Over 38% of the United States population depends on groundwater for drinking water, over 98% of self-supplied domestic water comes from groundwater [NGWA, 2016, USGS, 2010]. Of the two common nutrient contaminants in groundwater, nitrogen and phosphorus, high nitrogen concentration in drinking water is particularly harmful, as it can be fatal to infants [Fan and Steinberg, 1996, Spalding and Exner, 1993]. Phosphorus is generally considered safe for human consumption, but phosphorus in groundwater can leach to surface water and lead to surface water pollution. Crutch-field et al.

(1997) estimates that on average a household would be willing to pay $45 to $60 per month to reduce nitrate in drinking water to the EPA minimum safety standard.

There are many land use programs in the United States that aim at providing environmental benefits. CRP alone costs around $1.7 billion each year and is one of the most important conservation programs administrated by the U.S. Department of Agriculture. A primary goal of the CRP is to improve water quality. Researchers have focused on studying CRP enrollments' impact on surface water quality, and have reached conflicting results. A report from Farm Service Agency (FSA) shows CRP significantly reduces nutrient runoff; nitrogen and phosphorus leaving CRP land are 95% and 86% less respectively compared to cropped land [Farm Service Agency, 2012]. However, Sprague and Gronberg [2012] demonstrated there exists a positive relation between CRP area and nutrient export to surface water (both nitrogen and phosphorus), especially when soil erodibility is low or moderate.

Previous studies have found correlations between general agricultural land use and nitrate concentrations in groundwater in many countries [Böhlke, 2002, Strebel et al., 1989, Kumazawa, 2002, Viers et al., 2012, Gardner and Vogel, 2005], but few studies focus on how the CRP affects nutrient concentration in groundwater. The relation between CRP acreage and groundwater quality can be more complicated than surface water, and this area needs exploration. Compared to surface water, groundwater quality can be affected by many factors that are difficult to measure. Nutrient concentrations in groundwater can be affected by precipitation, soil quality, aquifer type, underground flow direction, depth of water table, among other factors [Nolan, 2001, Viers et al., 2012, Dinnes et al., 2002]. Nitrates are highly leachable and can easily reach groundwater compared to other forms of nitrogen such as ammonia. Wang et al. (2015) conduct a simulated rainfall experiment to study nitrate leaching. They find that about 50% of nitrate-nitrogen from total fertilizer applied to topsoil will stay in surface and bottom layers of the soil, and becomes a pollution source for groundwater. Nitrate is also very stable and persistent in groundwater under natural conditions [Bruggeman et al., 1995, Burow et al., 2007, Mastrocicco et al., 2011, Lerner and Harris, 2009]. The persistence of nitrate in groundwater and slow groundwater recharge/renew rate also makes it difficult to evaluate the relationship between land use and groundwater pollution.

Previous research on CRP and other land use programs has conducted cost-benefit analyses of different conservation programs and environmental policies. The relation between environmental benefits and land retirement costs is not straightforward and can depend on the environmental targets, productivity of the land, and policy design [Claassen et al., 2008]. Babcock et al. [1996] show that the targeting instrument of the CRP matters when optimally allocating a given budget, and different environmental benefits can correlate with CRP enrollment in different ways, even under the same policy. Most previous work studies surface water quality as a targeting instrument. This paper will provide a better view for future analyses that focus on groundwater benefits.

In this paper, I use data on CRP enrollment from National Resources Inventory (NRI) and water pollution data from United States Geological Survey (USGS). I use a fixed effects model to identify the relationship between CRP enrollment and nutrient pollution. The specification has advantages in its ability to control for unobserved factors that may contribute to nutrient concentration in groundwater. I find that the ratio of CRP land over the sum of CRP land and cropland (CRP/(CRP+cropland))[1] is negatively correlated with nitrate and phosphorus concentration in groundwater. In my preferred specification, a 1% increase of CRP ratio leads to a 7.9% reduction of nitrate concentration and 27% reduction of phosphorus concentration. However, the result is sensitive to the empirical specification. I also find that after the initiation of the CRP, counties with CRP land have around 55% less nitrate (dissolved) and 34% less nitrite (dissolved) in groundwater compared to counties that did not have CRP land.

This paper shows enrolling cropland into the CRP may reduce nitrate, nitrite and phosphate-phosphorus concentration in groundwater. I also show there is a lag between

[1]This ratio: (CRP/(CRP+cropland)), will be referred to as "CRP ratio" in the rest of this paper.

CRP enrollment and changes in nutrient concentrations in groundwater.That is, after CRP enrollment, changes in groundwater quality usually take years to become observable and also can last a few years after the land is dropped from CRP. This lag effect suggests when groundwater quality is a targeting instrument in program evaluation, the long-term effect should be considered.

The paper proceeds as follows: section 2 introduces the design and program history of the CRP, section 3 describes the datasets used in this paper, section 4 describes the empirical model, section 5 discusses the regression results, section 6 presents further analysis using the difference-in-differences method and the final section concludes.

2.2 The Conservation Reserve Program

The Conservation Reserve Program (CRP) was created in 1985 over concerns of high levels of soil erosion in the United States. It is a cost-share, rental payment program administered by the U.S. Department of Agriculture (USDA) Farm Service Agency (FSA). The primary goal of the CRP is to conserve and improve the natural environment, including soil, water quality, and wildlife habitat. The CRP is voluntary and requires farmers and landowners to sign 10 to 15-year contracts. To enroll land in CRP, farmers need to convert highly erodible or previous cropped land into conservation buffers[2] and long-term covers, such as grasslands, wildlife shelter planting[3], and riparian buffers[4]. Cropland needs to be planted to an agricultural commodity at least four of the previous six crop years. It also needs to be physically and legally capable of being planted in a normal manner to an agricultural commodity. Eligible land is either bid into the program and ranked using the

[2]Conservation buffers are small areas or strips of land in permanent vegetation, designed to intercept pollutant and manage other environmental concerns, including air, water, soil pollution. Conservation buffers include riparian buffers, filter strips, grassed waterways, shelterbelts, windbreaks, living snow fences, contour grass strips, cross-wind trap strips, shallow water areas for wildlife, field borders, alley cropping, herbaceous wind barriers, and vegetative barriers (NRCS).

[3]For example, Wildlife Food Plots that plant grains, such as rye, millet and buckwheat to provide food source and cover for many wildlife species, especially in extreme weather.

[4]Riparian buffer is a kind of conservation buffers, it traps sediment, nutrients and pollutants, recharges groundwater and provides better habitat for fish and other wildlife.

Environmental Benefits Index[5] (EBI) in general sign-ups, or is enrolled automatically (no bidding or ranking) via Continuous Conservation Reserve Program[6] (CCRP) if it meets certain eligibility criteria (NSAC, 2016). Thus, substantial selection bias exists when comparing outcomes of CRP and non-CRP land.

USDA was authorized to enroll up to 45 million acres of CRP land in 1985, and then the enrollment cap dropped to 24 million acres from 1985 to 2017. The enrollment cap has to do with commodity price. When crop prices are high, farmers have less incentive to retire their land. Figure 2.1 shows that less enrollment and tighter enrollment cap followed an increase in crop price in the early 90s. After crop price decreased in the late 90s, CRP enrollment increased, and the enrollment cap was raised. CRP is a costly program. Annual funding for CRP was about $1.7 billion in most years after 1991, which takes up a majority of major USDA conservation program spendings (USDA, 2016). Total nominal rental payments accumulate to almost $48 billion from 1987 to 2017. Average rental payment was $42.99/acre in 1986 and gradually increased to $72.61/acre by 2016. Figure 2.2 shows how CRP rental payment changed from 1985 to 2016.

Former research has reached conflicting conclusions on whether the CRP improves surface water quality, and I expect no less complication when it comes to groundwater. CRP may influence nutrient concentrations in groundwater through several mechanisms. By converting highly erodible land into buffers and pulling land out of agricultural production, the CRP reduces fertilizer use and nutrient runoff. Constructed and restored wetland can enhance the denitrification process, turning harmful nitrate to benign nitrogen gas (N_2), thus reducing nitrate concentrations in the field. On the other hand, grass filter strips and riparian buffers are installed to intercept nutrients before they enter surface water, increasing

[5]Current EBI factors include wildlife habitat benefits, water quality benefits, on-farm soil-retention benefits, benefits that will likely endure beyond the contract period, air quality benefits and cost

[6]For CCRP the land must be recognized as "marginal pastureland" that is bordered to a stream, creek, river, sink-hole, and/or duck nest. CCRP eligible practices include riparian buffers, wildlife habitat buffers, wetland buffers, filter strips, wetland restoration, grass waterways, shelterbelts, windbreaks, living snow fences, contour grass strips, salt-tolerant vegetation, and shallow water areas for wildlife.

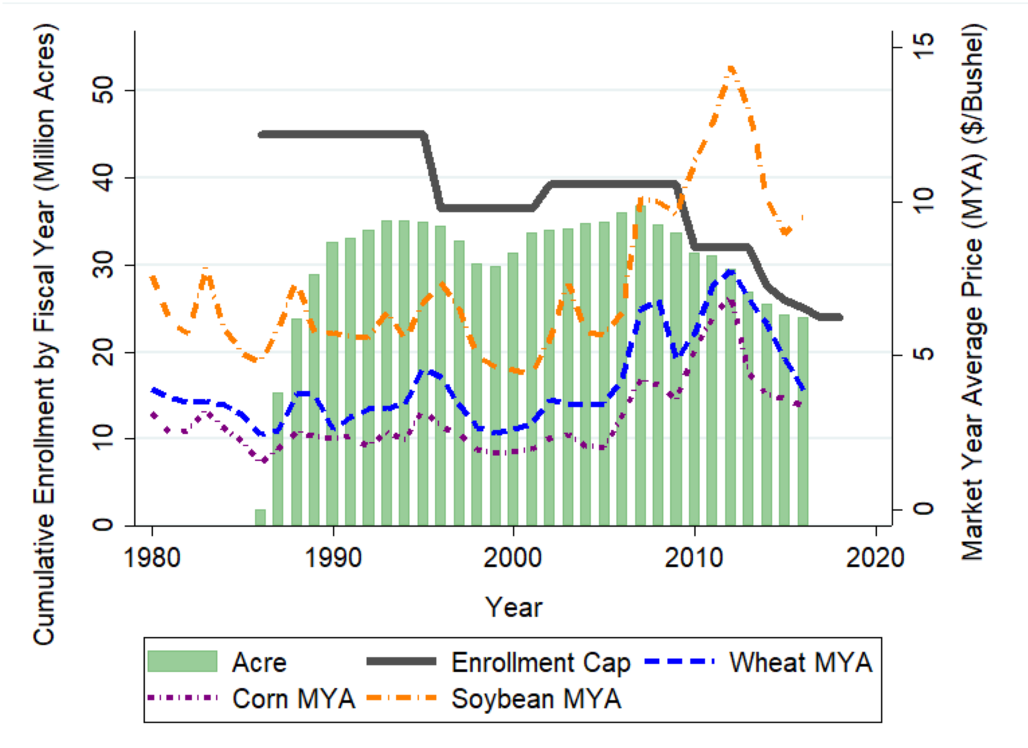


Figure 2.1: CRP enrollment cap, cumulative enrollment and crop prices
Notes: This figure shows cumulative CRP enrollment by fiscal year and CRP enrollment cap in million acres. It also plots the market year average price (MYA) of major field crops: wheat, corn, and soybean. MYA price is a weighted average of the monthly prices for the marketing year. It is used to calculate the current year actual crop revenue and determines farm bill programs payments.

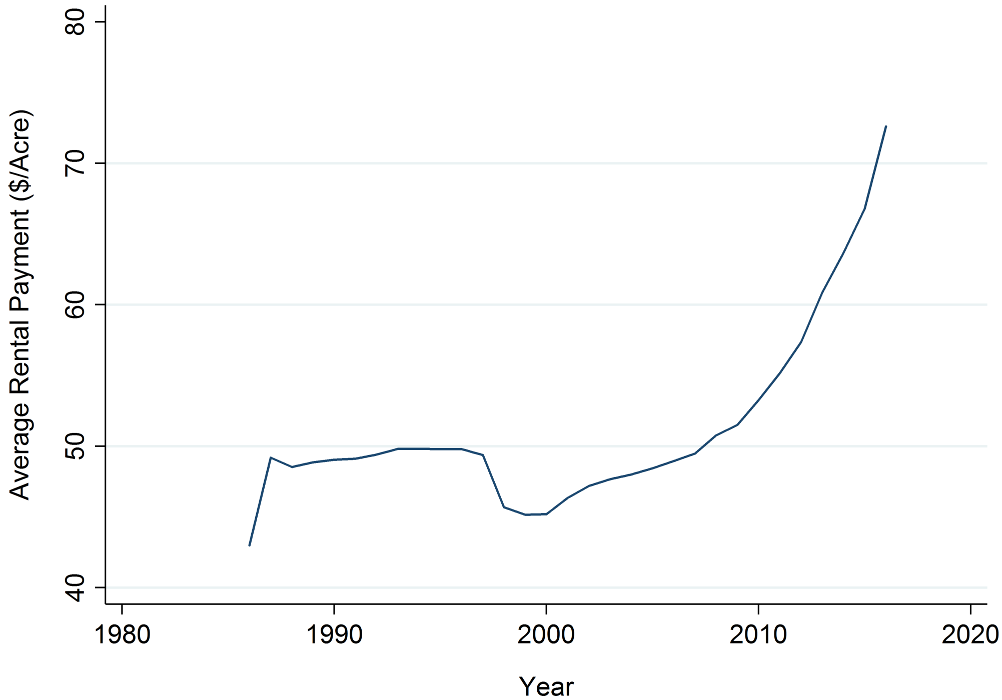
the amount of nutrients left in the field and potentially increases the amount of nutrients that leach into groundwater.

2.3 Data

Two major datasets are used in this paper: land use data from the National Resources Inventory (NRI) and water quality data from the U.S. Geological Survey (USGS).

2.3.1 Land use data

Land use data from the NRI provides county-level acreage of 12 categories of land use, including cultivated cropland, noncultivated cropland, rangeland, pasture land, urban, rural transportation, small water area, large water area, forest, federal land, minor land and CRP



Notes: This figure shows the average rental payment of CRP land started at $42.99/acre in 1986. It fell below $46/acre around 1998 and then steadily increased to $72.61/acre in 2016.

Figure 2.2: Average rental rate of CRP land

land. The dataset covers 1982, 1987, 1992, 1997, and every year from 2000 to 2010. I use linear interpolation to fill in the missing years from 1982 to 2000 for all categories except minor land and large water area. This is because large water area does not vary by year, and acreage of minor land in gap years is generated by subtracting all other categories from total acreage so that total acreage remains the same across years. I also run regressions using land use data without linear interpolation. I get similar results with or without linear interpolation, and the latter results are presented in Appendix[7].

Figure 2.3 provides a snapshot of CRP land by county in 1990, 2000 and 2010. This figure illustrates that the Midwest region, Montana, Washington, the adjacent area of Colorado, Kansas, Oklahoma and Texas are heavily enrolled in CRP, while counties along the

[7]See Table A.5 for regression results using original data with no interpolation.

Mississippi River have relatively less CRP enrollment over time. It also shows the density distribution of CRP enrollment does not vary much over time. Figure 2.4 presents a line plot of yearly CRP ratio[8] for eight counties from 1982 to 2010. This plot provides an example of the variation in CRP enrollment over both time and space. It also illustrates a common upward trend in enrollment over the first ten years of the program.

2.3.2 Water quality data

Groundwater quality data comes from the USGS. This dataset covers 1982 to 2010, and provides 19 categories of nutrient concentrations (Table 2.1). Different nutrient categories are collected at different months in different monitoring stations. There are nearly 900,000 observations from 75,252 distinct monitoring stations in the U.S. (Figure 2.5). For the ease of calculation and comparison, I converted all concentration values to nitrogen and phosphorus content. For example, "Nitrate, dissolved" means the concentration of NO_3^--N dissolved in the sample, not the concentration of NO_3^- [9].

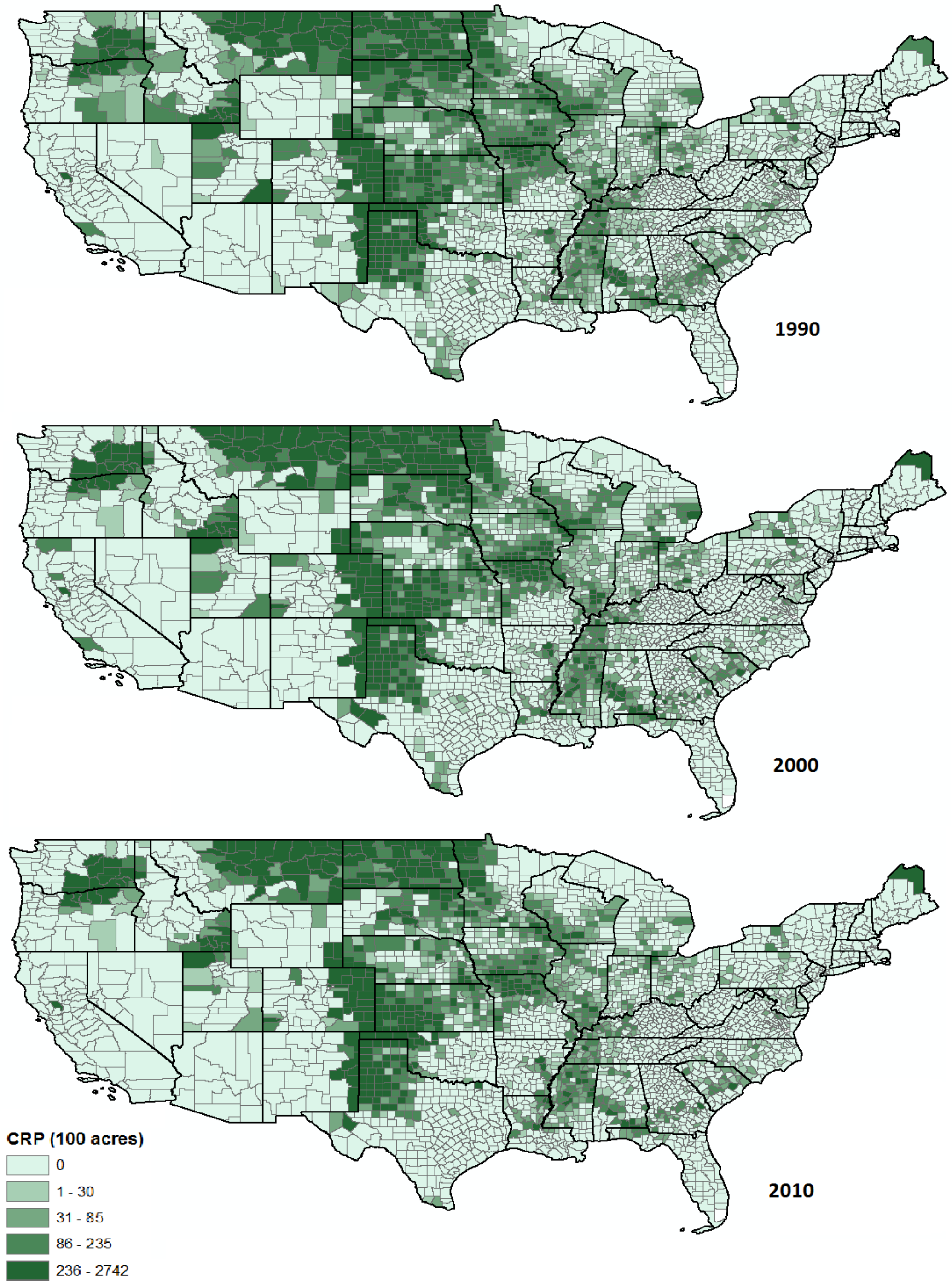
Of the 19 categories of nutrient categories in water quality dataset, I focus on nitrate concentration because it is a major health concern and is very persistent in groundwater. Phosphorus is a critical concern in surface water pollution, but its existence in groundwater only becomes a concern when it gets into surface water through the water cycle. In the water quality dataset, about 40% of records[10] are marked with "Not detected" or "Present Below Quantification Limit". I replace all these records with one-half of the corresponding "Historically lower limit reported" value. These non-reported records represent a large fraction of the data, so I explore the sensitivity of my results using censored data[11]. Less than 0.001% is marked as "Present Above Quantification Limit", and is replaced with "Historically upper limit reported" value. Nutrient concentration appears to be extremely

[8] CRP ratio $= \dfrac{acres\ of\ CRP\ land}{acres\ of\ CRP\ land\ and\ cropland}$

[9] Nitrogen-ion conversion: $NH_3 = NH_3$–N $\times$ 1.21589, $NH_4^+ = NH_4^+$–N $\times$ 1.28786, $NO_2^- = NO_2^-$–N $\times$ 3.28443, $NO_3^- = NO_3^-$–N $\times$ 4.42664. Phosphorus-ion conversion: $PO_4^{3-} = PO_4^{3-}$–P $\times$ 3.06619.

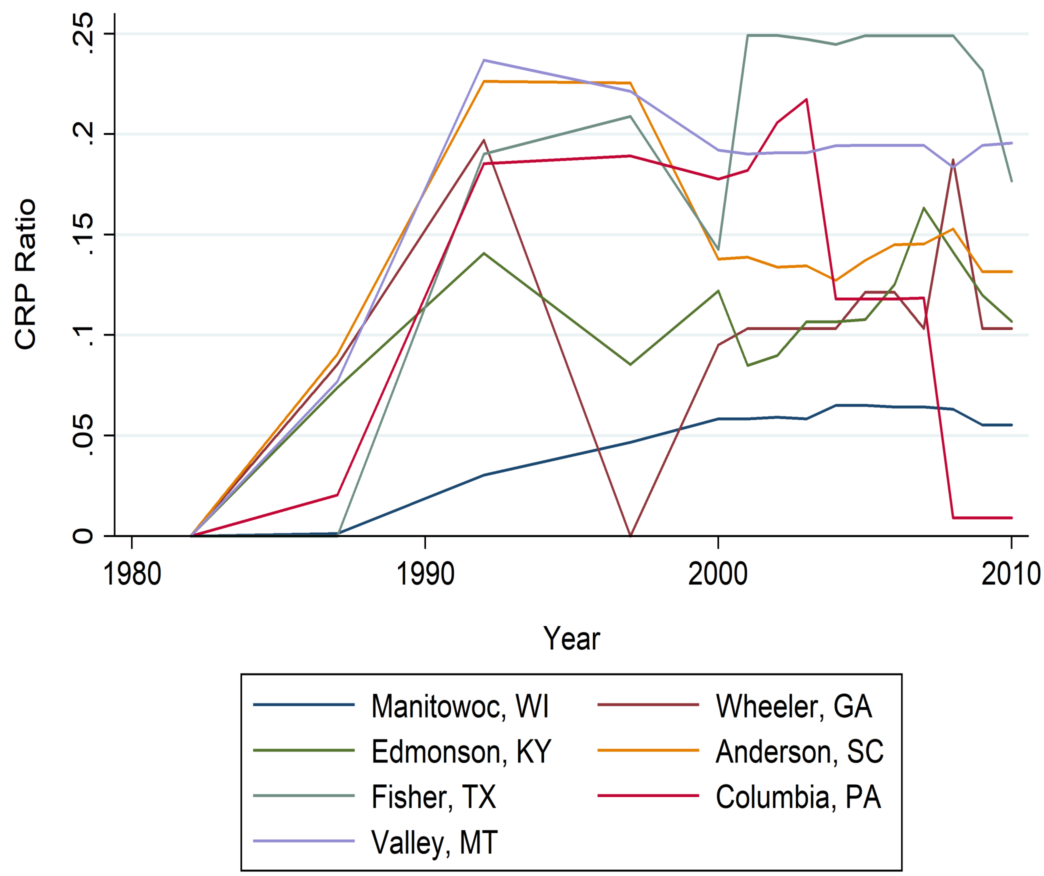
[10] 24.67% of nitrate observations

[11] See basic statistics and regression results using original censored data in Figure A.1 and Table A.6.



Notes: This figure shows the density and location of CRP land. It shows that though there exists variation in the density of CRP land, most CRP enrollment consistently comes from the same areas.

Figure 2.3: Snapshot of CRP land: 1990, 2000, and 2010



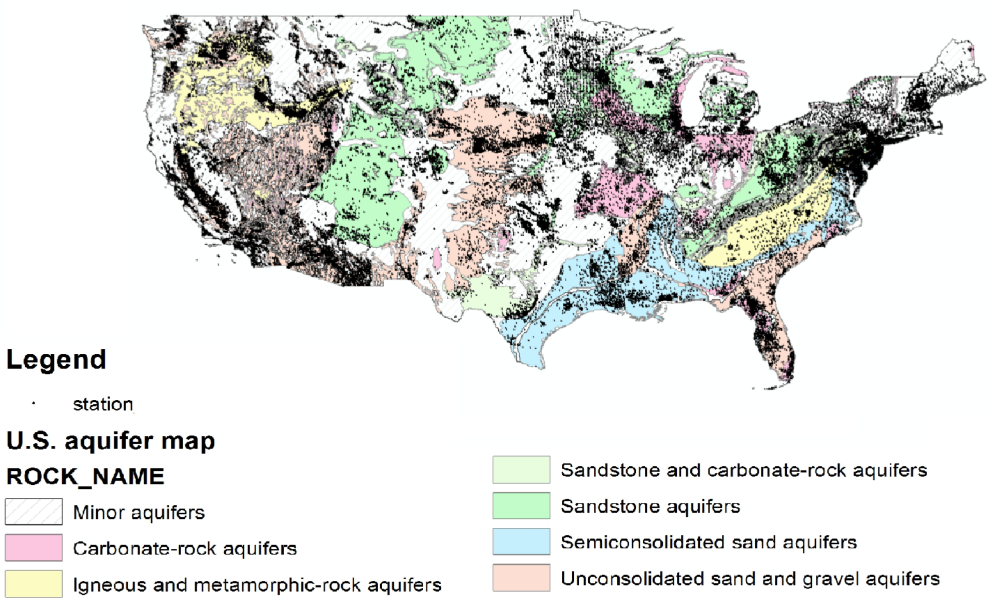
Notes: This figure shows the CRP ratio varies spatially. These eight counties are selected using random number generator from the dataset.

Figure 2.4: Variation in CRP ratio in eight counties

Table 2.1: Categories of nutrients in groundwater quality data

Nutrient	Frequency	Form	
		Dissolved	Total
Ammonia and ammonium	134,796	✓	✓
Inorganic nitrogen (nitrate and nitrite)	164,590	✓	✓
Kjeldahl nitrogen	62,083	✓	✓
Nitrate	128,277	✓	✓
Nitrite	118,669	✓	✓
Nitrogen	2,217	-	✓
Nitrogen, mixed forms	73,166	✓	✓
Organic nitrogen	70,858	✓	✓
Phosphate	115,390	✓	✓
Phosphate-phosphorus	9,184	-	✓
Phosphorus	72,365	-	✓

Notes: This table shows all measurements used when water quality data was collected. Some categories overlap but I keep them exactly as how they are categorized in the original water quality dataset. For example, "Nitrogen, mixed forms" measures the sum of NH_3, NH_4, organic nitrogen NO_2 and NO_3, which overlaps with several other categories. "nitrogen" is a vague definition compared with other categories.



Notes: Black dots in this map are monitoring stations.

Figure 2.5: Map of monitoring stations and aquifers by rock type

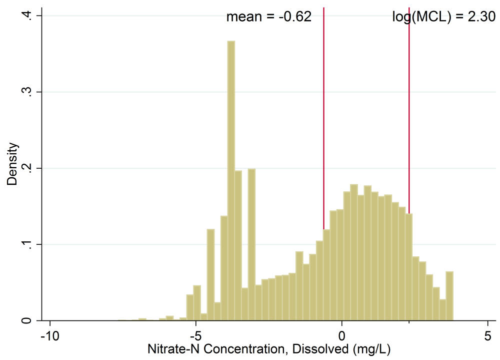
positively skewed in all categories. To limit the influence of outliers, I winsorise the readings at the 99th percentile for all nutrients.

To better understand the readings of nutrient concentrations in this dataset, I compare the data to Max Contaminant Level (MCL) set by EPA. EPA has established a drinking-water standard of 10 milligrams per liter (mg/L) for nitrate-nitrogen and 1 mg/L for nitrite-nitrogen [U.S. Environmental Protection Agency, 1995]. Nitrate concentrations in natural ground waters are usually less than 2 mg/L [Mueller et al., 1995]. Phosphorus is generally considered safe for human consumption, and U.S. EPA does not have an MCL for phosphorus in drinking water. Figure 2.6 present distribution of log of concentration values of nitrate, nitrite, and phosphorus. It also provides comparisons of the limit and actual nutrient concentrations. Figure 2.6 shows less than 10% observations have nitrate concentration above MCL and less than 1% observations have nitrite level greater than MCL.

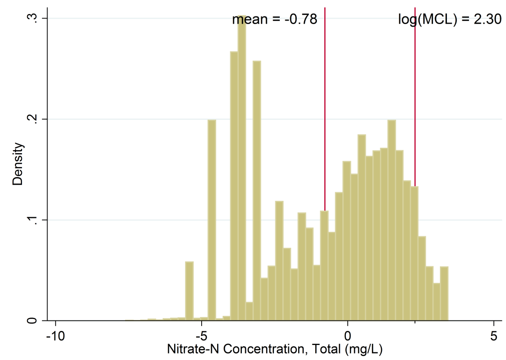
In the difference-in-differences analysis, I examine whether nutrient concentration in counties that have enrolled in the CRP (treatment group) is different from counties with zero CRP land (control group) in the years before and after the implementation of the CRP. I also use lead (dissolved) and cadmium (dissolved) to perform a placebo test. The primary source of lead in groundwater is dissolution of soil and earth crust, leaded gasoline, and water distribution system. Cadmium in groundwater mainly comes from mining activity, industrial waste, and combustion of fossil fuels. Since lead and cadmium in groundwater mostly come from non-agricultural activities, the placebo test will tell whether differences (if any) of nutrient concentrations between treatment group and control group are results of the implementation of CRP. Lead and cadmium data also come from the USGS. I use station-year level data for the difference-in-differences analysis instead of station-month level because of the limitation of the number of observations for pollutants.

2.3.3 Other datasets

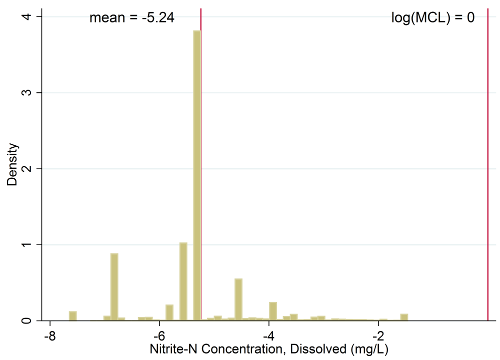
I include complete aquifer characteristic data by matching monitoring stations to the U.S. aquifer map from USGS (Figure 2.5). This process allows me to include a rock-type



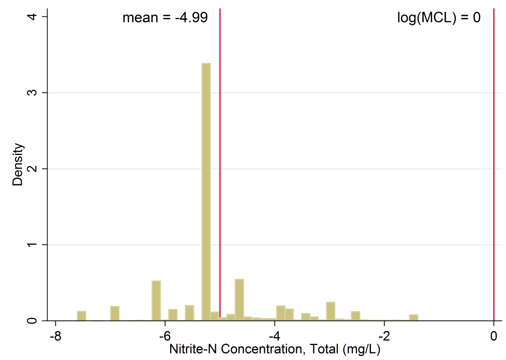
(a)



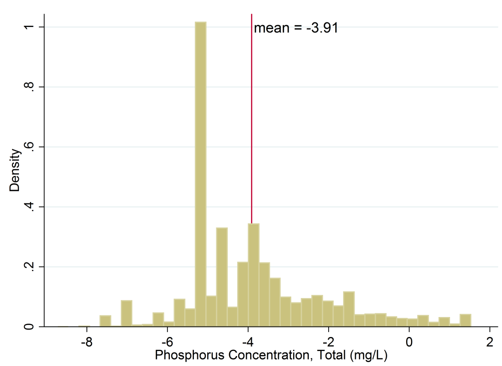
(b)



(c)



(d)



(e)

Notes: This figure shows the distribution of nitrate_N, nitrite_N and phosphorus concentration from water quality dataset using log of concentration readings. The MCL for nitrate-N is 10 mg/L, and 1 mg/L for nitrite-N. The mean value and MCL value are noted with red vertical lines in each figure. The mean of each distribution (in log value) is -0.62, -0.78, -5.24, -5, -3.92 respectively, and the percentage of observations that exceeds MCL is 9.33%, 6.87%, 0.27%, and 0.22% respectively (There is no MCL for phosphorus in the U.S.). The peaks below zero in figure (a), (b) and (e) are results from replacing "below quantification limit" or"not detected" values with one-half of corresponding " historically lower limit reported" values. Please see Figure A.1 for distributions of nutrient concentrations with no replacement for "below quantification limit" values.

Figure 2.6: Distribution of log of nutrient concentrations

fixed effect in the model as an approximation of the easiness of percolation. Rock type is not part of EBI, so it is exogenous to the decision of CRP enrollment. Rock types include carbonate-rock aquifers, igneous and metamorphic-rock aquifers, sandstone and carbonate-rock aquifers, sandstone aquifers, semiconsolidated sand aquifers, unconsolidated sand and gravel aquifers for major aquifers, and the rest belongs to minor (confined) aquifers. It is impractical to precisely sort the permeability of each aquifer by rock-type or by location. In general, carbonate-rock aquifers, sandstone and carbonate rock aquifers, unconsolidated sand and gravel aquifers have higher permeability and hydraulicly conductivity than the other types. Minor aquifers are also called confining units, which generally have low permeability (many are only permeable when fractured) and unproductive by unit. Aquifers with high permeability are more prone to contamination, and high hydraulic conductivity means easier transportation of contaminants in groundwater.

2.4 Model

This paper employs a fixed effects model to estimate the impact of CRP enrollment on nutrient concentrations in groundwater. I estimate the following model:

$$N_{icym} = f(ratio_{cy}) + \alpha_c + \alpha_y + \alpha_m + \alpha_r + \epsilon_{icym}, \tag{2.1}$$

where N_{icym} stands for log of nutrient concentration value at station i, in county c, for month m in year y; $f(ratio_{cy})$ is a function of land use at county c in year y; α_c is a county fixed effect; α_y and α_m are year and month fixed effects respectively; and α_r is a rock-type fixed effect. The rich fixed effects control for a number of potential confounding factors such as soil quality, underground flow direction and crop type. The ability to include rich fixed effects is a key contribution of this work.

In this paper I use two specifications of $f(ratio_{cy})$. The first is $f(ratio_{cy}) = \beta ratio_{cy}$, where $ratio_{cy}$ is a moving average of the CRP ratio (CRP/(CRP+cropland)). The second specification is $f(ratio_{cy}) = \gamma_1 R_{CRP_{cy}} + \gamma_2 R_{noncrop_{cy}} + \gamma_3 R_{rangeland_{cy}} + \gamma_4 R_{pastureland_{cy}} +$

$\gamma_5 R_{urban_{cy}}$, where $R_{(l)}$ is a moving average of the ratio of each land use l to acres of total land (acres of land use l/acres of total land). The category "noncrop" contains acres in forest, small water area, large water area, rural transportation, minor land and federal. Thus, the reference category in the second specification is cropland.

I use long-term moving average of land use variables in $f(ratio_{cy})$ because unlike surface water, groundwater has a much slower recharge rate. McMahon et al. [2011] shows mean groundwater residence time in different U.S. aquifers ranges extensively. Residence time varies mostly with the depth of groundwater: deeper groundwater usually has longer residence time. Drinking water wells are generally shallow groundwater that typically ranges from 10 to 60 feet deep. A regional research using Connecticut public well data shows the mean residence time of groundwater in such shallow aquifers is around five years [Starn and Brown, 2007]. The length of groundwater residence time provides a reference of how long nutrients in groundwater stay underground. Lerner and Harris [2009] also find that nitrate pollution in groundwater is a long-term problem, and contemporaneous changes in land use do not substantively affect nitrate accumulation. CRP contracts last 10 to 15 years, which means there is a long window for changes to occur to local soil quality, nutrient residue level in soil, and nutrient leaching rate. Without better knowledge of the exact hydrology profile of each monitoring station, I use different lengths of time when calculating the CRP land moving average[12], and my preferred order of moving average matches with findings in Starn and Brown (2007).

The first specification of $f(ratio_{cy})$ best describes the fact that most CRP land comes from putting cropland out of production. Results from this specification intuitively explain how retiring erodible land correlates with nutrient concentration in groundwater. However, it does not offer a bigger picture of how CRP land affects groundwater quality compared with other land use categories. The second specification of $f(ratio_{cy})$ provides a comparison

[12] See results in Figure 2.7

of the impact of different land use on nutrient concentrations in groundwater, using cropland as a reference.

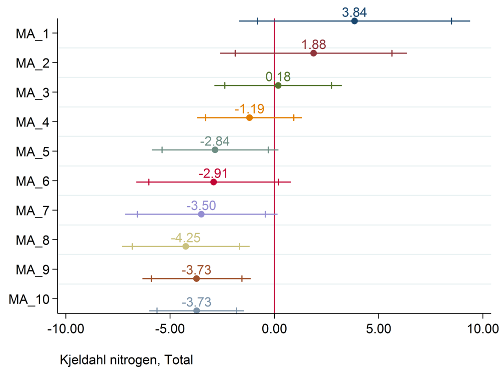
2.5 Empirical Result

I begin by selecting the preferred order of the moving average in $f(ratio_{cy})$ with the following equation:

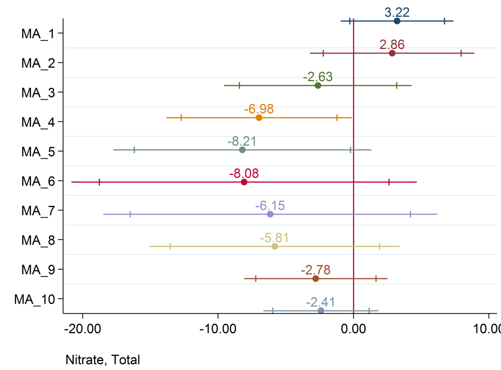
$$N_{icym} = MA_{cy,n} + \alpha_c + \alpha_y + \alpha_m + \alpha_r + \epsilon_{icym},\ n = 1, 2, ..., 10 \tag{2.2}$$

where $MA_{cy,n}$ stands for the moving average of CRP ratio in the past n years at county c in year y. Figure 2.7 plots how the regression results vary with different time range for representative nutrients. Y-axis in the plot is the moving average of CRP ratio in the past one to ten years. These plots show that the impact is smallest when only considering contemporaneous enrollment, and the impact increases as the moving average term increases. I find the greatest impact of CRP ratio occurs when considering the five-year moving average. This result is intuitive and is consistent with findings of shallow groundwater residence time in Starn and Brown (2007).

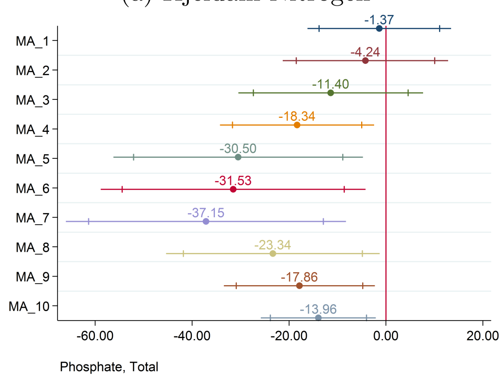
Table 2.2 presents regression results estimating equation (1) using the five-year moving average CRP ratio as the independent variable and total nitrate concentration as the dependent variable. Year fixed effect captures the time trend, and month fixed effect captures seasonality. County fixed effect captures time invariants unobservables such as soil quality, and rock fixed effect absorbs unobservables on aquifer features. Column (1) presents results from OLS model where no fixed effect is used. It shows the CRP ratio is positively correlated with total nitrate concentration in groundwater. Column (2) and (3) include time and seasonality controls and show a counter-intuitive, positive and statistically significant correlation between CRP and nitrate concentration. These results are consistent with the results in Sprague and Gronberg [2012], who find a positive correlation between CRP area and nutrient concentration in surface water using an OLS model. However, when county fixed effects are included in the model, the coefficient on the CRP ratio becomes



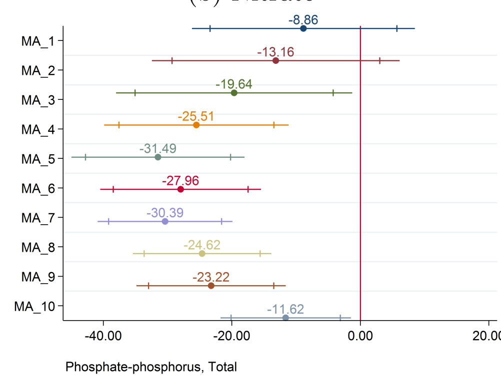
(a) Kjeldahl Nitrogen



(b) Nitrate



(c) Phosphate



(d) Phosphate-phosphorus

Notes: This figure shows regression results from equation (2). MA_n (n=1,2,3,...,10) in y-axis is the moving average of CRP ratio in the past n years. Each horizontal line represents 95 confidence interval and area between two marks represents 90 confidence interval.

Figure 2.7: Different orders of moving average of CRP ratio

negative and statistically significant, as is the result in column (5) when rock-type fixed effects are also included. Results in column (5) show if the CRP ratio increases by 1%, total nitrate concentrations in groundwater drop by about 7.9%. Table A.3 and Table A.4 contain regression results from equation (1) using regional data.

Table 2.2: Regression results of total nitrate concentration

DEPENDENT VARIABLES	(1) Nitrate Total	(2) Nitrate Total	(3) Nitrate Total	(4) Nitrate Total	(5) Nitrate Total
CRP ratio	2.609	2.688*	2.989**	-8.950*	-8.213*
	(1.812)	(1.495)	(1.389)	(5.084)	(4.843)
Year FE	No	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	Yes	Yes
County FE	No	No	No	Yes	Yes
Rock FE	No	No	No	No	Yes
Cluster	County	County	County	County	County
Obs.	10,495	10,495	10,495	10,416	10,416
R-squared	0.002	0.068	0.080	0.518	0.524

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table shows regression results estimating equation (1). The independent variable "CRP ratio" equals to the past five-year average of $\frac{Acres\ in\ CRP}{Acres\ in\ CRP\ and\ cropland}$. Note that when all levels of fixed effect are included, the coefficient of ratio is negative and statistically significant. This indicates when considering unobservables in all level, the higher the CRP ratio, the lower nitrate concentration in groundwater. From results in column (5), 1% increase of CRP ratio reduces total nitrate concentration in groundwater by 7.9%.

Table 2.3 shows the results of other important categories of nitrogen compound and phosphorus, and Table A.1 contains results of all 19 nutrient categories. Kjeldahl nitrogen measures the sum of NH_3 and organic nitrogen concentration. Column (3) in Table 2.3 shows a 1% increase in the five year moving average of CRP ratio leads to about 26.7% decrease in phosphate-phosphorus concentration in groundwater. These results indicate that allocation of cropland to CRP land contributes to better groundwater quality in the long run.

Table 2.3: Regression results of representative nutrients

DEPENDENT VARIABLES	(1) Kjeldahl Nitrogen Total	(2) Phosphate Total	(3) Phosphate-phosphorus Total
CRP ratio	-3.264*	-30.99**	-31.03***
	(1.745)	(13.44)	(6.658)
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes
Cluster	County	County	County
Obs.	11,780	5,078	5,848
R-squared	0.465	0.541	0.548

Robust standard errors in parentheses

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

Notes: Kjeldahl nitrogen measures the sum of NH_3 and organic nitrogen. The results shows that 1% increase in CRP ratio leads to 3.2%, 26.6% and 26.7% decrease in Kjeldahl nitrogen, phosphate, phosphate-phosphorus concentrations in groundwater respectively.

Table 2.4 presents regression results using the second CRP ratio specification on nitrate. Each independent variable represents the five-year moving average of one type of land use over total land acreage at county level, and by construction, cropland is the reference category. Similar to Table 2.2, when county fixed effects are omitted, I find that CRP land is positively correlated with total nitrate concentrations, and adding time trend and seasonality do not alter the result. Column (4) and (5) show that after adding county and rock-type fixed effect, the explaining power of this model increases. They also show the fraction of CRP land in total land is negatively correlated with total nitrate concentration in groundwater, but the results are statistically insignificant. This shows the regression results are sensitive to the specification of $f(ratio_{it})$ in the main model. The inclusion of other land use variables weakens the correlation between CRP land and nutrient concentration. The results are consistent with using the first specification of $f(ratio_{it})$ but are also much noisier. Table A.2 shows the regression results of all 19 categories of nutrients using the second specification.

Table 2.4: Regression results, second specification

	(1)	(2)	(3)	(4)	(5)
DEPENDENT VARIABLES	Nitrate Total	Nitrate Total	Nitrate Total	Nitrate Total	Nitrate Total
R_{CRP}	1.904	1.489	2.139	-42.43	-36.65
	(4.156)	(3.803)	(3.390)	(39.64)	(37.44)
$R_{noncrop}$	-0.624	-0.563	-0.442	0.144	0.983
	(1.562)	(1.205)	(1.116)	(8.192)	(8.215)
$R_{rangeland}$	2.737*	2.562**	2.706**	-10.37	-14.66**
	(1.625)	(1.247)	(1.140)	(7.663)	(6.109)
$R_{pastureland}$	-0.961	0.601	0.432	10.96	7.579
	(1.824)	(1.907)	(1.925)	(12.68)	(11.42)
R_{urban}	2.566	2.842**	2.872**	4.932	10.11
	(1.677)	(1.359)	(1.336)	(8.424)	(7.044)
Year FE	No	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	Yes	Yes
County FE	No	No	No	Yes	Yes
Rock FE	No	No	No	No	Yes
Cluster	County	County	County	County	County
Obs.	10,674	10,674	10,674	10,593	10,593
R-squared	0.106	0.161	0.170	0.521	0.527

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: $R_{(l)}$ stands for the five-year average of $\frac{Acre\ in\ land\ use\ l}{Total\ land\ acreage}$. This table shows the results are sensitive to the inclusion of other land use variables. $R_{noncrop}$ contains six non-cropland categories: forest, small water, large water, rural transportation, minor land and federal. The reference category in this specification is $R_{cropland}$.

2.6 Results from difference-in-differences analysis

To further examine and test whether the CRP has an impact on nutrient concentration in groundwater, I also perform a difference-in-differences analysis with station-year level data generated from the same dataset. Figure 2.1 shows how CRP enrollment changed with time. This bar chart shows that there was a transition period from 1986 to 1991 when the CRP was first implemented and CRP enrollment experienced substantial growth. During these years CRP enrollment increased dramatically from zero to around 33 million acres, then after 1991 CRP enrollment became stable at around 34 million acres. I take advantage of this dramatic change in transition period and consider the whole period as the event of "early enrollment of CRP". This allows me to construct a pre-event and post-event sample to perform a difference-in-differences analysis using counties with and without CRP land.

To study this, I estimate the regression:

$$N_{ij} = \beta_0 + \beta_1(post92 \times yescrp) + \beta_2 post92 + \omega_i + \omega_r + \epsilon_{ij} \tag{2.3}$$

where i stands for each monitoring station and j stands for year. $post92$ is a binary indicator that equals zero for observations before 1992 and one for observations after 1992. $yescrp$ is also a binary indicator which equals one if a county has enrolled in the CRP in the study period, and equals zero if a county never enrolled in the CRP. ω_i is a monitoring station fixed effect, and ω_r is the rock fixed effect. To estimate the before and after transition period effect, I use a subset of my main dataset, which contains only observations from 1982 to 1985, and from 1992 to 1997. It also consists of only counties that either never enrolled in the CRP in this period, or counties that enrolled in CRP in 1987 (the earliest record in the NRI database). As CRP contracts are long term, it makes sense for me to consider 1986 to 1991 as one event, which is the early enrollment of CRP. Therefore, 1982 to 1985 is the "pre-event" period, and 1992 to 1997 is the "post-event" period. This construction also takes into consideration that changes in land use may need years to affect groundwater quality.

Most nutrients from the main dataset have too few observations left for this model. Table 2.5 shows regression results of equation (2) using log value of nitrate, nitrite and inorganic nitrogen concentrations as the dependent variables. Column 1 and 2 in Table 2.5 show that after the transition period, counties that have enrolled in the CRP in 1987 have around 55% less nitrate (dissolved) and around 34% less nitrite (dissolved) in groundwater, compared with counties that did not enroll in the CRP. I also run regressions with cadmium and lead as placebo tests. Cadmium and lead are pollutants that are harmful to human health, and they most likely come from non-agricultural activities. Column 4 and 5 shows that enrolling in CRP does not have statistically significant effect on cadmium and lead concentrations in groundwater, and this provides evidence that the change of nutrient concentration in groundwater comes from the implementation of the CRP.

Table 2.5: Effects of the CRP on nutrient concentration in groundwater

	(1)	(2)	(3)	(4)	(5)
DEPENDENT VARIABLES	Nitrate Dissolved	Nitrite Dissolved	Inorganic Nitrogen Dissolved	Cadmium Dissolved	Lead Dissolved
post92×yescrp	-0.552**	-0.339**	-0.0722	0.263	-0.0520
	(0.217)	(0.165)	(0.0846)	(0.233)	(0.402)
post92 = 1	0.00869	-0.176***	0.0566	-0.277**	-1.809***
	(0.0788)	(0.0658)	(0.0595)	(0.112)	(0.221)
Station FE	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes
Cluster	Station	Station	Station	Station	Station
R-squared	0.934	0.732	0.933	0.672	0.758
Obs.	11,630	11,006	18,126	3,890	4,101
Stations	4266	4050	6715	1663	1740

Robust standard errors in parentheses

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

Notes: This table presents results from the CRP difference-in-differences analysis. Dependent variables are the log value of different nutrient concentration at the station-year level. $post92$ is an indicator variable for the post-transition period. yescrp is an indicator variable of whether a county was enrolled in the CRP in 1987. $post92 \times yescrp$ is the interaction term and is the interested variable in these regressions. The parameter of term $post92 \times yescrp$ presents that after transition period, what would be the difference of log value of nutrient concentration in counties that enrolled in the CRP compared to counties that was not enrolled in the CRP.

Equation (4) below is a more flexible form of equation (3).

$$N_{ij} = \sum_{j=2}^{29} \gamma_i(yescrp * year_j) + \sum_{j=2}^{29} \mu_i(year_j) + \alpha_i + \alpha_r + \epsilon_{ij} \qquad (2.4)$$

here i stands for monitoring station and j stands for year. α_i are station fixed effects, and α_r are rock fixed effects. This regression result shows year to year difference in counties that enrolled in the CRP and counties that did not enroll in the CRP, using 1986 as a reference year. Figure 2.8a shows that counties without and without CRP land experienced non-differential pre-trend of nitrate (dissolved) concentration in groundwater before 1986, After 1986, nitrate concentration in counties that enrolled in the CRP started to decrease, and became stable after about five years (after 1991). This is consistent with the lag effect discussed when estimating equation (1). Figure 2.8b shows there is no visible trend before and after 1986 for inorganic nitrogen, and Figure 2.8c shows there is a difference in trend before and after 1986 for nitrite. Figure 2.9a and Figure 2.9b present the results using concentrations of cadmium and lead in groundwater. Figure 2.9a shows cadmium did not have a change in trend. Lead seems to see decreasing trend after 1996, but no apparent change in trend between 1982 and 1996 and the decrease did not converge to a steady level. These figures are consistent with regression results listed in Table 2.5.

To better understand the result from the difference-in-differences method, I compare the point estimate in column 1, Table 2.5, to results from equation (1). The average CRP ratio in the treatment group in the difference-in-differences dataset is 0.108, and zero in the control group. From the results in Table 2.2, a 10.8% positive difference in CRP ratio means $10.8 \times 7.9\% \approx 85.3\%$ less nitrate concentration in groundwater. This number is larger in magnitude than the result in column 1 in Table 2.5 but is consistent with the converged level in Figure 2.9a. An 85.3% decrease seems like a drastic change, but the starting value of nitrate concentration in groundwater is a small number. For example, the mean nitrate concentration in the treatment group in the difference-in-difference sample is 5.28 mg/L, and an 85.3% decrease means a reduction of 4.5 mg/L dissolved nitrate in groundwater[13].

[13] The mean of CRP ratio in the full sample is 0.043. Mean nitrate concentration in the full sample is 0.54mg/L

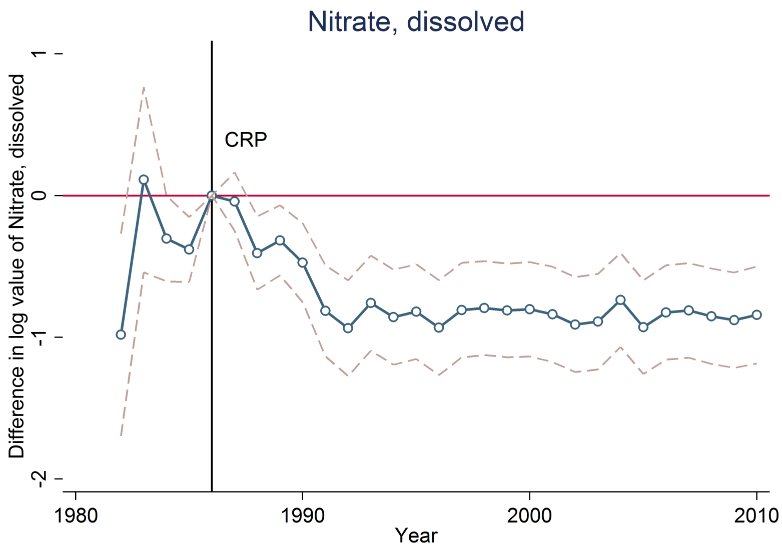


Figure 2.8a: Year to year difference, nitrate

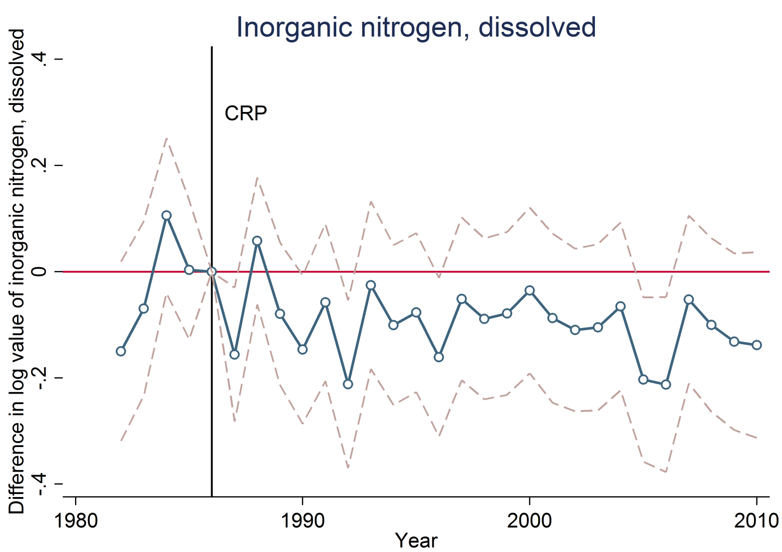
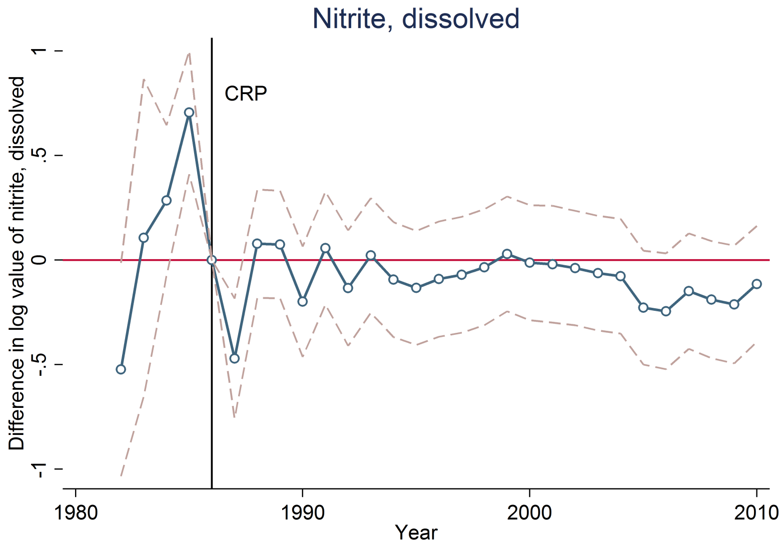


Figure 2.8b: Year to year difference, inorganic nitrogen



Notes: These figures plot the year to year difference in nutrient concentration between counties that enrolled in the CRP and counties with zero CRP land, using 1986 as the reference year. Dashed lines represent 95 percent confidence intervals. The dependent variable is log value of nutrient concentration at station-year level. Regressions control for year and station fixed effects.

Figure 2.8c: Year to year difference, nitrite

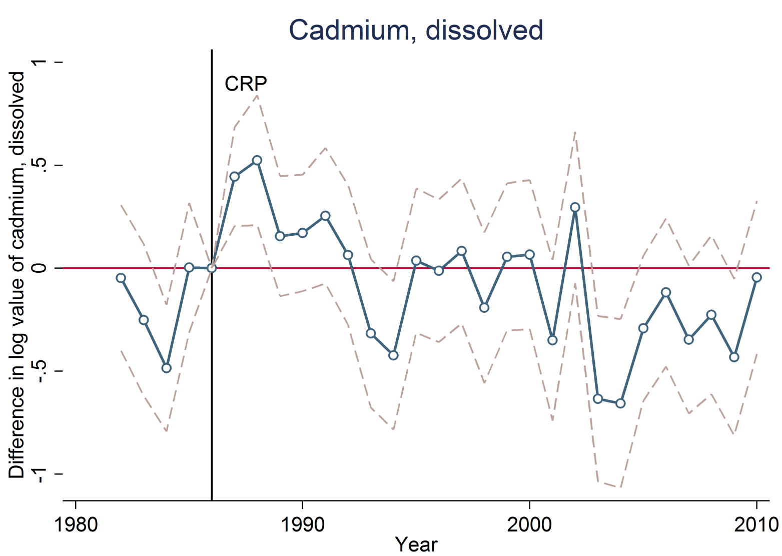
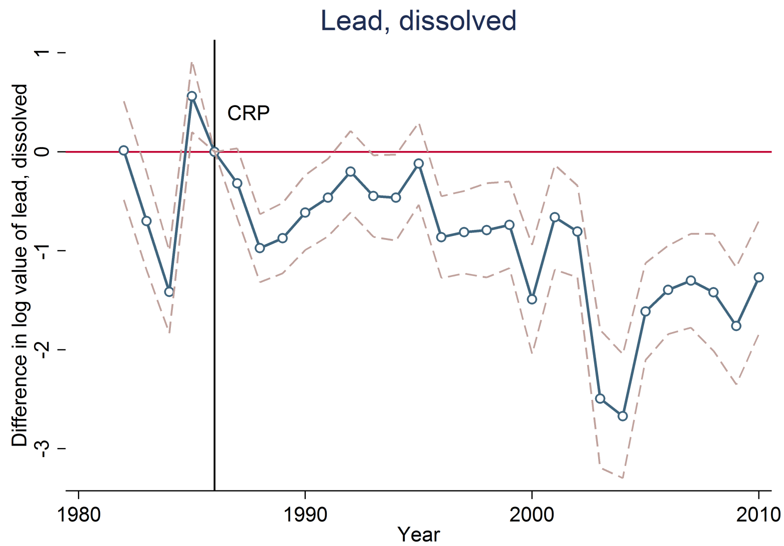


Figure 2.9a: Year to year difference, cadmium



Notes: These figures plot year to year difference of cadmium or lead concentrations between counties that enrolled in the CRP and counties with zero CRP land, using 1986 as the reference year. Dashed lines represent 95 percent confidence intervals. Dependent variable is log value of nutrient concentration at station-year level. Regressions control for year and station fixed effects.

Figure 2.9b: Year to year difference, lead

2.7 Conclusions and discussion

This paper studies how land use, especially land enrolled in CRP, influences nutrient concentrations in groundwater. Using a fixed effects model, I find evidence that long-term moving average of CRP ratio is negatively correlated with nitrate concentration in groundwater: 1% increase in the five-year moving average of CRP ratio leads to 7.9% reduction in nitrate concentration and 27% reduction in phosphorus concentration in groundwater. However, I find this result is sensitive to model specification. After adding other land use variables to the main model, there is weak evidence (consistent but noisy) that long-term moving average of the fraction of CRP land in total land is negatively correlated with nutrient concentrations in groundwater. I further examine whether counties that enrolled in the CRP experienced different trends in nutrient concentration compared with counties that did not enroll in the CRP with a difference-in-differences analysis. I find that in the few years after the implementation of the CRP, counties that enrolled in the CRP has 55% less nitrate and 34% less nitrite in groundwater, compared with counties that did not enroll in the CRP.

Apart from finding that CRP enrollment has a negative effect on nitrate concentration in groundwater, I also find that there exists significant lag effect in this result. It takes at least five years after enrolling land in CRP for changes in nutrient concentrations in groundwater to become detectable. This finding is inspiring as nutrient pollution in groundwater is extremely difficult to deal with, such that prevention from future pollution is usually the best way to improve groundwater quality. This is also an important indicator that early termination of CRP contracts may lead to unexpected loss of groundwater quality. Groundwater is currently one category in EBI with relatively small weight (25 out of 100 in water quality category), but with findings in this paper, I hope to provide evidence that CRP enrollment may have a more substantial impact on groundwater quality than previously studied.

2.8 Bibliography

Babcock, B. A., Lakshminarayan, P., Wu, J., and Zilberman, D. (1996). The economics of a public fund for environmental amenities: a study of crp contracts. *American Journal of Agricultural Economics*, 78(4):961–971.

Böhlke, J.-K. (2002). Groundwater recharge and agricultural contamination. *Hydrogeology Journal*, 10(1):153–179.

Bruggeman, A., Mostaghimi, S., Holtzman, G., Shanholtz, V., Shukla, S., and Ross, B. (1995). Monitoring pesticide and nitrate in virginia's groundwater- a pilot study. *Transactions of the ASAE*, 38(3):797–807.

Burow, K., Dubrovsky, N., and Shelton, J. L. (2007). Temporal trends in concentrations of dbcp and nitrate in groundwater in the eastern san joaquin valley, california, usa. *Hydrogeology Journal*, 15(5):991–1007.

Claassen, R., Cattaneo, A., and Johansson, R. (2008). Cost-effective design of agri-environmental payment programs: Us experience in theory and practice. *Ecological economics*, 65(4):737–752.

Crutch-field, S., Cooper, J., and Hellerstein, D. (1997). Benefits of safer drinking water: The value of nitrate reduction. agricultural economic report 752. *US Department of Agriculture, Economic Research Service, Food and Consumer Economics Division*.

Dinnes, D. L., Karlen, D. L., Jaynes, D. B., Kaspar, T. C., Hatfield, J. L., Colvin, T. S., and Cambardella, C. A. (2002). Nitrogen management strategies to reduce nitrate leaching in tile-drained midwestern soils. *Agronomy journal*, 94(1):153–171.

Fan, A. M. and Steinberg, V. E. (1996). Health implications of nitrate and nitrite in drinking water: an update on methemoglobinemia occurrence and reproductive and developmental toxicity. *Regulatory toxicology and pharmacology*, 23(1):35–43.

Farm Service Agency (2012). Conservation reserve program: annual summary and enrollment statistics.

Gardner, K. K. and Vogel, R. M. (2005). Predicting ground water nitrate concentration from land use. *Ground water*, 43(3):343–352.

Kumazawa, K. (2002). Nitrogen fertilization and nitrate pollution in groundwater in japan: Present status and measures for sustainable agriculture. *Nutrient Cycling in Agroecosystems*, 63(2-3):129–137.

Lerner, D. N. and Harris, B. (2009). The relationship between land use and groundwater resources and quality. *Land Use Policy*, 26:S265–S273.

Mastrocicco, M., Colombani, N., Castaldelli, G., and Jovanovic, N. (2011). Monitoring and modeling nitrate persistence in a shallow aquifer. *Water, Air, & Soil Pollution*, 217(1-4):83–93.

McMahon, P., Plummer, L., Böhlke, J., Shapiro, S., and Hinkle, S. (2011). A comparison of recharge rates in aquifers of the united states based on groundwater-age data. *Hydrogeology Journal*, 19(4):779.

Mueller, D. K., Hamilton, P. A., Helsel, D. R., Hitt, K. J., and Ruddy, B. C. (1995). Nutrients in ground water and surface water of the united statesan analysis of data through 1992. *US Geological Survey Water-Resources Investigations Report*, 95(4031):74.

NGWA (2016). Groundwater use in the united states of america.

Nolan, B. T. (2001). Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the united states. *Ground Water*, 39(2):290–299.

NRCS. Buffer strips: Common sense conservation.

NSAC. Conservation reserve program.

Spalding, R. F. and Exner, M. E. (1993). Occurrence of nitrate in groundwatera review. *Journal of environmental quality*, 22(3):392–402.

Sprague, L. A. and Gronberg, J. A. M. (2012). Relating management practices and nutrient export in agricultural watersheds of the united states. *Journal of environmental quality*, 41(6):1939–1950.

Starn, J. J. and Brown, C. J. (2007). *Simulations of ground-water flow and residence time near Woodbury, Connecticut*. US Department of the Interior, US Geological Survey.

Strebel, O., Duynisveld, W., and Böttcher, J. (1989). Nitrate pollution of groundwater in western europe. *Agriculture, ecosystems & environment*, 26(3-4):189–214.

U.S. Environmental Protection Agency (1995). Drinking water regulations and health advisories.

USDA (2016). Background.

USGS (2010). Groundwater use in the united states.

Viers, J. H., Liptzin, D., Rosenstock, T. S., Jensen, V. B., Hollander, A. D., McNally, A., King, A. M., Kourakos, G., Lopez, E. M., Mora, N. D. L., Fryjoff-Hung, A., Dzurella, K. N., Canada, H., Laybourne, S., McKenney, C., Darby, J., Quinn, J. F., and Harter, T. (2012). Assessing nitrate in californias drinking water with a focus on tulare lake basin and salinas valley groundwater.

Wang, H., Gao, J.-e., Li, X.-h., Zhang, S.-l., and Wang, H.-j. (2015). Nitrate accumulation and leaching in surface and ground water based on simulated rainfall experiments. *PloS one*, 10(8):e0136274.

3. RELIABLE DRINKING WATER SUPPLY AND CITIES' RESILIENCE TO DROUGHT

Access to safe and reliable drinking water is essential to economic activities, especially during extreme weather events. Drinking water facilities' choice of supply source determines their ability to maintain supply for domestic needs and economic activities under unexpected weather shock. This paper examines how the choice of drinking water supply source affected population resilience during droughts in the last century. We use the most comprehensive dataset on municipal drinking water facilities from the 1920s to the 1960s, a period that witnessed a substantial expansion in the use of groundwater as a municipal drinking water supply source as well as a series of severe droughts in the United States. Using a fixed effect model we find evidence that having groundwater as a source of municipal water supply helps with population resilience during drought, especially for communities that locate on major aquifers. This finding provides evidence that the choice of supply source matters in maintaining population during drought and provides insights on the design of drinking water infrastructure in developing areas.

3.1 Introduction

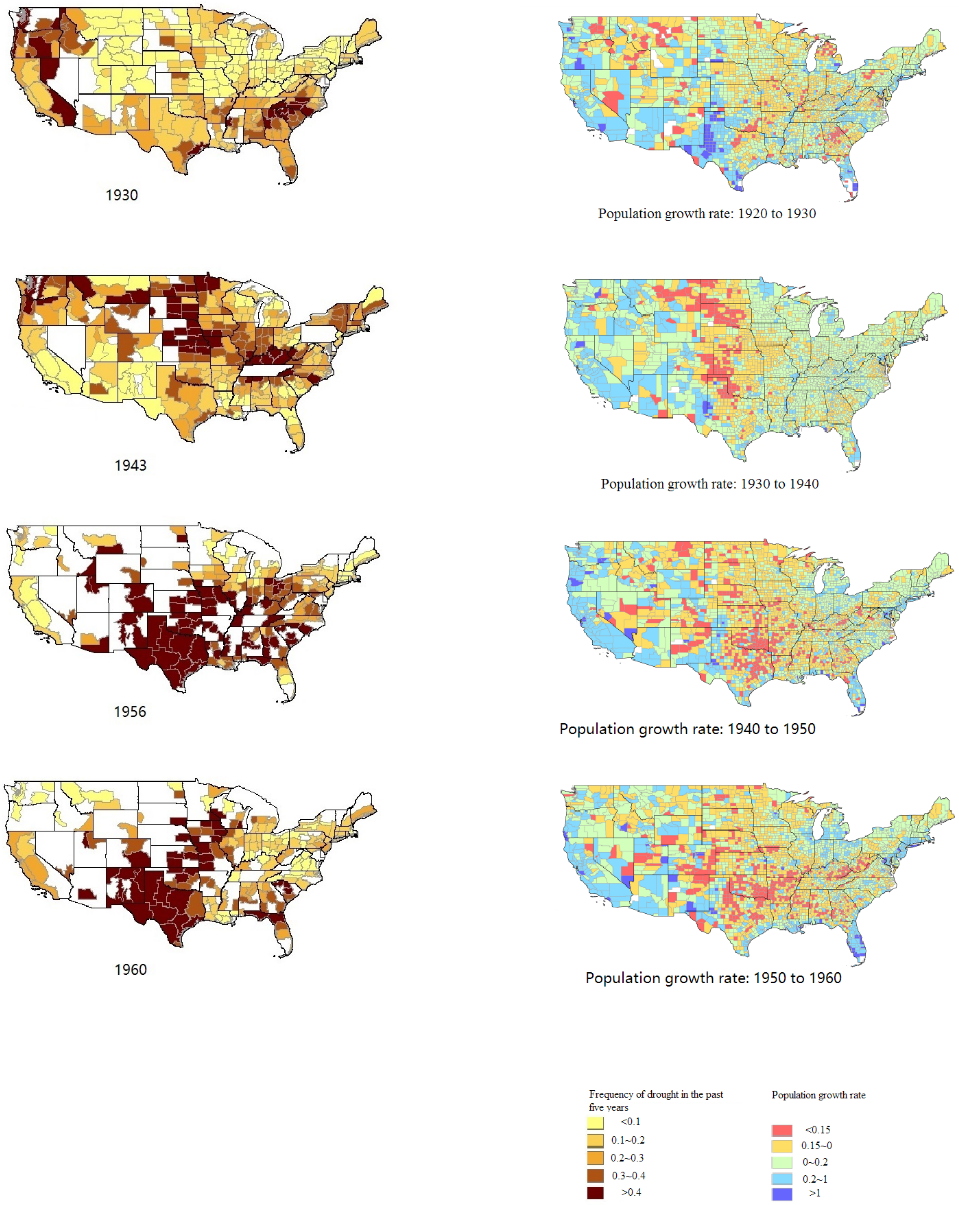
Water security is essential to human beings, especially during extreme weather events. City developers face the challenge of designing drinking water infrastructure that can tolerate future population growth, increasing economic activities and extreme weather shocks. This paper examines the relationship between choices of municipal drinking water supply sources and population resilience during drought. Previous work has studied how climate change and severe weather events affect fresh water system, land use and urbanization process in the United States. Another topic that was widely studied is how to improve public drinking water system to adapt to increasing need in modern communities and more frequent droughts due to climate change [Mosley, 2015, Nace et al., 1965, Singh et al., 2014, Hansen and Libecap, 2003]. Many researchers have also pointed out the importance of groundwater during drought in modern societies [Foster, 2001, Taylor et al., 2013, Baker et al., 2004, Mosley, 2015, Singh et al., 2014]. However, few studies examined the beginning of modern drinking water infrastructure to evaluate how groundwater matters during drought when the technology was not well developed. This paper fills this gap by exploring the relationship between choice of drinking water supply source and population resilience during drought in the U.S. We compiled the most comprehensive drinking water facility dataset that covers a wide time range from the 1920s to the 1960s. With a fixed effect model, we find that in large communities, having groundwater as drinking water supply source significantly improve the resilience of population during drought, especially in communities located on a major aquifer.

Organized municipal drinking water supply systems appeared as early as ancient Greece. Today drinking water infrastructure remains a vital part of city planning and is a long-term investment that expects a life of over 100 years (Marshall, 2008). In the United States, the government spent about \$10 billion annually on water mains in the 1920s and about \$30 billion annually by the early 1950s. A large part of U.S. drinking water infrastructure people now rely on was built in mid 20th century, and the cumulative cost of that original investment exceeds \$2 trillion (EPA, 2011). By studying the beginning of modern drinking water infrastructure history, one gains valuable information on the function of drinking water infrastructure on city growth and population resilience under an environment with relatively few regulations.

Increasing population in large communities means larger demand for treated water. It is known that groundwater, due to its generally high quality and easiness to access, has always been the preferred source of municipal water supply, for both domestic and industrial uses

[Foster, 2001]. However, whether different sources of municipal water supply has impacted urbanization is typically challenging to answer ex-ante. Most cities grow fastest during a relatively wet period, and city planners and dwellers may overlook the possibility of severe drought in the future [Nace et al., 1965]. The United States went through three severe drought periods in the 20th century: the Dust Bowl drought in the 1930s, the 1950s drought that hit Southwestern United States (especially Texas) the most, and the 1987-1989 drought which started from the Southeastern United States and destroyed crops almost nationwide. Of these three periods, the first two are covered by dataset used in this paper.

The maps in the right column of Figure 3.1 present census population growth rate at the county level from the 1920s to 1960s. They show that during the 1930s, the Great Plains region experienced the most severe population loss. In fact, approximately 250,000 people moved to from Midwestern states to California by 1940, and around 2.5 million population left the Plains states in the 1930s [Hansen and Libecap, 2003, Martin, 2008]. Figure 3.1 also shows that population loss areas in the 1950s and 1960s coincide with drought-hit areas during that period. Notice that though California, especially Southern California, was stricken by drought in the 1950s, it did not experience population loss. The municipal drinking water facilities dataset shows that in California, the number of water facilities that have access to groundwater supply increased significantly after 1945, compared with the rest of the country (five out of six districts in the sample have added groundwater supply sources). This finding supports the assumption that having access to groundwater supply helps with population resilience during drought.



Notes: Maps on the left of Figure 1 describes frequency of drought in the past five years at each survey year. Maps on the right of Figure 1 presents the population growth rate using census population. Notice how drought hit areas in 1943 drought map coincides with the areas that experienced most population loss in population loss in population growth rate map, also how population loss areas move from north great plain areas down to Texas, also matches the drought hit areas in the 1950s.

Figure 3.1: Population growth rate and drought hit areas

This paper takes advantage of detailed records of supply source and population supplied information from the drinking water facility dataset and repeated occurrence of droughts in the last century. We focus on the access to groundwater supply source in this paper because this period of interest witnessed the start of adopting groundwater supply source in municipal drinking water facilities in the whole nation. Many developing countries and districts still use similar drilling technology and face similar regulations as the United States did during that period. We hope this research can provide insight on better utilization of water resources and water infrastructures that can help with population resilience under extreme weather shock.

The paper proceeds as follows: section 2 describes the datasets used in this paper, section 3 provides summary statistics, section 4 contains model and regression results, and the final section concludes.

3.2 Data

The data are composed of two major components: municipal water supply data and drought index constructed from the Palmer drought index.

3.2.1 Municipal water supply data

This municipal water supply dataset is the first to aggregate detailed infrastructure information from the 1920s to 1960s in the U.S. All data are digitized from the original municipal drinking water treatment facility surveys carried out mostly by Department of Public Health. The data contain 11 distinct years: 1924, 1930, 1943, 1945, 1954, 1956, 1958, 1960, 1962, 1963 and 1964.

The 1924 data come from the Filtration Plant Census [Gillespie, 1925] and contain information on water filtration plants with a capacity of 1 million gallons per day (M.G.D) or greater. The data include the name of each city, its source of water supply, filtration capacity, population supplied, date of installation and simple remarks of what treatment technology was used. The 1930 data (Wolman, 1933) on water purification plants are categorized into plants with treatment other than simple chlorination and plants with only chlorination. It appears that in the first case surface sources of supply takes up around 85% of all facilities, while in the latter surface sources of supply and ground sources of supply are almost equal. 1943 data ("National Census of Water", 1943) contains water treatment plants that serve communities with a population of 100 or more in the continental United States. In 1943 and later years, water supply surveys provide more detailed information regarding capacity, water storage, treatment technology and ownership. 1945 data ("Inventory of Water", 1945) is the only year that contains data of both water treatment facilities and sewage facilities.

1954, 1956, 1960, 1962, 1964 and part of 1958 data come from a series of surveys: Municipal water facilities – communities of 25,000 population and over by U. S. Public health service ("Municipal Water Facilities", 1956-1964). Data from these years are more consistent in format than other years and have most detailed facility characteristics. However, they also contain much fewer observations than 1945, 1958 and 1963 surveys ("Municipal Water Facilities", 1958 and 1963) due to their population threshold.

Though water supply data are at facility level, many large cities have multiple municipal water supply facilities. To best preserve the details from this unique dataset, all water supply data are aggregated to community (mostly city or township) level for the main analysis. Populations are summed to the community level, and supply source is categorized as whether the whole community has groundwater supply source or surface water supply source (both are binary indicators).

3.2.2 Palmer drought index data

Palmer drought severity index (PDSI) (Historical Palmer Index) is used as an indicator of the severity of droughts in this paper. The data come from National Oceanic and Atmospheric Administration (NOAA). PDSI data are recorded monthly from over 300 climate divisions and usually range from -6 to 6. Moderate drought is classified by PDSI value between -2 and -2.99, and severe drought is classified by PDSI value below -3, and extreme drought by PDSI value below -4. This paper uses the five-year frequency of PDSI below -2 to establish a measure of how often an area experiences drought.

3.2.3 U.S. aquifer data

Aquifer information comes from USGS website. We use ArcGIS to map observations in our water facility dataset to U.S. aquifers. Districts that overlap with both major and minor aquifer are categorized as where their center point locates. Minor aquifer, or confining unit, means such areas are underlain by low-permeability deposits and rocks, and have very limited groundwater supply. Large areas in the eastern, northeastern and north-central parts of United States are also categorized as minor aquifers because these areas are underlain by crystalline rock which has minimal permeability. However as these rocks extend over large areas, they are the only reliable source of water supply in many places and provides large amount of groundwater ("Aquifer Basics", 2016). Of all the 19,581 distinct districts in the drinking water dataset, 12,117 districts locate on major aquifers and the rest 7,464 districts locate on minor aquifers. All districts are mapped to one of these two aquifer categories.

3.3 Summary statistics

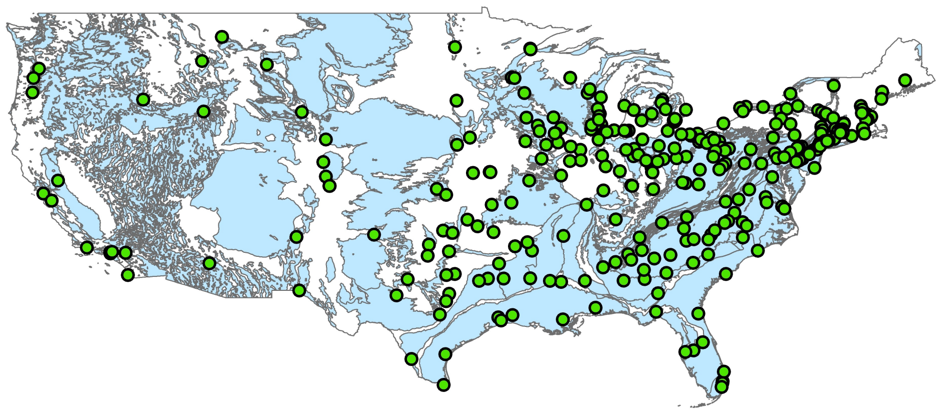
Table 3.1 summarizes the number of observations in our data by year. The whole dataset contains 58,228 observations from 11 survey years and 19,581 unique districts. The number of observations per year varies between 492 and 18,557 because some surveys contain only communities with over 25,000 population, and other surveys contain communities with over 100 population. Another way to look at the dataset is to check whether it provides a series of reasonably consistent observations over all the years. In the full dataset there are 11 time periods, and if we keep only the observations that appear at least nine times in this dataset, this leaves us with 3,027 observations in total from 299 distinct districts. From now on we will refer to this more balanced dataset as the "small panel" and the whole dataset as the " large panel". In the small panel, there are 207 districts located on major aquifers and 91 districts located on minor aquifers. Figure 3.2 and Figure 3.3 present the location of the facilities in the large and small panels.

Table 3.1: Number of observations by year

Year	large panel	small panel
1924	518	170
1930	2515	284
1943	3574	249
1945	17170	277
1954	493	292
1956	486	291
1958	12809	299
1960	554	295
1962	679	285
1963	15191	291
1964	804	294
total number of observations	54793	3027
number of unique facilities	19574	299

Notes: The highly imbalanced results of column 1 is a result of different design of the surveys. In 1945, 1958 and 1963, facilities that serve over 100 people are provided in the survey, while in 1954, 1956, 1960, 1962 and 1964 the survey only covers communities with population greater than 25,000.

Figure 3.4 describes the change of supply sources over the years in the small panel. It shows the percentage of "ground only" sources grew fast from 1924 to 1930, then kept almost constant after 1930. Also, it shows the percentage of "both ground and surface" steadily increased over time. Two possible reasons have contributed to the sharp change from 1924 to 1930 in this plot: 1924 survey criteria is based on treatment plant capacity (one million gallons per day and over), while other surveys set the criteria by population supplied. This results in about 100 fewer observations in 1924 than other years in the small panel, which means many districts that were surveyed in 1924 did not appear enough times (eight more times) in the following surveys. Another reason is the increased demand for groundwater for municipal and irrigation use. By the 1920s, many areas had such urgent demand for water that they faced the danger of overdevelopment. At this time, geology and hydrology investigation made it possible to widely use groundwater (Hornbeck and

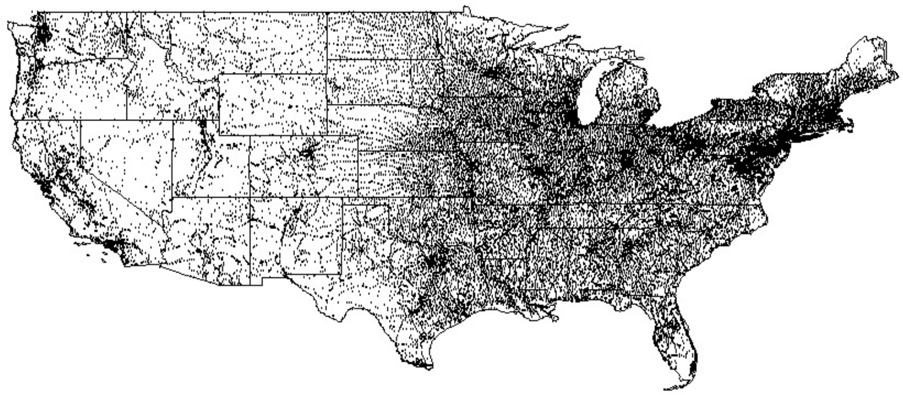


Notes: Light blue areas ar major aquifers, blank areas are minor aquifers. Green dots are observations in the small panel.

Figure 3.2: Map of districts in the small panel

Keskin, 2014; "The 1920's", 2016). Furthermore, with the development of diesel, electric pumps and drilling technology borrowed from oil drilling such as internalized power source and portable tool, groundwater became major water supply in places with limited access to surface water in the following two decades [Lund et al., 2014]. Figure 3.4 shows that districts with only surface supply were adding ground sources and switch into "both ground and surface" category over the years, which provides evidence of increasing need for new supply sources in those years.

Figure 3.5 describes the distribution of supply sources by region. It shows that West region had much higher percentage of both ground and surface sources than the other three regions. In the West region, around 66 percent of "surface only" places have added ground source from 1924 to 1954. This change mostly comes from the state of California: from 1945 to 1954, five out of six of the districts in California with only surface sources have added ground sources.



Notes: Light blue areas are major aquifers, blank areas are minor aquifers. Green dots are observations in the large panel.

Figure 3.3: Map of districts in the large panel

3.4 Model and regression results

This paper employs a fixed effect model to estimate the impact of different drinking water supply sources on population growth. We estimate the following model:

$$y_{it} = \beta_0 + \beta_1 ng_{it} + \beta_2 d_{it} + \beta_3 ng_{it} d_{it} + \alpha_i + \alpha_t + \epsilon_{it} \tag{3.1}$$

where ng_{it} is a binary indicator of drinking water supply source at district i and year t: $ng_{it} = 1$ means there is no groundwater (only surface water) as supply source, $ng_{it} = 0$ means there is groundwater or both ground and surface water as supply source. d_{it} is a drought index created from the Palmer Drought Severity Index. It measures the frequency of drought in the past five years and is denoted "*drought_5*" in regression result tables. $ng_{it} \times d_{it}$ is the interaction term of water supply source and drought frequency, and β_3 is the coefficient of interest that estimates how different supply source affects population growth during drought years. y_{it} is the log of the estimated population supplied in district i during year t, which are from the water dataset. Though the dataset also provides other variables that may represent the prosperity or resilience of a district, such as census population,

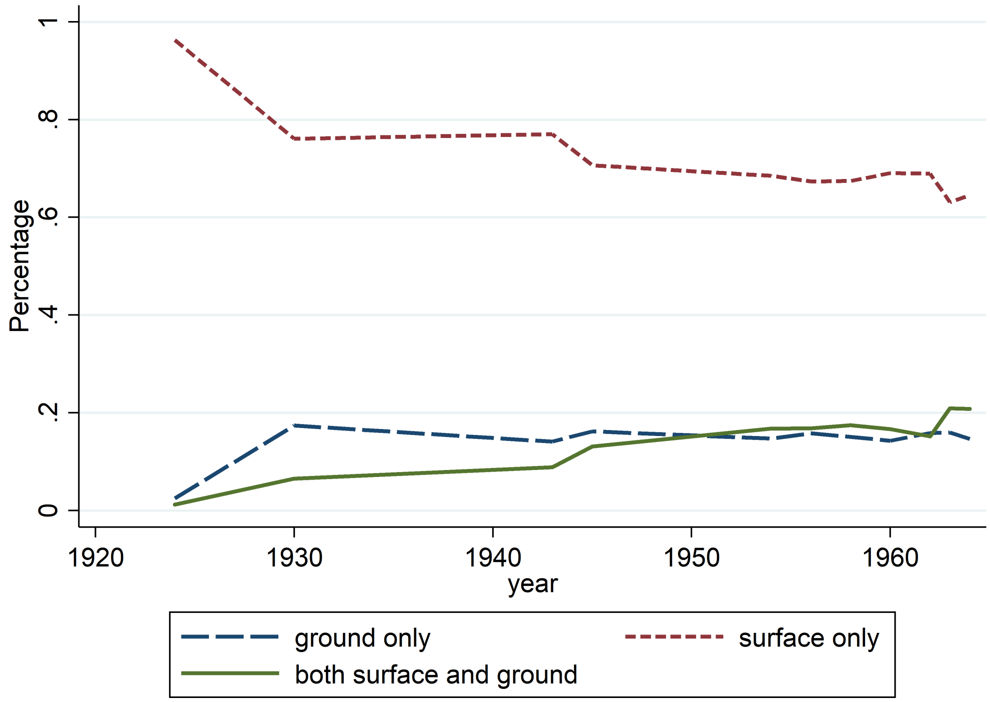


Figure 3.4: Percentage of different sources by year

number of accounts served and number of meters, we choose to use estimated population served because census population is decennial, and number of accounts and number of meters contain a significant amount of missing values. α_i stands for district fixed effect, and α_t stands for year fixed effect. We also run regression using alternative expression as a check of sensitivity, that is, we use "no surface water" as supply source indicator in the equation[1]. Results from this regression are presented in the Appendix.

[1]The alternative specification is

$$y_{it} = \gamma_0 + \gamma_1 ns_{it} + \gamma_2 d_{it} + \gamma_3 ns_{it} d_{it} + \rho_i + \rho_t + \xi_{it}, \tag{3.2}$$

where ns_{it} is a binary indicator for surface water supply: $ns_{it} = 1$ means there is no surface water (only groundwater) as supply source, $ns_{it} = 0$ means there is surface water or both ground and surface water as supply source. Note these two specifications are not symmetric by construction (ng_{it} is not the complement set of ns_{it}). Thus running both specifications allows us to test the sensitivity of our model to different classifications of supply source. γ_3 is the coefficient of interest, ρ_i stands for district fixed effect, and ρ_t stands for year fixed effect.

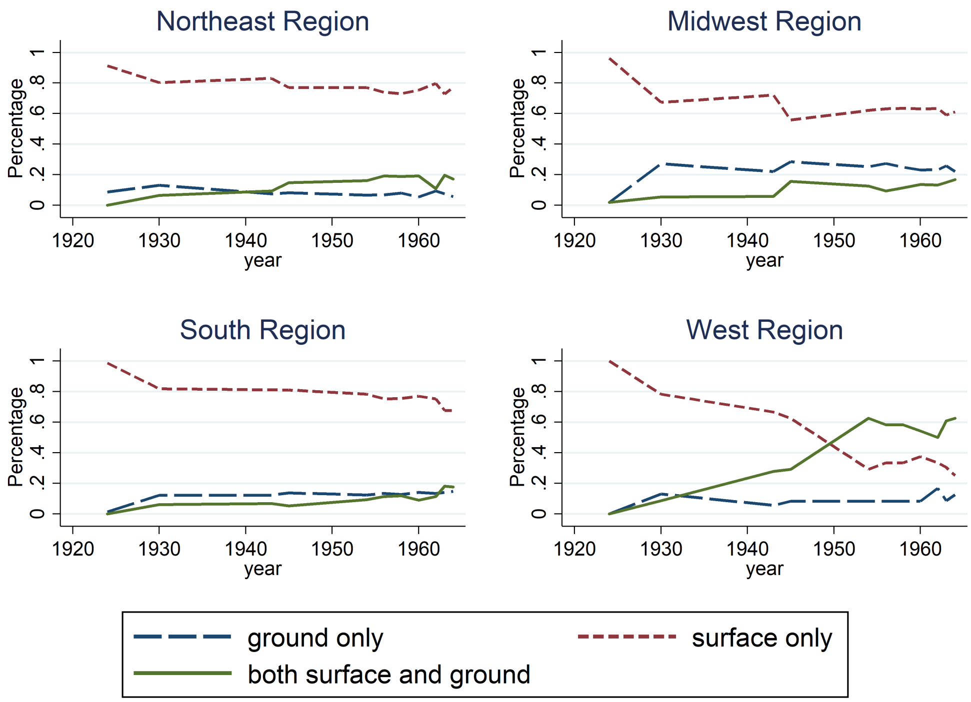


Figure 3.5: Percentage of different supply source by region

3.4.1 Results from the small panel

Table 3.2 contains regression results from the small panel[2]. All observations of supply source and estimated population supplied are at district-year level, and drought index *drought_5* is at climate division-year level. The water facility dataset shows that it is common for districts in major cities to share the same source of supply. Column (1) to column (3) show results using community (district) fixed effect with standard errors clustered at the district level. We also include regressions on subsets of the small panel based on aquifer categories. Columns (2) and (5) show results from observations located on major aquifers, columns (3) and (6) are results from observations located on minor aquifers.

Since the small panel contains only observations that appear at least nine times in the whole dataset and most water facility surveys use population as a threshold, this means the small panel contains mostly large communities. The regression results show that among

[2]See Table B.1 for results estimating equation (2).

Table 3.2: Regression results from small panel

	(1)	(2)	(3)	(4)	(5)	(6)
Small panel: Log Estimated Population Supplied						
ng × drought_5	-0.389***	-0.496***	-0.0773	-0.347***	-0.467***	0.0301
	(0.122)	(0.158)	(0.116)	(0.133)	(0.168)	(0.176)
ng	0.123	0.181	-0.0457	0.219*	0.296*	-0.0754
	(0.108)	(0.152)	(0.0419)	(0.121)	(0.162)	(0.144)
drought_5	0.241***	0.255**	0.133**	0.191**	0.223*	0.0452
	(0.0830)	(0.108)	(0.0526)	(0.0910)	(0.117)	(0.112)
Community FE	Yes	Yes	Yes	No	No	No
County FE	No	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor
Observations	2,732	1,912	820	2,733	1,913	820
Communities	299	208	91	299	208	91
Counties	267	189	83	167	189	83
R-squared	0.858	0.828	0.951	0.750	0.727	0.917

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table represents regression results from estimating equation (1) using data in the small panel, where observations appear at least nine times in the whole dataset. Term $ng \times drought\_5$ is the coefficient of interest, and it represents the effect of "no groundwater" on population during drought.

these relatively large communities, lacking groundwater during droughts has a negative impact on the population. This confirms the assumption that having access to groundwater supply source helps with population resilience during drought.

Column (2) shows that for communities located on major aquifer where groundwater is relatively easy to access, lacking groundwater supply source during drought has a greater negative impact on the population than the general case. On the other hand, the regression results show that for large communities located on minor aquifers, the supply source does not have a statistically significant impact on the population.

3.4.2 Large panel

Table 3.3 presents our results using the large panel[3]. The regression results show that having surface water as the single supply source is correlated with larger population. Column (2) and (5) are regression results using the major aquifer subset, and column (3) and (6) are results using the minor aquifer subset. The $ng \times drought\_5$ term in column (1) is negative and statistically significant, which shows that lacking access to groundwater supply source decreases population during drought, consistent with the regression results from the small panel.

The interaction term of column (6) is positive and statistically significant. This may indicate that county level fixed effect may not have captured every unobserved characteristics, and supports our choice of using community level fixed effect. Not only the choice of drinking water supply source may have impact on population resilience, this choice can also be affected by other variables. The choice of supply source could be a result of continuous drought, and it could also be the result of expanding population.

3.4.3 Regression results by region

The drought maps in Figure 3.1 show that drought severity and population growth rate vary by region. To test for heterogeneity among different regions, we also run regressions by region. It seems our results are consistent but noisy (Table 3.4 and Table 3.5). Results using the West region data and the Northeast region data are very different between the small panel and large panel. It seems that smaller communities in the West region were significantly affected by lacking groundwater during drought, however, larger communities were not affected by water supply source during drought. On the other hand, it seems that in the Northeast region, it is larger communities that were affected the most by the lack of groundwater supply source. The distinction indicates possible heterogeneity in water supply source decisions between large and small communities, and among different regions.

[3]See Table B.1 for results estimating equation (2).

Table 3.3: Regression results from large panel

	(1)	(2)	(3)	(4)	(5)	(6)
Large panel: Log Estimated Population Supplied						
ng×drought_5	-0.220***	-0.305***	-0.0672**	-0.0379	-0.225***	0.209**
	(0.0242)	(0.0347)	(0.0317)	(0.0608)	(0.0843)	(0.0843)
ng	0.0975***	0.112***	0.0597**	0.770***	0.828***	0.745***
	(0.0203)	(0.0278)	(0.0259)	(0.0350)	(0.0457)	(0.0552)
drought_5	0.201***	0.250***	0.0995***	0.00845	0.0694*	-0.0401
	(0.0186)	(0.0243)	(0.0250)	(0.0306)	(0.0368)	(0.0518)
Community FE	Yes	Yes	Yes	No	No	No
County FE	No	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor
Observations	36,094	23,860	12,234	39,408	25,650	13,678
Communities	20,820	12,953	7,867	20,820	12,953	7,867
Counties	3088	2423	1436	3088	2423	1436
R-squared	0.969	0.966	0.977	0.545	0.576	0.571

Robust standard errors in parentheses

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

Notes: This table represents the regression results using data from large panel. Column (2) and (5) are results using the major aquifer subset of the large panel, and column (3) and (6) are results using the minor aquifer subset. The term $ng \times drought\_5$ is the coefficient of interest, which represents the impact of choice of supply source on population during drought.

3.5 Conclusion

In this paper, we take advantage of a unique and comprehensive drinking water infrastructure dataset and study how drinking water supply source affects population resilience during droughts. Using a fixed effect model, we find evidence that having access to groundwater supply helps maintain city population during drought, especially for larger communities and communities located on major aquifers. Our results are robust to different classifications of water supply sources. We run regressions using both balanced subsample and imbalanced (but significantly larger size) subsample; we also run regressions on a national scale and by different region. Our results are mostly consistent but can get noisy with imbalanced subsample.

Table 3.4: Regression results by region, small panel

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable: Log Estimated Population Supplied												
Northeast region							**Midwest region**					
ng×drought_5	-1.673*	-2.651*	0.00896				-0.123	-0.0268	-0.226			
	(0.918)	(1.379)	(0.270)				(0.138)	(0.186)	(0.255)			
ng	0.596	1.004	-0.0159				0.0325	0.0194	0.0249			
	(0.448)	(0.735)	(0.0776)				(0.0491)	(0.0644)	(0.0763)			
drought_5	1.368	2.218*	-0.264	-0.171	-0.243	-0.261	0.0154	-0.0633	0.0797*	-0.0484	-0.0377	-0.0683
	(0.850)	(1.238)	(0.283)	(0.173)	(0.316)	(0.167)	(0.101)	(0.178)	(0.0399)	(0.0884)	(0.0847)	(0.188)
ns×drought_5				1.705	2.078	-0.0852				-0.0271	-0.140	0.106
				(1.762)	(2.982)	(0.358)				(0.152)	(0.224)	(0.185)
ns				-1.063	-1.749	0.136*				-0.0308	-0.00143	-0.0413
				(0.884)	(1.439)	(0.0783)				(0.0447)	(0.0720)	(0.0443)
Observations	654	351	303	654	351	303	875	573	302	875	573	302
R-squared	0.747	0.683	0.950	0.753	0.698	0.950	0.946	0.932	0.978	0.946	0.932	0.978
Community FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor	All	Major	Minor	All	Major	Minor
Region	NE	NE	NE	NE	NE	NE	MW	MW	MW	MW	MW	MW
Panel	Small	Small	Small	Small	Small	Small	Small	Small	Small	Small	Small	Small

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable: Log Estimated Population Supplied												
South region							**West region**					
ng×drought_5	-0.271*	-0.328	-0.180				0.115	0.176	0.100			
	(0.161)	(0.209)	(0.168)				(0.244)	(0.336)	(0.104)			
ng	0.0439	0.0587	0.0353				0.160	0.224	-0.00971			
	(0.136)	(0.176)	(0.0969)				(0.118)	(0.173)	(0.122)			
drought_5	0.0652	0.0598	0.293*	-0.168	-0.219	0.125	0.108	0.147	-0.125	0.0335	0.0107	-0.0904
	(0.129)	(0.153)	(0.164)	(0.133)	(0.178)	(0.234)	(0.128)	(0.165)	(0.0822)	(0.109)	(0.183)	(0.103)
ns×drought_5				0.425**	0.462**					0.555**	0.660**	-0.480
				(0.179)	(0.211)					(0.199)	(0.297)	(0.385)
ns				-0.601	-0.606					-0.278	-0.426	0.119
				(0.762)	(0.764)					(0.256)	(0.383)	(0.0673)
Observations	967	800	167	967	800	167	218	170	47	218	170	47
R-squared	0.832	0.817	0.915	0.835	0.820	0.914	0.965	0.949	0.996	0.965	0.949	0.996
Community FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor	All	Major	Minor	All	Major	Minor
Region	South	South	South	South	South	South	West	West	West	West	West	West
Panel	Small	Small	Small	Small	Small	Small	Small	Small	Small	Small	Small	Small

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Regression results by region, large panel

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable: Log Estimated Population Supplied												
Northeast region							**Midwest region**					
ng×drought_5	0.0553	-0.217	0.396**				-0.0517*	-0.0142	-0.0297			
	(0.144)	(0.195)	(0.161)				(0.0308)	(0.0613)	(0.0377)			
ng	0.0349	0.120	-0.0803**				0.113***	0.0344	0.168***			
	(0.0583)	(0.0949)	(0.0394)				(0.0306)	(0.0353)	(0.0498)			
drought_5	-0.0467	0.172	-0.172	-0.0037	0.0029	0.156	0.0678***	0.121***	0.00938	0.00791	0.139***	-0.0457
	(0.120)	(0.173)	(0.154)	(0.0794)	(0.124)	(0.104)	(0.0238)	(0.0310)	(0.0376)	(0.0328)	(0.0527)	(0.0452)
ns×drought_5				-0.0228	0.155	-0.355				0.0667**	-0.0213	0.0690*
				(0.191)	(0.236)	(0.255)				(0.0314)	(0.0532)	(0.0389)
ns				-0.130	-0.223	0.0240				-0.123***	-0.0359	-0.193***
				(0.0899)	(0.141)	(0.0552)				(0.0330)	(0.0377)	(0.0548)
Observations	6,006	3,205	2,801	6,006	3,205	2,801	11,293	6,123	5,170	11,293	6,123	5,170
R-squared	0.963	0.943	0.982	0.963	0.944	0.982	0.983	0.985	0.982	0.983	0.985	0.982
Community FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor	All	Major	Minor	All	Major	Minor
Region	NE	NE	NE	NE	NE	NE	MW	MW	MW	MW	MW	MW
Panel	Large	Large	Large	Large	Large	Large	Large	Large	Large	Large	Large	Large

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable: Log Estimated Population Supplied												
South region							**West region**					
ng×drought_5	-0.139***	-0.0971**	-0.143***				-0.475***	-0.606***	-0.199*			
	(0.0324)	(0.0449)	(0.0418)				(0.0684)	(0.0818)	(0.116)			
ng	0.0874**	0.0899**	0.0634				0.0549	0.0957**	-0.0473			
	(0.0349)	(0.0409)	(0.0619)				(0.0413)	(0.0479)	(0.0820)			
drought_5	0.0763***	0.126***	-0.0334	-0.0506	0.0388	-0.162**	0.472***	0.510***	0.340***	0.0984**	0.0529	0.189***
	(0.0288)	(0.0348)	(0.0627)	(0.0363)	(0.0481)	(0.0690)	(0.0430)	(0.0512)	(0.0730)	(0.0431)	(0.0526)	(0.0720)
ns×drought_5				0.132***	0.0927**	0.126***				0.420***	0.517***	0.140
				(0.0303)	(0.0419)	(0.0389)				(0.0656)	(0.0794)	(0.108)
ns				-0.215***	-0.221***	-0.172**				-0.236***	-0.246***	-0.230**
				(0.0537)	(0.0662)	(0.0768)				(0.0505)	(0.0576)	(0.108)
Observations	12,533	9,912	2,621	12,533	9,912	2,621	6,262	4,620	1,642	6,262	4,620	1,642
R-squared	0.965	0.963	0.974	0.965	0.963	0.974	0.964	0.965	0.954	0.964	0.965	0.954
Community FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor	All	Major	Minor	All	Major	Minor
Region	South	South	South	South	South	South	West	West	West	West	West	West
Panel	Large	Large	Large	Large	Large	Large	Large	Large	Large	Large	Large	Large

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.6 Bibliography

Baker, L. A., Brazel, A., and Westerhoff, P. (2004). Environmental consequences of rapid urbanization in warm, arid lands: case study of phoenix, arizona (usa). *WIT Transactions on Ecology and the Environment*, 72.

EPA, U. Drinking water infrastructure needs survey and assessment.

Foster, S. (2001). The interdependence of groundwater and urbanisation in rapidly developing cities. *Urban water*, 3(3):185–192.

Gillespie, C. (1925). Filtration plant census, 1924. *Jour. Am. WW Assoc*, 14:123–142.

Hansen, Z. K. and Libecap, G. D. (2003). Small farms, externalities, and the dust bowl of the 1930's. Technical report, National Bureau of Economic Research.

Hornbeck, R. and Keskin, P. (2014). The historically evolving impact of the ogallala aquifer: Agricultural adaptation to groundwater and drought. *American Economic Journal: Applied Economics*, 6(1):190–219.

Lund, J., Munevar, A., Taghavi, A., Hall, M., and Saracino, A. (2014). *Integrating storage in California's changing water system.*

Marshall. Drinking water infrastructure needs of new york state.

Martin (2008). Dust bowl migration. 14(4).

Mosley, L. M. (2015). Drought impacts on the water quality of freshwater systems; review and integration. *Earth-Science Reviews*, 140:203–214.

Nace, R. L., Pluhowski, E. J., et al. (1965). *Drought of the 1950's with special reference to the midcontinent.* US Government Printing Office.

Singh, V. P., Khedun, C. P., and Mishra, A. K. (2014). Water, environment, energy, and population growth: Implications for water sustainability under climate change. *Journal of Hydrologic Engineering*, 19(4):667–673.

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., et al. (2013). Ground water and climate change. *Nature Climate Change*, 3(4):322–329.

USGS (2004). The 1920's.

4. THE EFFECTIVENESS OF PHOSPHORUS LAWN FERTILIZER BANS IN FLORIDA

Phosphorus runoff has been a major environmental issue in the United States for decades. Over-application of phosphorus fertilizer in both farm and non-farm sectors can cause surface water pollution. This paper looks into the effect of a series of rainy season phosphorus lawn fertilizer bans in Florida from 2006 to 2015. These bans require zero phosphorus lawn fertilizer application from June to September in eleven Florida counties. By studying the before-ban and after-ban changes on lawn fertilizer purchases in ban and non-ban counties, we show that these bans result in an average of 21.7% decrease of fertilizer purchases in ban counties.

4.1 Introduction

Phosphorus pollution causes serious water quality problems worldwide. Excess phosphorus in surface waters results in dramatic changes to aquatic ecosystems and leads to the loss of biodiversity. Phosphorus pollution comes from both agricultural and urban activities, such as over-application of phosphorus fertilizer, soil erosion, and urban runoff. High phosphorus levels contribute to harmful algal blooms in U.S. coastal waters and cost an estimated $82 million annually, or $4,162 per kilometer of coastline[1]. This estimation includes losses in public health, commercial fisheries, recreation and tourism, and monitoring and management sectors [Granéli and Turner, 2006].

Florida has been suffering from nutrient over-enrichment in waters for decades. Phosphorus is a key contributor to harmful algae blooms in inland and coastal waters in Florida.

[1] Averaged over the 14-year period from 1987 to 2000, in 2005 dollars

High phosphorus levels in Florida lakes have destroyed sport fish population, and fast-growing algae blooms have led to death of coral reefs and underwater vegetation in coastal area. Poor water quality also endangers human health [Dutzik and Baliga, 2004]. Algae outbreaks have brought real social cost to Florida. In a 2012 report, the total use and non-use economic value of Florida's clean water amounts to $1.3 to $10.5 billion dollars[2] annually[3] [Stanton and Taylor, 2012]. Different policies have been implemented to control phosphorus in Florida over the years, such as building Stormwater Treatment Areas, requiring best management practices, creating surface or groundwater storage for seasonal water surpluses and constructing water quality credit markets.

This paper focuses on a policy that aims at reducing phosphorus from home lawn fertilizer application during rainy season. Beautiful green lawns are highly valued by people. However, the social cost of maintaining green lawns exceeds the private cost. In addition to Florida, eleven states[4] have enacted laws that restrict applying phosphorus fertilizer to lawn and turf since 2002 [Miller, 2012]. Many Florida soils are naturally high in phosphorus, and rainy season means heavier runoff which contributes to transporting excess phosphorus to local waterways. The ban was first initiated in Sarasota County, Florida in 2007, and ten more coastal counties in Florida adopted this ban from 2008 to 2015.

In this paper, we use a restricted weekly scanner dataset at store level from Nielsen, obtained through the Kilts Center for Marketing and the University of Chicago. It exploits spatial and temporal variations in Florida fertilizer ordinances and examines the effectiveness of the bans by studying how fertilizer consumers in ban counties responded to this rainy season ban. We find evidence that ban counties experienced a drop in fertilizer purchases as a result of the ban.

[2] In 2010 dollars

[3] The authors calculated potential non-use value for improvements to Florida's water quality to be $448 million (using average EPA 2010 willingness-to-pay (see U.S. Environmental Protection Agency (2010) Section 13.2 for details)) and $3.5 billion (using maximum willingness-to-pay). Total use and non-use value ($1.3 to $10.5 billion) is calculated as three times the above non-use value ($448 million to $3.5 billion). See Page 24 of Stanton and Taylor (2012) for details.

[4] Namely Illinois, Maine, Maryland, Michigan, Minnesota, New Jersey, New York, Vermont, Virginia, Washington, Wisconsin and Florida.

This paper contributes to the existing literature on incomplete environmental regulations. Previous work on the effectiveness of similar urban phosphorus fertilizer regulations often analyzes from an environmental science perspective, which means the effectiveness of such regulations is measured by changes in post-regulation phosphorus concentrations in surface water and soil[5] [Vlach et al., 2010, Lehman et al., 2009, MDA, 2007, Hochmuth et al., 2012]. Alternatively, we examine the effectiveness of phosphorus lawn fertilizer bans from an economic perspective. We look into how this series of bans affect consumer behavior, and we examine their effectiveness by testing whether there is spillover in sales. We provide evidence that fertilizer sales in ban counties during summer decreased after the bans were implemented, and there is strong spillover effect on fertilizer sales to before-ban months. In addition, we show that the series of incomplete regulations did not result in spillover effect on fertilizer sales between ban counties and non-ban, border counties, which agrees with our assumption of a ban on use.

Also, much of existing environmental economics literature on incomplete regulations focus on air policies, and our work contributes to the literature by studying a water policy. Our paper also differs from previous work on incomplete regulations of household phosphorus product such as Cohen and Keiser (2017) by studying a different good[6]. The two studies are also different in that this paper studies a use ban which requires residents to follow the regulations voluntarily , compared with Cohen and Keiser (2017) which studies a

[5] For example, Vlach et al. (2010) conducted a paired watershed study between three sites in Plymouth, Minnesota where phosphorus lawn fertilizer was restricted, and three sites in Maple Grove, Minnesota where phosphorus lawn fertilizer was allowed. This study found evidence that phosphorus fertilizer restriction leads to decreased phosphorus level in the surface water. Lehman et al. (2009) analyzed changes in total and dissolved phosphorus concentration in Huron River following a phosphorus fertilizer restriction in the city of Ann Arbor. They found evidence of reduced phosphorus concentration but suggested the reduction in phosphorus was also subject to overlapping policies. Other studies that focus on phosphorus runoff (including turf-grass runoff and watershed runoff) after Minnesota enacted a statewide law to restrict the application of phosphorus lawn fertilizer have reached inconclusive results on the efficacy of this regulation [MDA, 2007]. Hochmuth et al. (2012) provides a scientific review on phosphorus lawn fertilizer regulations and discusses potential unintended consequences of the Florida rainy season phosphorus bans from an environmental science perspective. They focus on the potential unintended changes of phosphorus concentration as a result of the complex interaction between landscape management, plant species, nutrient uptake, rainfall, irrigation management and overlapping regulations.

[6] Cohen and Keiser (2017) studies the unintended consequences of incomplete and overlapping regulations by looking into bans on phosphate in dishwasher detergent in Spokane, Washington.

sales ban that directly put restrictions on the supply side. Our work also contributes to the environmental economics literature by taking advantage of a series of bans implemented in different counties, different years but in the same state. This allows us to isolate the average effect of the series of bans and minimize possible impacts from overlapping policies, compared to previous studies that focus on a one-time statewide or municipal regulation.

The rest of the paper proceeds as follows: section 2 introduces ban details, section 3 describes the dataset used in this paper, section 4 presents the empirical model, section 5 discusses the preliminary regression results, and the final section concludes.

4.2 Ban details

The rainy season phosphorus ban in Florida was first implemented in Sarasota County in 2007. Rainy season means June 1st to September 30th, and this ban is also referred to as a "summer ban". This ban requires that no phosphorus fertilizer shall be applied to lawn (turf) during rainy season. It aims at household use of fertilizer, which means it does not ban sales of fertilizer (except in Pinellas County) and it only affects non-farm sector. This ban is not the only restriction on the home use of fertilizer in Florida. Before the series of rainy season ban there already existed other restrictions on lawn and turf fertilization, such as restrictions on total nutrient content per application or per year and restrictions on fertilizer application under storm/flood/hurricane warning.

One example of a related regulation on fertilizer application is a restriction on the total content of nitrogen and phosphorus per application and all year round. For example, Charlotte County requires that phosphorus content per application cannot exceed 0.25 pounds P_2O_5 / 1000 square feet and per year application cannot exceed 0.50 pounds P_2O_5 / 1000 square feet [Charlotte County, 2008]. These restrictions are typical among all Florida counties. These related regulations are mostly according to Florida Green Industry Best Management Practice (GI-BMP) manual, which was first established in 2002 [FDEP, 2010]. These rules also have exemptions for golf courses, public recreational & stadium fields, bona fide farm operations, newly established turf and/or landscape plants (first 60 days after installation or planting), damaged turf and/or landscape plants (60 days) and damaged turf and/or landscape plants (during rainy season, with proper document). Other than rainy season phosphorus ban, Pinellas County, Manatee County and St. Lucie County also ban phosphorus lawn fertilizer all year round, except for where phosphorus deficiency has been demonstrated by a soil analysis test performed by a State of Florida certified lab.

Only Pinellas County bans both the use and sale of phosphorus lawn fertilizer. All other ban counties only ban the use of phosphorus lawn fertilizer. This means the rainy season ban on phosphorus lawn fertilizer requires households to follow the rules voluntarily, and there are no requirements of monitoring or testing exactly what fertilizer is applied to lawn and turf. If a person is found to violate the ban, the government will issue a warning for the first time. One will not be fined unless he/she violates the ban multiple times, as a Cocoa Beach code enforcement officer commented: "...you want to educate them before they ever do get fined..." (Florida Today). To help residents pick the right fertilizer that complies with the ban, Florida also requires all fertilizer producers to label the content of nutrient on fertilizer packages clearly. Usually, there are three numbers that represent the primary nutrients: nitrogen(N)-phosphorus(P)-potassium(K). A string of 10-10-10 means this bag of fertilizer contains 10% of N, P, K each. Thus, consumers in ban counties need to choose from fertilizers that are labeled "0" in the middle if they need to fertilize home lawns during the rainy season.

After Sarasota County adopted rainy season ban countywide in 2007, Lee County, Orange County, Manatee County, Pinellas County, Charlotte County, Brevard County, Indian River County, Volusia County, St. Lucie County and Martin County have gradually adopted this ban from 2008 to 2015. Table 4.1 and Figure 4.1 show the temporal and spatial variation of counties that adopted this ban. Figure 4.1 also shows that all ban counties are along the coastline. Ban counties are filled with light to dark blues indicating early to late implementation year. Shaded area means the county bans phosphorus all year round instead of only rainy season blackout.

Table 4.1: Ban Counties and Details

County	Ban Start Year	Ban Type	Ban Period	Number of Stores	Number of Observations
Sarasota	2007	Ban on Use	Rainy Season Ban	42	1,774
Lee	2008	Ban on Use	Rainy Season Ban	73	3,100
Charlotte	2008	Ban on Use	Rainy Season Ban	16	694
Orange	2010	Ban on Use	Rainy Season Ban	129	5,536
Pinellas	2010	Ban on Sale	All Year Ban	128	4,571
Manatee	2011	Ban on Use	All Year Ban	37	1,715
Brevard	2014	Ban on Use	Rainy Season Ban	65	2,896
Indian River	2014	Ban on Use	Rainy Season Ban	15	827
Volusia	2014	Ban on Use	Rainy Season Ban	74	3,274
St. Lucie	2014	Ban on Use	All Year Ban	30	1,062
Martin	2015	Ban on Use	Rainy Season Ban	16	687
Total				625	26,136

4.3 Data

The main dataset used in this paper is a restricted store-level scanner dataset[7] obtained through the Kilts Center for Marketing and the University of Chicago[8]. This dataset provides information on weekly store-level purchases of all brands of lawn fertilizers and other products. It provides detailed information such as brands, UPC, and description of content for each entry. The detailed information allows us to separate lawn fertilizers from other lawn product such as flower food, pesticide, herbicide and more. We aggregate weekly sales revenue, which is unit × price, to monthly store-level sales of total lawn fertilizer. We include all stores in Florida dataset, and we match stores to each county using county FIPS code provided in this dataset. The final data used in the following analysis are store-month

[7]Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

[8]The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

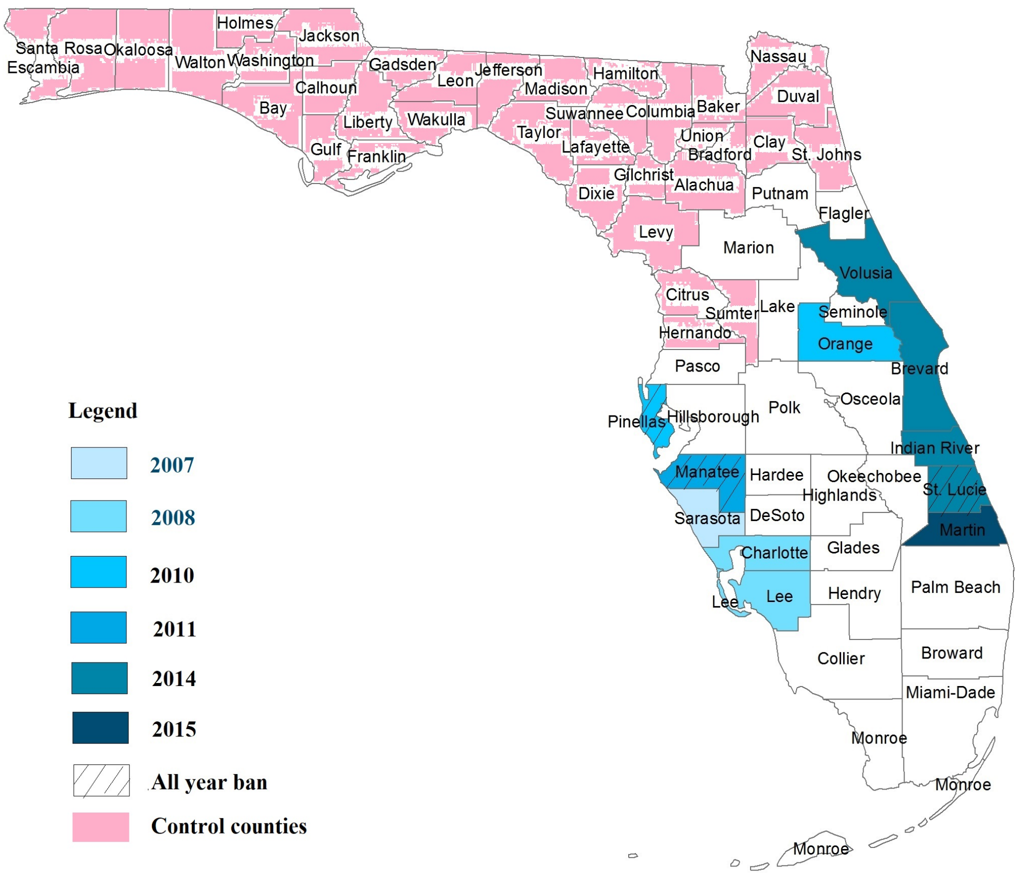


Figure 4.1: Ban Counties in Florida

level, and the final sample contains data of 290 different lawn fertilizers from 2,163 stores in 67 counties, spanning from 2006 to 2015 with a total of 92,755 observations. Table 4.2 shows summary statistics of ban counties in the final sample. It shows that there are much more stores that sell lawn fertilizers among counties in the south Florida (where ban counties and non-ban border counties are) than in the central and north Florida.

Table 4.2: Summary statistics of final sample

County Type	Number of Counties	Number of Stores	Number of Observations
Ban County	11	625	26,136
Non-ban Non-adjecent County	35	476	23,322
Non-ban Adjecent County	17	674	28,871
Total	63	1,775	78,329

We also use labor force data from Bureau of Labor Statistics and Economic data from Bureau of Economic Analysis to bring in more control variables that describe the economic and labor force status. These data are county-year level. We add these control variables (income, unemployment rate, and population) to test the sensitivity of our results to the recession during our sample period.

4.4 Model

To study how consumers responded to the rainy season ban, we estimate the average effect of this series of bans on consumer behavior in ban counties. We use a fixed model to control for time-invariant and space-invariant unobservables, such as store characteristics, potential different preferences in home use fertilizer in different counties and seasonality in fertilizer purchases,

The fixed effects model we begin with is:

$$ln(sales_{icm}) = \beta_0 + \beta_1 post_{icm} * ban\_month_{icm} + \beta_2 post_{icm} * nonban\_month_{icm} + \alpha_i + \alpha_m + \epsilon_{icm}, \tag{4.1}$$

where $sales_{icm}$ stands for the aggregated fertilizer sales at store i, county c, year-month m. β_1 and β_2 are our coefficient of interest, which represent the aggregate effect of the ban on months the bans apply, and spill over effect to non-ban months. $post_{icm}$ is a binary indicator that equals one for store i in county c in and after the June this phosphorus ban was first implemented in this county, and zero otherwise. ban_month_{icm} ($nonban\_month_{icm}$) is a binary indicator that equals one (zero) for store i in county c in year-month m that the phosphorus ban is applied, and equals zero (one) for all months this ban does not apply in a county. α_i and α_m are store and year by month fixed effects respectively, and ϵ_{icm} is the error term.

To explore potential leakage of sales to months immediately proceeding or following ban months separately, we estimate a slightly different specification:

$$\begin{aligned} ln(sales_{icm}) = \gamma_0 + \gamma_1 post_{icm} * spring_m + \gamma_2 post_{icm} * ban\_month_{icm} \\ + \gamma_3 post_{icm} * fallwinter_m + \eta_i + \eta_m + \epsilon_{icm}, \end{aligned} \tag{4.2}$$

where $spring_m$ equals one if month equals to March, April, May, and $spring_m$ equals zero otherwise. $fallwinter_m$ equals one if month equals to October, November, December, January, February, and $fallwinter_m$ equals zero otherwise. γ_1 and γ_3 are coefficients of interest. γ_1 and γ_3 provide more information of potential spillover effect of the ban than β_1 and β_2 from equation (1).

The following equation is used to plot the average effect of this phosphorus ban by year, using the year before the ban as reference year:

$$ln(sales_{icm}) = \rho_0 + \Sigma_{n=2}^{N}(\rho_n * Year_{c,T-n}) + \Sigma_{m=0}^{M}(\rho_m * Year_{c,T+m})_{+}\omega_i + \omega_m + \epsilon_{icm} \tag{4.3}$$

where $Year_{c,T-n}$ equals one for county c in n year before the start-year of the ban, and T stands for the start-year of the ban. $Year_{c,T-n}$ equals zero for county c in all other years. $Year_{c,T+m}$ is the same for county c in years after the start-year. The year right before the ban $Year_{c,T-1}$ is the reference year, so it does not appear in this equation.

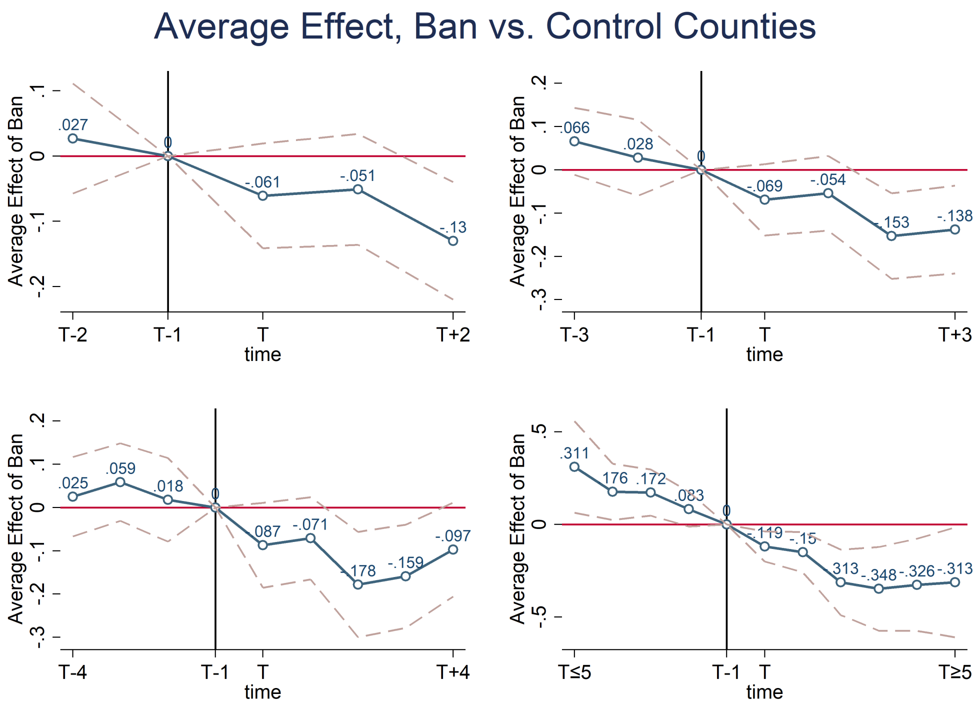
We estimate equation (3) to illustrate the differential sales between ban counties and non-ban, non-border counties. We show that ban counties and non-ban, non-border counties share non-differential sales pattern before the ban, but sales in ban counties started to fall in about two years after the ban and stayed at a lower level. Non-differential pre-trend provides evidence that our identification is valid in that the treatment group and control group satisfy the parallel trend assumption prior to the ban (after controlling for different fixed effect). This allows us to attribute the difference in post-ban sales between two groups to the implementation of this summer ban.

4.5 Results

Table 4.3 presents regression results from equation (1). It shows that on average, fertilizer purchases drop by around 21.68% in ban counties. About a third of this decrease is from sales in ban months, and two thirds are from sales in non-ban months. For robustness check, we also control for income, unemployment rate, population temperature and precipitation, and our results are consistent. Column (8) shows there is no significant change in fertilizer sales in non-ban, border counties compared with control counties. This indicates no spatial spillover of this ban, which agrees with our assumption. We assume that since this ban only applies to fertilizer use instead of sale, consumers in ban counties are likely to (1) reduce fertilizer purchase because there is no need to stock up fertilizer; (2) have no incentive to buy phosphorus fertilizer in non-ban counties even when local supply is low, since they cannot apply it as a result of the use ban.

Table 4.4 presents regression results from equation (2). In equation (2), non-ban months are split into before-ban season and post-ban seasons. Column (4) shows that the bans are most effective in reducing fertilizer sales during spring, or before-ban season. This is reasonable as spring is growing season, and most people fertilize lawns relatively often during spring to keep grass green and healthy. The strong seasonality in sales data confirms this behavior. Fertilizer sales rise sharply every spring and gradually drop through the rest of a year. Column (8) shows fertilizer sales in non-ban, border counties experienced statistically significant changes in different seasons. This is unexpected and against the results from estimating equation (1), and we do not yet have the explanation of this inconsistent result.

Figure 4.2 and Figure 4.3 present plots of the average effect of this ban in different time length before and after the implementation of summer ban, using different subsets of the final sample. We use the last year before the ban as the reference year. Figure 4.2 shows on average, this ban has reduced fertilizer purchases in ban counties by about 17% on average in the following years after implementation compared with control counties, which is consistent with the regression results in Table 4.3.



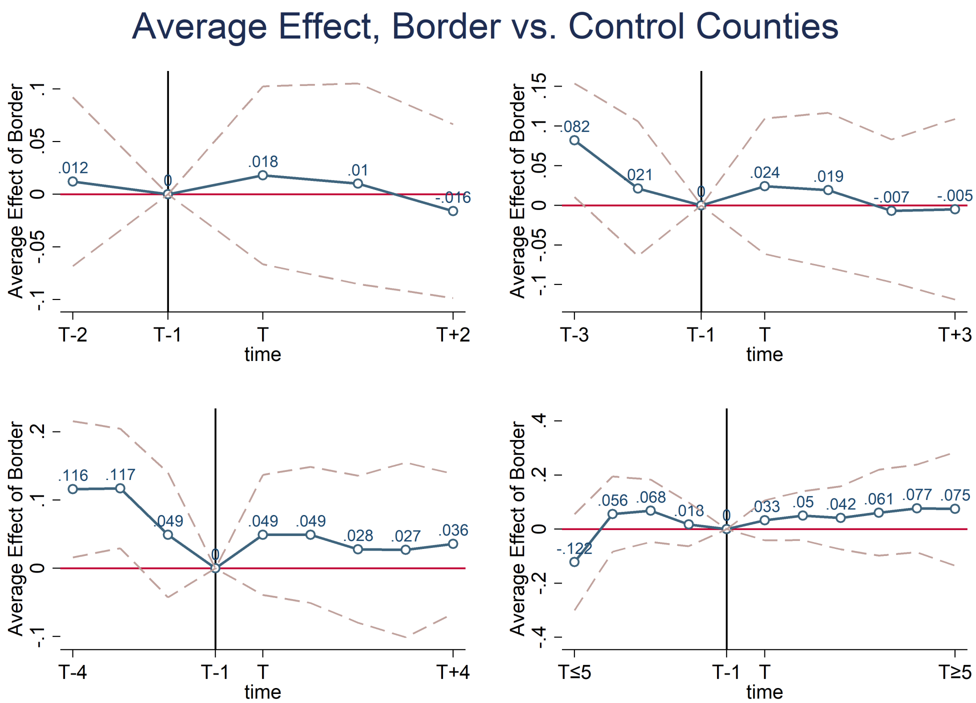
Notes: This figure presents regression results from equation (3), using observations in ban counties and non-ban, non-border counties. T-1 is the last year before the implementation of this ban, and we use this year as reference year.

Figure 4.2: Average effect in ban counties

4.6 Conclusions and discussion

This paper examines the effectiveness of a series of rainy season bans on phosphorus lawn fertilizer application in Florida. We use fixed effects model to study the average effect of the ban. We find evidence that the ban effectively reduced home lawn fertilizer purchases in ban counties, and our results are robust and consistent. We also find spillover effect of this ban on fertilizer sales to before-ban season among ban counties. We find weak evidence that there is no spatial spillover of this ban, that is, we show this ban did not change total fertilizer sales in non-ban, border counties.

Due to the limitedness of data, we cannot tell the exact phosphorus contents from fertilizer sales data. Also, we can only observe the purchases of lawn fertilizer, instead of



Notes: This figure presents regression results from equation (3), using observations in non-ban, border counties and non-ban, non-border counties. We use the earliest ban start year in a county's bordered ban counties as a "start year" for non-ban, border counties. We also use the last year before the "start year" as reference year.

Figure 4.3: Average effect in border counties

the actual application time and amount of fertilizer. This means we cannot provide strong evidence on exactly how much phosphorus is applied to home lawns, and whether there exist changes in actual application due to the ban. All we can observe is that this ban did result in a negative and statistically significant change in fertilizer purchasing pattern. For a ban on use instead of sale, this change in consumer behavior shows the ban is effective and results in little spatial spillover.

Table 4.3: Average effect of the ban

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ban VS. Control				Border VS. Control			
post×ban month	-0.0724	-0.178***	-0.179***	-0.0824*	-0.0206	-0.113***	-0.117***	-0.0557
	(0.0481)	(0.0489)	(0.0488)	(0.0497)	(0.0374)	(0.0414)	(0.0418)	(0.0419)
post×non-ban month	0.0323	-0.0723	-0.0742	-0.162***	0.0812*	-0.00678	-0.00923	-0.0644
	(0.0485)	(0.0460)	(0.0462)	(0.0459)	(0.0452)	(0.0457)	(0.0455)	(0.0455)
ln(income)			0.321	0.304			0.229	0.155
			(0.381)	(0.382)			(0.333)	(0.331)
ln(population)			0.493	0.621			-0.201	-0.103
			(0.802)	(0.802)			(0.543)	(0.538)
ln(unemployment rate)			-0.463***	-0.224			-0.384***	-0.0694
			(0.161)	(0.159)			(0.144)	(0.146)
$\overline{temperature}$				0.227***				0.209***
				(0.0154)				(0.0159)
$\overline{temperature}^2$				-0.00363***				-0.00371***
				(0.000421)				(0.000429)
$\overline{precipitation}$				-0.000590***				-0.000950***
				(0.000188)				(0.000177)
$\overline{precipitation}^2$				1.20e-06***				1.74e-06***
				(3.70e-07)				(3.63e-07)
Observations	49,436	49,436	49,436	49,436	52,173	52,173	52,173	52,173
R-squared	0.802	0.803	0.803	0.807	0.803	0.803	0.804	0.806
Store FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ban County×year FE	No	Yes	Yes	Yes	-	-	-	-
Border County×year	-	-	-	-	No	Yes	Yes	Yes
Cluster	Store	Store	Store	Store	Store	Store	Store	Store
Ban Period	Both	Both	Both	Both	Both	Both	Both	Both

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table shows the regression results from equation (1). Column (1) to (4) are regression results from the subsample of ban counties and control counties (non-ban, non-border counties). Column (5) to (8) are regression results from the subsample of non-ban, border counties and control counties. "post" is a binary indicator that equals one for each ban county after the first month in the first year this phosphorus ban was implemented; otherwise equals zero. For any border county, "post" follows the earliest ban county of all its bordered ban counties. "Both" in "Ban period" means the subsample contains both counties that enacted summer phosphorus ban and counties that enacted all year phosphorus ban. Data used in this regression is store-month level. Column (1) and (5) shows results without ban (border) county by year fixed result. The rest columns show results after adding ban (border) county by year fixed effect, ie, assuming fertilizer purchases in ban (border) counties and control counties present different trend. We also add other controls of economics and labor force, and controls of temperature and precipitation. Column (4) and (8) shows that compared with non-ban, non-border counties, fertilizer purchase in ban counties dropped after the ban, while fertilizer purchase in non-ban border counties did not change. It also shows in ban counties, fertilizer purchase in non-ban months dropped about twice as much as in ban months. This table shows fertilizer purchase in ban counties after the implementation of the ban dropped by about 21.7% on average.

Table 4.4: Average effect of the ban by season, summer ban only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Ban VS. Control				Border VS. Control			
post×ban month	-0.0938**	-0.209***	-0.187***	-0.0221	0.113**	0.00236	0.00746	-0.0821*
	(0.0456)	(0.0462)	(0.0465)	(0.0464)	(0.0451)	(0.0455)	(0.0457)	(0.0454)
post×spring	-0.136***	-0.268***	-0.252***	-0.234***	0.197***	0.0977**	0.111**	0.110**
	(0.0475)	(0.0519)	(0.0515)	(0.0511)	(0.0466)	(0.0457)	(0.0458)	(0.0454)
post×fallwinter	0.212***	0.0971*	0.111*	-0.0559	-0.221***	-0.334***	-0.319***	-0.150***
	(0.0653)	(0.0586)	(0.0583)	(0.0586)	(0.0478)	(0.0535)	(0.0532)	(0.0553)
ln(income)			0.460	0.437			0.282	0.266
			(0.419)	(0.417)			(0.363)	(0.363)
ln(population)			-0.137	-0.134			0.188	0.211
			(0.855)	(0.852)			(0.728)	(0.726)
ln(unemployment rate)			-0.551***	-0.389**			-0.457***	-0.170
			(0.158)	(0.157)			(0.150)	(0.152)
$\overline{temperature}$				0.199***				0.184***
				(0.0173)				(0.0177)
$\overline{temperature}^2$				-0.00298***				-0.00313***
				(0.000480)				(0.000472)
$\overline{precipitation}$				-0.000795***				-0.00102***
				(0.000198)				(0.000191)
$\overline{precipitation}^2$				1.55e-06***				1.73e-06***
				(3.81e-07)				(3.81e-07)
Observations	42,091	42,091	42,091	42,091	43,207	43,207	43,207	43,207
R-squared	0.803	0.804	0.804	0.807	0.802	0.803	0.803	0.805
Store FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ban County×year FE	No	Yes	Yes	Yes	-	-	-	-
Border County×year	-	-	-	-	No	Yes	Yes	Yes
Cluster	Store	Store	Store	Store	Store	Store	Store	Store
Ban Period	Summer	Summer	Summer	Summer	Summer	Summer	Summer	Summer

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table shows results from equation (2) using two different subsamples of our data. "spring" equals one for March, April and May, and "fallwinter" equals one for October, November, December, January and February. This table shows the ban effectively decreased fertilizer purchase in the months preceding ban months in ban counties.

4.7 Bibliography

Charlotte County (2008). Fertilizer ordinance.

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

Dutzik, T. and Baliga, N. (2004). Phosphorus pollution in floridas waters: The need for aggressive action to protect floridas rivers and streams from nutrient runoff. https://frontiergroup.org/sites/default/files/reports/FL-Phosphorus-Pollution-in-Floridas-Waters-text.pdf.

FDEP (2010). Florida friendly best management practices for protection of water resources by the green industries. http://turfpath.ifas.ufl.edu/pdf/BMPmanual.pdf.

Granéli, E. and Turner, J. T. (2006). *Ecology of harmful algae*, volume 189.

Hochmuth, G., Nell, T., Unruh, J. B., Trenholm, L., and Sartain, J. (2012). Potential unintended consequences associated with urban fertilizer bans in floridaa scientific review. *HortTechnology*, 22(5):600–616.

Lehman, J. T., Bell, D. W., and McDonald, K. E. (2009). Reduced river phosphorus following implementation of a lawn fertilizer ordinance. *Lake and Reservoir Management*, 25(3):307–312.

MDA (2007). Report to the minnesota state legislature: Effectiveness of the minnesota phosphorus lawn fertilizer law. https://www.mda.state.mn.us/ /media/Files/protecting/waterprotection/07phoslawreport.pdf.

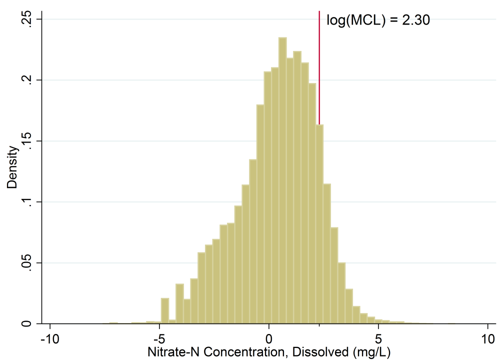
Miller, K. (2012). State laws banning phosphorus fertilizer use.

Stanton, E. Q. and Taylor, M. (2012). Valuing floridas clean waters. *Stockholm Environment Institute - U.S. Center* http://earthjustice.org/sites/default/files/ValuingFloridasCleanWaters.pdf.

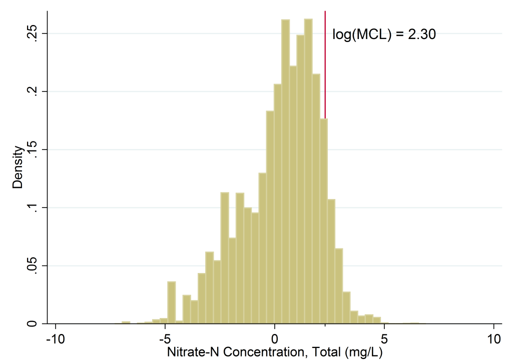
U.S. Environmental Protection Agency (2010). Economic analysis of final water quality standards for nutrients for lakes and flowing waters in florida. https://www.epa.gov/sites/production/files/2015-07/documents/florida-economic-anaylis-report.pdf.

Vlach, B., Barten, J., Johnson, J., and Zachay, M. (2010). Stormwater treatment: Assessment and maintenance case study# 9: Assessment of source reduction due to phosphorus-free fertilizers.

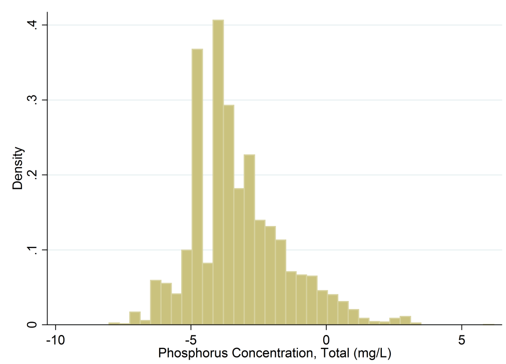
APPENDIX A. ADDITIONAL MATERIAL FOR CHAPTER 1



(a)



(b)



(c)

Notes: This figure shows distributions of nutrient concentrations with no replacement for "below quantification limit" values.

Figure A.1: Distribution of log of nutrient concentrations (no replacement)

Table A.1: Regression results from fixed effect model

Dep. VARIABLES	(1) Ammonia and ammonium Dissolved	(2) Ammonia and ammonium Total	(3) Inorganic nitrogen Dissolved	(4) Inorganic nitrogen Total	(5) Kjeldahl nitrogen Dissolved	(6) Kjeldahl nitrogen Total	(7) Nitrate Dissolved	(8) Nitrate Total	(9) Nitrite Dissolved
CRP ratio	0.152	-3.431	0.324	7.382	1.459	-3.264*	1.213	-8.950*	-0.932*
	(0.982)	(5.299)	(1.293)	(8.118)	(1.505)	(1.745)	(1.902)	(5.084)	(0.518)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County
Obs.	86,349	12,814	106,412	14,903	35,876	11,780	82,613	10,416	80,023
R-squared	0.363	0.608	0.392	0.497	0.406	0.465	0.378	0.518	0.410

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Dep. VARIABLES	(10) Nitrite Total	(11) Nitrogen mixed forms Dissolved	(12) Nitrogen mixed forms Total	(13) Organic nitrogen Dissolved	(14) Organic nitrogen Total	(15) Phosphate Dissolved	(16) Phosphate Total	(17) Phosphorus Dissolved	(18) Phosphorus Total	(19) Phosphate-phosphorus Total
CRP ratio	-3.724	-0.717	1.008	0.619	-5.498**	-1.333	-30.99**	0.00163	5.524***	-31.03***
	(4.097)	(1.277)	(2.375)	(0.826)	(2.192)	(1.100)	(13.44)	(1.706)	(2.136)	(6.658)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County	County
Obs.	8,527	44,851	10,086	41,202	11,620	81,614	5,078	34,893	17,658	5,848
R-squared	0.557	0.325	0.380	0.370	0.471	0.381	0.541	0.465	0.499	0.548

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table contains regression results from estimating equation (1). Independent variable "CRP ratio" stands for five-year moving average of $\frac{CRP\ acreage}{(CRP+cropland)\ acreage}$.

Table A.2: Regression results after adding more land use variable

Dep. VARIABLES	(1) Ammonia and ammonium Dissolved	(2) Ammonia and ammonium Total	(3) Inorganic nitrogen Dissolved	(4) Inorganic nitrogen Total	(5) Kjeldahl nitrogen Dissolved	(6) Kjeldahl nitrogen Total	(7) Nitrate Dissolved	(8) Nitrate Total	(9) Nitrite Dissolved
R_{CRP}	2.340	1.866	1.850	-4.064	6.993***	-10.12	2.869	-42.43	-2.410
	(2.320)	(16.81)	(2.928)	(23.15)	(2.276)	(8.046)	(6.176)	(39.64)	(1.763)
$R_{noncrop}$	-2.190	-29.40*	6.936	-29.09**	-7.543***	3.316	6.913	0.144	1.679
	(2.029)	(16.64)	(4.536)	(14.40)	(2.328)	(8.870)	(5.432)	(8.192)	(1.961)
$R_{rangeland}$	5.142	-8.616	1.075	-18.78*	13.12***	2.119	1.714	-10.37	0.280
	(3.304)	(12.01)	(3.115)	(10.46)	(4.092)	(11.02)	(3.595)	(7.663)	(2.746)
$R_{pastureland}$	-1.308	-0.367	5.603*	-12.48	-6.047*	-7.464	1.172	10.96	-2.010
	(2.535)	(9.829)	(3.339)	(11.61)	(3.308)	(4.601)	(3.466)	(12.68)	(2.371)
R_{urban}	-0.246	-1.183	0.633	-5.662	-5.515***	-5.453	-0.152	4.932	2.972**
	(1.931)	(6.129)	(1.972)	(7.662)	(1.949)	(4.721)	(2.229)	(8.424)	(1.421)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County
Ob.	87,965	12,995	108,064	14,950	36,347	11,830	84,312	10,593	81,749
R-squared	0.361	0.607	0.396	0.499	0.412	0.468	0.380	0.521	0.424

Dep. VARIABLES	(10) Nitrite Total	(11) Nitrogen mixed forms Dissolved	(12) Nitrogen mixed forms Total	(13) Organic nitrogen Dissolved	(14) Organic nitrogen Total	(15) Phosphate Dissolved	(16) Phosphate Total	(17) Phosphorus Dissolved	(18) Phosphorus Total	(19) Phosphate-phosphorus Total
R_{CRP}	-11.56	-1.049	7.592	1.819	-16.74***	-2.244	-94.24	0.791	28.26***	-44.95***
	(10.91)	(2.308)	(7.710)	(2.022)	(6.275)	(2.654)	(67.28)	(4.380)	(7.440)	(13.80)
$R_{noncrop}$	-57.16**	-2.160	3.257	-1.561	-0.0933	-2.476	-17.23	-15.04***	11.51*	7.594
	(25.33)	(2.487)	(7.959)	(1.982)	(4.923)	(2.507)	(16.04)	(4.568)	(6.563)	(13.95)
$R_{rangeland}$	12.82*	2.194	19.24***	13.18***	11.86*	4.863**	34.64*	5.068	26.73***	33.34**
	(7.491)	(2.693)	(7.224)	(2.865)	(6.874)	(2.134)	(20.15)	(5.865)	(7.178)	(15.43)
$R_{pastureland}$	-19.54*	4.291*	-0.705	1.914	-1.979	2.175	6.457	-4.015	11.88*	4.786
	(11.62)	(2.359)	(2.746)	(2.065)	(4.161)	(2.383)	(11.93)	(4.495)	(6.056)	(7.587)
R_{urban}	-16.11*	-2.101	4.624	-4.507***	-4.567	2.073	10.22	-0.0168	10.26**	9.486
	(9.289)	(2.056)	(3.769)	(1.661)	(3.437)	(1.432)	(12.59)	(3.091)	(4.605)	(10.88)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County	County
Obs.	8,704	45,680	10,159	41,996	11,827	83,160	5,241	35,447	17,761	6,011
R-squared	0.569	0.330	0.384	0.375	0.475	0.382	0.538	0.474	0.502	0.543

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Independent variable $R_{(l)}$ is the moving average of $\frac{acreage\ in\ land\ use\ l}{Total\ land\ acreage}$ in the past five years.

Table A.3: Regression results, corn belt

DEPENDENT VARIABLES	Kjeldahl nitrogen Total		Nitrate Total		Phosphate Total		Phosphate-phosphorus Total	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CRP ratio	-0.394		38.65*		216.9		-27.79***	
	(1.522)		(22.58)		(302.8)		(6.416)	
R_{CRP}		-7.682		133.4*		-4,878		-34.94***
		(8.958)		(65.76)		(3,121)		(9.389)
$R_{noncrop}$		18.94		99.12***		194.9***		13.93
		(22.66)		(26.86)		(63.60)		(15.53)
$R_{rangeland}$		56.55		312.3		-		-
		(41.27)		(477.8)		-		-
$R_{pastureland}$		-36.37		-92.14***		82.98		-0.983
		(23.77)		(22.18)		(120.4)		(7.872)
R_{urban}		-2.933		155.7***		282.6***		18.24
		(16.26)		(33.58)		(53.56)		(23.09)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County
Obs.	1,262	1,262	527	527	320	320	1,019	1,019
R-squared	0.502	0.504	0.392	0.395	0.793	0.801	0.565	0.565

Robust standard errors in parentheses

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

Notes: This regression use observations from Illinois, Indiana, Iowa, Nebraska, Minnesota, North Dakota, South Dakota, Michigan, Wisconsin, Ohio, Kentucky. Compared with Table 2 and Table 4, one can see that as I restrict the sample to Corn Belt counties, sample size drops dramatically and data gets much noisier.

Table A.4: Regression results by region

DEP. VAR.	Kjeldahl nitrogen Total		Nitrate Total		Kjeldahl nitrogen Total		Nitrate Total		Kjeldahl nitrogen Total		Nitrate Total	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
CRP ratio	-2.350***		0.317		-3.243***		-0.186		17.77**		-19.14	
	(0.821)		(4.629)		(0.343)		(5.985)		(8.645)		(17.21)	
R_{CRP}		-15.04***		-89.06*		-16.03***		5.625		45.29		-57.78
		(5.258)		(48.43)		(3.275)		(10.35)		(29.45)		(56.05)
$R_{noncrop}$		-6.836		-107.6***		-3.867		-71.51		-5.819		9.128
		(10.31)		(38.77)		(4.387)		(54.36)		(7.885)		(66.54)
$R_{rangeland}$		-1.239		-40.18***		-7.174*		-11.52		42.05		-239.8
		(7.192)		(13.29)		(3.905)		(56.32)		(67.90)		(239.2)
$R_{pastureland}$		9.661		-13.24		6.057		1.516		14.88**		20.42
		(6.006)		(12.71)		(5.686)		(32.93)		(5.867)		(26.06)
R_{urban}		-10.62***		32.43**		-2.637		-29.78		9.835**		67.73
		(2.941)		(15.33)		(6.088)		(37.71)		(4.361)		(65.28)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County	County	County	County
Obs.	2,650	2,652	2,112	2,112	2,349	2,351	1,975	1,975	1,966	1,974	975	981
R^2	0.568	0.573	0.398	0.402	0.400	0.402	0.419	0.420	0.258	0.265	0.408	0.411
Region	West	West	West	West	South	South	South	South	MS River	MS River	MS River	MS River

Robust standard errors in parentheses

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

Notes: West means observations from Arizona, California, Colorado, Idaho, Kansas, Oklahoma, Texas, Oregon, Utah, Washington and Wyoming. South stands for Kansas, Texas, Colorado and New Mexico. MS River means states along Mississippi River: Minnesota, Iowa, Wiconsin, Illinois, Missouri, Arkansas, Tennessee, Mississippi, Louisiana and Kentucky. The regression results are noisy but consistent with previous results using national dataset. I only include Kjeldahl nitrogen and nitrate because other categories appear to be too noisy.

Table A.5: Regression results using original CRP data with no interpolation

DEP. VAR.	(1) Ammonia and ammonium Dissolved	(2) Ammonia and ammonium Total	(3) Inorganic nitrogen Dissolved	(4) Inorganic nitrogen Total	(5) Kjeldahl nitrogen Dissolved	(6) Kjeldahl nitrogen Total	(7) Nitrate Dissolved	(8) Nitrate Total	(9) Nitrite Dissolved
CRP ratio	1.106	-1.517	-1.934	-19.08	7.604***	-4.557***	-1.139	-5.678	-0.280
	(1.395)	(8.114)	(1.280)	(12.06)	(2.778)	(1.719)	(1.315)	(5.832)	(0.617)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County
Obs.	47,900	4,878	54,684	6,067	15,111	4,661	45,948	3,953	44,588
R-squared	0.386	0.704	0.399	0.590	0.486	0.583	0.387	0.468	0.453

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

DEP. VAR.	(10) Nitrite Total	(11) Nitrogen mixed forms Dissolved	(12) Nitrogen mixed forms Total	(13) Organic nitrogen Dissolved	(14) Organic nitrogen Total	(15) Phosphate Dissolved	(16) Phosphate Total	(17) Phosphorus Dissolved	(18) Phosphorus Total	(19) Phosphate-phosphorus Total
CRP ratio	-0.115	0.662	-1.422	2.913***	-7.761**	-1.750*	-22.25**	6.911***	8.817**	-22.46***
	(1.926)	(1.271)	(2.766)	(1.088)	(3.053)	(0.979)	(9.082)	(2.276)	(4.322)	(6.666)
Cluster	County	County	County	County	County	County	County	County	County	County
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County	County
Obs.	2,978	24,176	4,390	21,578	4,614	45,149	1,825	17,213	8,051	2,530
R-squared	0.778	0.358	0.379	0.376	0.559	0.418	0.665	0.531	0.531	0.622

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table shows regression results using original land use data without linear interpolation. It shows the results are consistent with using the main dataset in this paper but noisy.

Table A.6: Regression results using censored data

DEP. VAR.	(1) Ammonia and ammonium Dissolved	(2) Ammonia and ammonium Total	(3) Inorganic nitrogen Dissolved	(4) Inorganic nitrogen Total	(5) Kjeldahl nitrogen Dissolved	(6) Kjeldahl nitrogen Total	(7) Nitrate Dissolved	(8) Nitrate Total	(9) Nitrite Dissolved
CRP ratio	0.439	-2.157	0.436	4.523	2.572	-1.712	1.185	-9.059***	-2.020**
	(1.430)	(3.423)	(0.824)	(8.681)	(2.324)	(1.516)	(1.010)	(3.469)	(0.971)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County
Obs.	46,468	8,577	80,923	10,302	20,375	8,420	63,028	6,682	16,080
R-squared	0.372	0.587	0.329	0.425	0.403	0.421	0.314	0.488	0.428

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

DEP. VAR.	(10) Nitrite Total	(11) Nitrogen mixed forms Dissolved	(12) Nitrogen mixed forms Total	(13) Organic nitrogen Dissolved	(14) Organic nitrogen Total	(15) Phosphate Dissolved	(16) Phosphate Total	(17) Phosphorus Dissolved	(18) Phosphorus Total	(19) Phosphate-phosphorus Total
CRP ratio	16.29*	2.408*	-2.819	2.242	-2.579	-2.754*	-26.18*	-2.988	1.053	-18.93**
	(8.383)	(1.276)	(2.003)	(1.561)	(1.591)	(1.425)	(13.66)	(2.524)	(1.883)	(7.370)
Cluster	County	County	County	County	County	County	County	County	County	County
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County	County	County	County	County
Ob.	2,371	22,758	4,189	14,035	6,543	56,887	2,511	23,557	12,629	3,187
R-squared	0.489	0.370	0.502	0.349	0.415	0.352	0.366	0.415	0.478	0.436

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table shows regression results using original censored water quality data, meaning there is no replacement for "below quantification limit" records (these type of records consists of around 40% of water quality data). It shows the results are consistent with using the main dataset in this paper.

APPENDIX B. ADDITIONAL MATERIAL FOR CHAPTER 2

Table B.1: Regression results, small panel, equation (2)

	(1)	(2)	(3)	(4)	(5)	(6)
Small panel: Log Estimated Population Supplied						
ns×drought_5	0.312*	0.422*	-0.0387	0.247	0.393	-0.00522
	(0.183)	(0.249)	(0.0998)	(0.193)	(0.257)	(0.128)
ns	-0.376	-0.552	0.0729*	-0.402*	-0.552*	-0.0951
	(0.269)	(0.383)	(0.0392)	(0.233)	(0.330)	(0.135)
drought_5	-0.0447	-0.0967	0.0847	-0.0577	-0.104	0.0700
	(0.0671)	(0.0899)	(0.0970)	(0.0719)	(0.0968)	(0.102)
Observations	2,732	1,912	820	2,733	1,913	820
R-squared	0.859	0.830	0.951	0.751	0.728	0.916
Community FE	Yes	Yes	Yes	No	No	No
County FE	No	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor
	(1)	(2)	(3)	(4)	(5)	(6)
Large panel: Log Estimated Population Supplied						
ns×drought_5	0.178***	0.235***	0.0565*	0.0210	0.170**	-0.175**
	(0.0229)	(0.0324)	(0.0296)	(0.0516)	(0.0702)	(0.0698)
ns	-0.200***	-0.211***	-0.157***	-1.125***	-1.160***	-1.131***
	(0.0269)	(0.0366)	(0.0336)	(0.0353)	(0.0451)	(0.0580)
drought_5	0.0270	0.0197	0.0428	-0.0126	-0.0931	0.126**
	(0.0211)	(0.0296)	(0.0295)	(0.0431)	(0.0603)	(0.0581)
Observations	36,094	23,860	12,234	39,408	25,650	13,678
R-squared	0.970	0.966	0.977	0.566	0.594	0.593
Community FE	Yes	Yes	Yes	No	No	No
County FE	No	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Aquifer	All	Major	Minor	All	Major	Minor

Robust standard errors in parentheses

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

Notes: This table represents regression results from estimating equation (2) using data in the large panel. Term $ns \times drought\_5$ is the coefficient of interest, and it represents the effect of "no surface water" on population during drought.