Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Cowles Foundation Discussion Papers | Cowles Foundation

11-1-2019

'Follow the Data' — What Data Says About Real-world Behavior in Commons Problems

Caleb M. Koch

Heinrich H. Nax

Follow this and additional works at: https://elischolar.library.yale.edu/cowles-discussion-paper-series

 Part of the Economics Commons

Recommended Citation

Koch, Caleb M. and Nax, Heinrich H., "'Follow the Data' — What Data Says About Real-world Behavior in Commons Problems" (2019). *Cowles Foundation Discussion Papers*. 53.
https://elischolar.library.yale.edu/cowles-discussion-paper-series/53

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

"FOLLOW THE DATA"— WHAT DATA SAYS ABOUT REAL-WORLD BEHAVIOR IN COMMONS PROBLEMS

By

Caleb M. Koch and Heinrich H. Nax

November 2019

COWLES FOUNDATION DISCUSSION PAPER NO. 2198



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

"*Follow the data*" — What data says about real-world behavior in commons problems[†]

First draft: August 2018
This draft: October 2019

Working paper

Caleb M. Koch[a], Heinrich H. Nax[b]

[a]*Computational Social Science, Dept. of Humanities, Social and Political Sciences, ETH Zürich, Switzerland*
[b]*Behavioral Game Theory, Dept. of Humanities, Social and Political Sciences, ETH Zürich, Switzerland*

Abstract

We test the game-theoretic foundations of common-pool resources using an individual-level dataset of groundwater usage that accounts for 3% of US irrigated agriculture. Using necessary and sufficient revealed preference tests for dynamic games, we find: (i) a rejection of the standard game-theoretic arguments based on strategic substitutes, and instead (ii) support for models building on reciprocity-like behavior and strategic complements. By estimating strategic interactions directly, we find that reciprocity-like interactions drive behavior more than market and climate trends. Taken together, we take a step toward developing more realistic models to understand groundwater usage, and related issues pertaining to tragedy of the commons and commons governance.

Keywords: Common-pool resources – US agriculture – groundwater – dynamic game theory – revealed preferences – panel data – identification

[†]This paper would not have been possible without support. We thank participants at conferences and seminars hosted by ETH Zürich, University of Zürich, KAUST, University of Oxford, and Stony Brook University for their generous feedback. We also thank Rahul Deb, Rod DeBuhr, Andreas Diekmann, Dana Divine, Stefano Duca, Julia Franck, Philip Grech, Dirk Helbing, Barbara Ikica, Nate Jenkins, Yuichi Kitamura, Marie Krausnick, Jonathan Newton, Bary Pradelski, John Quah, Tatsuyoshi (Yoshi) Saijo, Koji Shirai, Scott Snell, Aaron Young, and, last but not least, Peyton Young for their support. All views expressed herein are ours and do not necessarily reflect the views of others, in particular, the Upper Big Blue District, the Upper Republican District or the Nebraska Department of Natural Resources Districts. Any remaining mistakes are our own.

Email addresses: calebkoch@ethz.ch (Caleb M. Koch), hnax@ethz.ch (Heinrich H. Nax)

November 20, 2019

Contents

1 Introduction **4**
- 1.1 Motivation . . . 4
- 1.2 Main contributions . . . 7
- 1.3 Structure of the paper . . . 9

2 Our data **9**
- 2.1 Groundwater in the American Midwest . . . 9
- 2.2 Governing groundwater: Nebraska's Natural Resources Districts . . . 10
- 2.3 Empirical regressors . . . 12

3 Rationalizing data with dynamic games **15**
- 3.1 Groundwater dynamic games . . . 15
  - 3.1.1 Revenue . . . 16
  - 3.1.2 Intra-season groundwater depletion . . . 16
  - 3.1.3 Cost of pumping groundwater . . . 17
  - 3.1.4 Future strategies and payoffs . . . 17
  - 3.1.5 Definition of groundwater dynamic games . . . 18
- 3.2 Integrating a quota policy . . . 19
- 3.3 GW dynamic games can rationalize any dataset . . . 19

4 Revealed preference test of strategic behavior **22**
- 4.1 Strategic substitutes vs. complements . . . 24
- 4.2 A test based on revealed preferences . . . 25
- 4.3 Test setup . . . 26
- 4.4 Results . . . 27
- 4.5 Robustness check I: Larger N . . . 31
- 4.6 Robustness check II: Network structure . . . 32

5 Estimating strategic interactions **33**
- 5.1 Data . . . 33
- 5.2 Regression setup . . . 34
- 5.3 Identification . . . 35
- 5.4 Testing procedure . . . 36
- 5.5 Results . . . 37
  - 5.5.1 Robustness of IVs . . . 39
  - 5.5.2 Main findings . . . 40
- 5.6 Predicting revealed preferences . . . 42

6 Outlook **44**

A Proofs **48**
A.1 Brief summary of nonsmooth optimization 48
A.2 Proof of Theorem 1 48
A.2.1 (B) $\Longrightarrow$ (A) 48
A.2.2 (A) $\Longrightarrow$ (B) 50
A.3 Proof of Theorem 2 51
A.3.1 (D) $\Longrightarrow$ (C) 51
A.3.2 (C) $\Longrightarrow$ (D) 52

B Instrumental variables **53**
B.1 Derivation of IVs 54

C Robustness checks **55**
C.1 Revealed preferences on different network structures 55
C.2 Strategic interactions on different network structures 57

"For that which is common to the greatest number has the least care bestowed on it. Everyone thinks chiefly of his own, hardly at all of the common interest," (Aristotle, *Politics*, Book II, Chapter 3),

"Nothing is more useful than water: but it will purchase scarce any thing; scarce any thing can be had in exchange for it," (Adam Smith, *Wealth of Nations*, 1776, Book I, Chapter 4)

1. Introduction

1.1. Motivation

One of the biggest challenges facing the 21th century is sustaining natural resources in the face of a growing global population and economy. Global resource extraction has more than tripled since 1970, outpacing global population growth by 50% (International Resource Panel (IRP), 2019). Since 2000, the growth rate in global resource extraction has matched the global GDP growth rate. Looking forward, natural resource extraction is expected to double by 2050 (see the same IRP 2019 report).

Of all the natural resources, groundwater—which is the focus of our own study—plays a particularly important role in the sustenance of economic systems, and yet faces serious depletion risks. More than 95% of usable and accessible freshwater is groundwater (Gleick and Palaniappan, 2010), which makes it somewhat unsurprising that almost 80% of global agriculture relies on groundwater irrigation (see UNESCO reports from 2009 and 2012). The present rate of groundwater extraction, however, is unsustainable. Nearly every major groundwater basin has been flagged as *"depleting due to over-appropriation"* (Gleeson et al., 2012) because of irrigation, including Northern India, Southern California, and—the focus of our own study—the American Midwest.[1] Groundwater depletion has also been linked to other problems such as climate change (Green et al., 2011; Taylor et al., 2013) and rising sea levels (Konikow, 2011; Wada et al., 2012).

The central goal of this paper is to test two major, yet opposing, streams of research that try to explain why groundwater levels are depleting at accelerating rates, despite the fact that nobody wants groundwater to deplete.

[1]Groundwater-usage in the American Midwest (i.e., of the High Plains Aquifer) is a classic CPR case study (Coman, 1911; Ostrom, 2011), illustrating the various challenges transitioning from collective access investments initially to today's risks of over-usage. See Gleick (2002, 2003), Foster and Chilton (2003), Giordano (2009), Aeschbach-Hertig and Gleeson (2012), and Famiglietti (2014).

The first major stream, which is the neoclassical, textbook answer in economic theory, is commonly referred to as the *"Tragedy of the Commons"* (Lloyd, 1833; Gordon, 1954; Scott, 1955; Hardin, 1968). According to this hypothesis, the key issue is that a rational, self-interested individual views resource sustainability as diametrically opposed to maximizing profits. Any profit-maximizing user would extract resources until the marginal rate of return equaled the marginal cost of extraction. One user extracting as such is seldom harmful because the entire cost of depletion is borne. The problem, however, is when hundreds if not thousands of users are all trying to maximize profit. If 99 out of 100 users are all pumping groundwater sustainably, then the 1 can profit by pumping more groundwater while only bearing 1/100th of the cost of depletion. All users thinking this way, the conclusion is that a population using a resource strictly for profit will deplete the resource—hence the "Tragedy of the Commons". Inherent in this hypothesis is the divergence between individual and collective rationality: nobody wants the resource to deplete, but it depletes nevertheless.[2]

The second, more 'behavioral' stream of research puts forward a very different story, arguing that the self-interested forces are not the sole nor principal drivers of human behavior. While the previous stream hypothesizes that sustainable pumping incentivizes unsustainable pumping, this stream posits various mechanisms that all result in *reciprocity-like behavior*, whereby sustainable pumping can incentivizes *sustainable* pumping. From more neoclassical game theory, such mechanisms include non-cooperative collusion (Green and Porter, 1984; Rotemberg and Saloner, 1986), social norms and conventions (Young, 1993; Burke and Young, 2010; Young, 2015), and shared intentions (Newton, 2017). From behavioral and experimental game theory, mechanisms include conditional cooperation (Fischbacher et al., 2001), social preferences (Rabin, 1993; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Fehr and Schmidt, 2006), interdependent preferences (Sobel, 2005), interactive preferences (Levine, 1998; Nax et al., 2015), social pressuring (Gaechter and Fehr, 1999; Masclet, 2003), and collective identities (Eckel and Grossman, 2005; Charness et al., 2007). Perhaps most notably, Ostrom (1990) asserted that reciprocity-like patterns emerge because of local traditions, norms, values, social ties, etc.[3]

[2] In the relevant theory literature, CPR interactions are typically modeled as dynamic-stochastic games à la Shapley (1953). For some prominent examples, see—in order of appearance—Levhari and Mirman (1980), Reinganum and Stokey (1985), Clemhout and Wan (1985), Negri (1989), Dockner and Van Long (1993), Dutta and Sundaram (1993), Clemhout and Wan (1993), Gaudet et al. (2001), Mirman and To (2005), Sorger (2005), Wirl (2007, 2008), Koulovatianos (2015), and Jaskiewicz and Nowak (2015).

[3] See Ostrom (1965, 1998, 2000, 2002, 2003, 2005, 2007, 2009) for some of her most relevant work in this area. See Lara (2015) for an overview over Ostrom's work and how her perspective on and use of game theory evolved.

The above disagreement between the neoclassical, textbook answer and more behavioral answers can be interpreted as a debate between strategic substitutes versus complements. The former asserts that sustainable behavior incentivizes *un*sustainable behavior—i.e. groundwater usage decisions are *strategic substitutes*—and groundwater generally depletes. By contrast, the latter asserts that sustainable behavior incentivizes sustainable behavior—i.e. groundwater usage decisions are *strategic complements*—and groundwater levels can be sustainable. This debate remains unsettled, particularly for the more general class of common-pool resources (CPRs) under which groundwater falls, because of conflicting empirical evidence. In a framed field experiment, Noussair et al. (2015) find that fishermen behave in a way that is *"consistent with standard economic theory that assumes selfish preferences and non-cooperative behavior"* (p. 413). Similar evidence was found in an experiment by Stoop et al. (2012). By contrast, through experimental and survey data, Rustagi et al. (2010) find evidence that *"conditional cooperation and costly monitoring explain success in forest commons management"* (p. 961). Beyond the conflicting evidence, external validity concerns have been voiced (Voors et al., 2011).

What continues to be missing is empirical evidence based on *observational* CPR data—especially with groudnwater data—that circumvents external validity concerns and that can delienate between strategic substitutes versus complements. We see two reasons why such studies are missing until now. First, as is argued by Poteete and Ostrom (2008), there is a lack of real-world data availability, and what is available lacks the detailed microscopic (individual-level) information necessary to test the strategic foundations of CPR usage. This is beginning to change as monitoring technologies have improved, which incidentally made our study possible. Second—and we shall come back to this below—the appropriate test procedures were not available until recently.

To date, most research on natural resources and the tragedy of the commons has focused on fisheries, while groundwater has played a comparatively minor role in this discussion. This is because, in part, fishery depletion was a serious issue in the years leading up to and that followed Hardin's (1968) popularization of the tragedy of the commons. Gordon (1954), focusing almost exclusively on fisheries, is credited as the first to formulate the tragedy of the commons as an economic model (see Banzhaf et al., 2018, who proposed a revealed preference test for Gordon's model). Smith (1969) advanced the economic theory of natural resources using fisheries as the leading example. The seminal work by Levhari and Mirman (1980), capitulating tragedy of the commons as a dynamic non-cooperative game, was dedicated to fisheries. Fisheries have been the subject of many case studies (Jensen,

2007), field experiments (Fehr and Leibbrandt, 2011; Carpenter and Seki, 2011; Stoop et al., 2012), framed field experiments (Noussair et al., 2015), and artefactual field experiments (Torres-Guevara and Schlueter, 2016). Beyond theory and experiments, Huang and Smith (2014) *"conduct the first empirical investigation of common-pool resource users' dynamic and strategic behavior at the micro-level using real-world data"*—and this data was from a North Caroline fishery. These are but a few of the many economic studies on fisheries. In the end, this literature has greatly contributed to the long-term economic viability of global fisheries, as the policies developed therein *"have been successful in some 150 major fisheries of 170 species in seventeen countries ... Since 1986, the system has been effective, largely eliminating overfishing, restoring stock to sustainable levels, and increasing fishermen's profits,"* (Stavins, 2011, p. 91).

However, the success of fisheries have not been transferred to groundwater because, as noted above, such data has not been available until recently. And much like the fishery literature, studying the micro-behavior underlying groundwater usage can have far-reaching policy implications, especially since strategic complements/substitutes result in starkly different policy recommendations. Based on the more neoclassical, textbook model, Hardin (1968) claimed that only centralized authorities and private property institutions could sustain CPRs in the long-run. By contrast, Ostrom (1990) built on strategic complements and reciprocity-like behavior in order to advocate for policies that foster self-governance, local norms, etc. While global groundwater depletion trends persist, there exists room to further study which, if either, policy paradigm is most appropriate in the context of groundwater. We return to this thought in the discussion section.

1.2. Main contributions

In this paper, with policymaking implications in mind, we take a critical look at the strategic foundations of CPRs. We do so by analyzing a unique, individual-level dataset of groundwater-usage from Nebraska, US—which sits on top of one of the world's largest groundwater reservoirs—that accounts for three percent of US irrigated agriculture. Groundwater is renewable in the regions that we study, however, notable depletion exists. The dataset documents farmer-level groundwater pumping and includes a rich set of control variables including land-size, soil quality, longitude-latitude coordinates, climate data, market data, as well as farmer-level information about whether groundwater recharges or depletes due to pumping. Importantly, this dataset also describes farmer-level pumping in policy-free and quota-policy areas, which gives us a somewhat unique opportunity to study how a policy can change farmers' behavioral incentives.

As noted above, our central question is whether groundwater-usage decisions are strategic substitutes or complements. We address this question, first by developing a new revealed preference test for dynamic games, and second developing a new procedure for empirically analyzing networked panel data.

Our revealed preference test pits both hypotheses against each other. We define two non-parametric classes of "Groundwater Dynamic Games", the first with strategic substitutes and the second with strategic complements such that each postulates the exact opposite strategic behavior. These games allow groundwater levels to dynamically evolve and farmers to plan future farming seasons accordingly. Building on recent work by Carvajal et al. (2013), we characterize both classes of dynamic games as revealed preference tests with a parametric, simple, and easy-to-implement linear system of equations. Importantly, we show that these tests can be used to study policy-free and quota-policy datasets, which allows us to compare real-world behavior from both regions in an unbiased manner.

While the first test has the advantage of being non-parametric, it crucially relies on game-theoretic arguments and thereby precludes the opportunity to test for more behavioral drivers of groundwater pumping. Therefore, our second test complements the first by making explicit parametric assumptions but avoiding any game-theoretic assumptions. By building on Mutl and Pfaffermayr (2011), we pose the question of strategic substitutes versus complements by incorporating a strategic interaction term into a regression framework. This allows us to test for positive (or strategic complements) versus negative (or strategic substitutes) directly while controlling for a range of possibly confounding factors. Importantly, we make a methodological contribution by allowing the network that governs strategic interactions to be *unknown*. This allows us to leverage our latitude-longitude data to test behavior under a wide rage of possible mechanisms that could drive reciprocity-like behavior. Our test is assured to be consistent and robust to endogeneity thanks to Mutl and Pfaffermayr (2011).

When applied to the dataset from Nebraska, both empirical tests point to the same answer: we find a resounding rejection of the standard tragedy of the commons framework and, even more surprisingly, strong evidence supporting models that build on reciprocity-like behavior, i.e. models that asset behavior exactly contrary to what tragedy of the commons would predict. Our findings also shed light on the influence of institutions on groundwater pumping behavior. When we compare data from the policy-free and the policy region, it is clear that introducing quota schemes renders farmers more likely to behave as if they are in a tragedy of the commons situation. Taken together, we take a step toward developing more realistic models to understand groundwater usage, and related questions regarding the tragedy of the

commons and its governance.

1.3. Structure of the paper

The paper is structured as follows. In Section 2, we describe the various datasets used in our analyses. In Section 3, we present a game-theoretic framework that forms the basis of our analysis. Our revealed preference test and results are in Section 4. We empirically estimate strategic interactions in Section 5. Finally, we conclude in Section 6.

We have three appendices. Proofs are relegated to Appendix A. The estimation strategy for our regression framework is provided in Appendix B. Robustness checks are reported in Appendix C.

2. Our data

We study individual-level irrigation decisions from two districts above the High Plains Aquifer (HPA) system (often referred to as the Ogallala Aquifer, which is actually just one of the aquifers in the system). Our two districts lie in Nebraska, US. This dataset is rich on an individual-level, which ultimately makes our test of game-theoretic foundations of tragedy of the commons possible. But before we describe the dataset itself, we would like to describe the context of the data in more detail.

2.1. Groundwater in the American Midwest

Nebraska (the "Cornhusker State") is one of the key contributors to US agriculture, being the number one red meat producer, second pinto bean producer, and third corn producer among the US states as per 2014 (Nebraska Department of Agriculture, 2014). Nebraska's agricultural productivity also competes on a global scale, being the twelfth largest player internationally in terms of irrigated land. Agriculture is the main economic activity in Nebraska with farm-related income contributing over $23 billion (24% of Nebraska GDP in 2013) and a quarter of jobs coming from agriculture.

Nebraskan agriculture crucially relies on freshwater from the HPA, which stretches across eight US states from southern South Dakota to northern Texas.[4] Approximately 90% of groundwater pumped from the HPA is used to irrigate crops. But like most other aquifers around the world, the HPA is depleting due to over-pumping. Current estimates suggest that 15% of the aquifer has been depleted since the 1930s, and the current *rate* of depletion signals

[4]To give an idea of its size: if spread across the US, the HPA would cover all fifty states with 1.5ft of water (`https://www.scientificamerican.com/article/the-ogallala-aquifer/`).

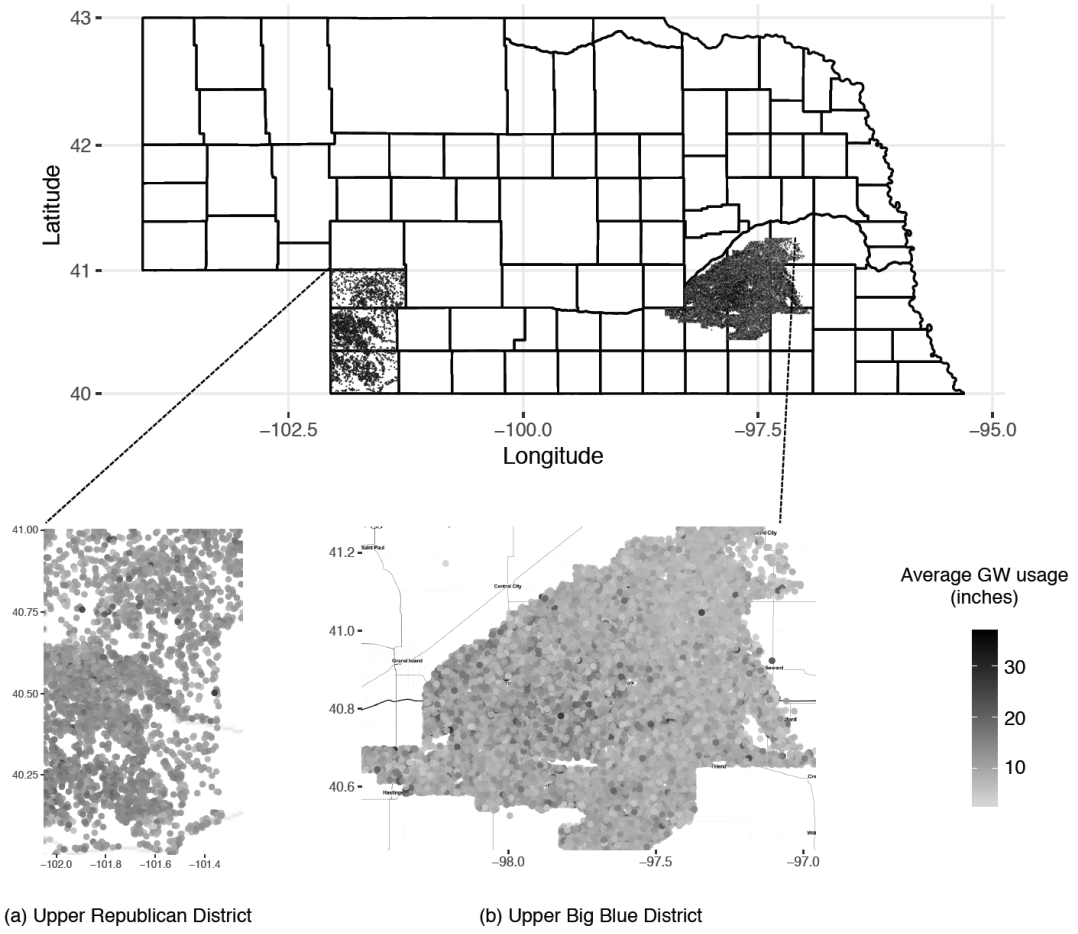


Figure 1: Average water-per-acre usage in the (a) Upper Republican District from 2001–14 and (b) Upper Big Blue District from 2007–2014. One inch = one inch of groundwater on top of one acre of land, which is 27,157 gallons.

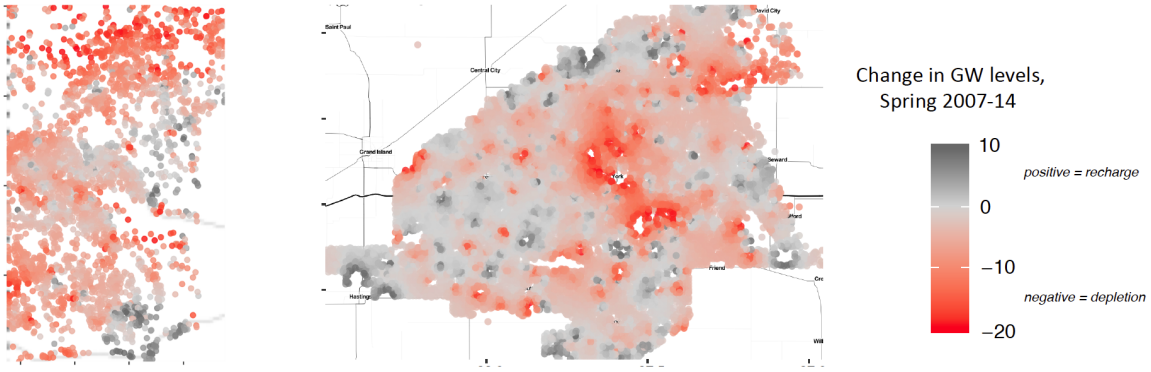
that further decline in groundwater levels is likely (Haacker et al., 2016; Steward and Allen, 2016). Nebraska, with nearly two-thirds of all the groundwater from the HPA, is particularly dependent on the HPA and particularly critical for its sustainability.[5]

2.2. Governing groundwater: Nebraska's Natural Resources Districts

Groundwater has been at the heart of policymaking in the American Midwest for a long time. But at first, policy makers struggled to incentivize farmers to pump groundwater

[5]The HPA is considered a *renewable* CPR in Nebraska, since snowmelt from the Rocky Mountains and annual rainfall are, with moderate levels of pumping, sufficient to sustain groundwater levels.

Groundwater depletion/recharge in the URD and UBB from 2007–14



(a) Upper Republican District (b) Upper Big Blue District

Figure 2: Well-per-well change in water table levels (i.e. groundwater levels) from Spring 2017–14 in the (a) Upper Republican District and (b) Upper Big Blue District.

because it was prohibitively costly to build and maintain irrigation wells. The first article in the first issue of the AER, "Some Unsettled Problems of Irrigation" (Coman, 1911), was dedicated to collective action problems associated with coordinating farmers to use irrigation technologies. However, the situation changed fundamentally in the 1960s with the invention of the center pivot irrigator, which made access to groundwater easier and cheaper. Center pivot irrigation proliferated the number of farmers pumping groundwater from the HPA for agricultural purposes (see fn. 8). Groundwater depletion was inevitable. In response to depleting groundwater levels, in 1969 Nebraska founded 23 Natural Resources Districts (NRDs) to govern groundwater pumping. In fact, each district has full jurisdiction over governance of all natural resources within the district borders—which includes groundwater. Since then, the NRDs have adopted very different policies to govern groundwater pumping, and these differences play an important role in our empirical analysis below.

Our data concern individual-level groundwater-usage behavior in two of the 23 NRDs, namely the "Upper Big Blue" NRD (UBB) and "Upper Republican" NRD (URD). Studying the UBB and URD provides a somewhat unique opportunity to compare behavior in different policy environments, as the UBB and URD adopted starkly different groundwater dynamics and governance approaches.

The "Upper Big Blue" District — The UBB is the largest and most irrigated district in Nebraska, alone accounting for 15% of Nebraska's irrigated acres and 2% of the US's irrigated acres. Given the size, it is perhaps surprising that this district takes a policy-free approach to regulating groundwater usage. Farmers pump groundwater without restrictions.

The only regulatory measure taken by the UBB was requiring all farmers to install pumping meters on irrigation wells in 2006. This is why the dataset begins in 2007, and we have data through 2014.[6]

The "Upper Republican" District — The URD is roughly half the size of the UBB in terms of groundwater irrigated land. Unlike the UBB, the URD was the first district in Nebraska to introduce a *quota policy* to govern individual-level groundwater usage.[7,8] This policy requires farmers to pump less groundwater than a certain set quota over the course of five years (with usage being monitored via pumping meters like in the UBB); farmers who fail to do so are subject to a fine. The first quota was implemented from 1982–87, allowing farmers to pump 80" (=80" of groundwater per acre $\approx$ 2 million gallons). Since, this quota was decreased in an attempt to chase declining groundwater levels: the quota was 75" for 1988–92, 72.5" for 1993–97 and the two subsequent five year time windows, 67.5" for 2008–12, and 65" for 2013–17. Our data comprises the years 2001–14.[9]

2.3. Empirical regressors

Our empirical analysis utilizes data from different sources concerning (i) groundwater usage collected by the UBB and the URD, (ii) groundwater depletion, soil types and other controls collected by the University of Nebraska at Lincoln (UNL), (iii) market prices and land rental rates from the US Department of Agriculture,[10] (iv) and rain and temperature from the US National Oceanic and Atmospheric Administration.[11] Below, we categorize and describe the key variables pertinent to our study.

Farmer's groundwater usage — Our 'main' datasets are the annual UBB and URD records concerning farmer-level groundwater usage. For each of the roughly 10,000+ farmers in the UBB, we have data on total groundwater usage for each season from 2007–14, which amounts

[6]We would like to cordially thank Rod DeBuhr, Marie Krausnick, and Scott Snell at the UBB for permission and access to UBB data, as well as conversations that have greatly contributed to our study.

[7]The US Geological Survey estimated that approximately 150 million acre-feet (=4.9×10^{13} gallons) of renewable groundwater underlies the URD.

[8]In 1965, only 50,000 acres in the URD (or 3% of the URD) were being irrigated because of the sandier soil that, without excessive pumping practices, were comparatively impractical to farm. With the proliferation of center pivot irrigation in the late 1960s through the 1970s, this number increased by more than a quarter million and plateaued to about 425,000 acres of irrigated cropland (Stephenson, 1996).

[9]We would like to express our gratitude to Nate Jenkins and Julia Franck at the URD for permission and access to URD data, as well as conversations that have greatly contributed to our study.

[10]`https://www.nass.usda.gov/Statistics_by_State/Nebraska/index.php`

[11]`https://www.noaa.gov`

to nearly 100,000 observations. For each of the 3,000+ farmers in the URD, we have data on the total groundwater usage for each season from 2001–14, which amounts to nearly 50,000 observations. Data include (i) latitude-longitude coordinates of farmers' farmland and (ii) the number of acres the groundwater serviced. We plot the spatial location of farmers in the UBB and URD in Figure 1; black-white shading represents the average groundwater usage of each farmer.

Neighbors — Because the dataset includes latitude-longitude coordinates, we have the opportunity to account for the networked strategic interactions of groundwater usage. In other words, we have a way to control for *with whom* a farmer interacts strategically. There are four reasons we shall emphasize a farmer's *neighbors* as those who play an important role in strategic interactions. First, neighbors are a natural starting point as farmers are more likely to observe, communicate and strategically interact in close proximity than with more distant farmers. Second, the importance of locality was also emphasized by UBB farmers in qualitative interviews we made in 2015. Third, and in line with Ostrom's local governance principles, informal institutions—such as moral and social norms—operate on a local level. Finally, there are so-called *"local drawdown"* effects (Theis, 1941; Cooper and Jacob, 1946; Spalding and Khaleel, 1991) operating at the hydrological level. Local drawdown is caused by pumps drawing 'local' groundwater in close proximity first, which causes local depletion of neighbors' groundwater-levels—season-long pumping gradually increases the set of neighbors on whom a farmer induces local drawdown. The hydrology literature typically defines a *"radius of influence"* in which local drawdown is measurable by the end of the farming season, which ranges from 200ft to a mile (Cashman and Preene, 2001).

Groundwater dynamics — The HPA is a so-called *"water table aquifer"* because there exists a clear *"upper-boundary"* of the aquifer—often referred to as the *water table level*—that moves closer to the surface-level if the resource is recharging and further away if the resource is depleting (Gutentag et al., 1984; McGuire et al., 2003). We follow geologists, hydrologists and government officials in using water tables as proxies for depletion risks. Researchers at UNL kindly provided us with relevant data for UBB and URD for our years of interest.[12] The data include measurements at more than 500 locations in the UBB and URD, time-stamped as Spring or Fall, which allows us to empirically track groundwater depletion within season, as well as resource recharge between seasons. Based on this data, we use standard spatial

[12]We are grateful to Dana Divine and Aaron Young for helping us include this data in our anlayses.

Table 1: Descriptive statistics, 2008–12. GW = groundwater.

Panel A	*Upper Big Blue District*					*Upper Republican District*				
	(2008)	(2009)	(2010)	(2011)	(2012)	(2008)	(2009)	(2010)	(2011)	(2012)
Number of observations	10,375	10,546	10,435	10,426	10,714	3,209	3,209	3,209	3,209	3,209
GW-usage$_i^t$	5.79	8.27	6.38	5.85	13.81	12.07	9.53	10.72	10.18	19.04
(GW-usage$_i^t$ SD)	(4.2)	(5.1)	(4.3)	(4.1)	(6.9)	(4.01)	(4.12)	(4.09)	(4.19)	(5.49)
Groundwater control variables										
Spring GW$_i^t$ *(ft)*	81.18	81.89	81.98	81.68	82.00	113.6	114.39	115.61	114.47	111.91
(Spring GW$_i^t$ SD)	(12.2)	(12.3)	(12.2)	(12.6)	(11.9)	(37.9)	(37.8)	(39.6)	(38.8)	(36.4)
Fall GW$_i^t$ *(ft)*	82.89	82.67	80.70	78.47	84.58	117.97	117.62	112.64	110.56	108.49
(Fall GW$_i^t$ SD)	(9.8)	(8.5)	(8.6)	(8.8)	(7.5)	(40.33)	(39.6)	(23.1)	(26.3)	(20.9)
Annual control variables										
Price of corn^t *($ / bushel)*	5.65	3.55	3.68	7.17	7.17	5.65	3.55	3.68	7.17	7.17
Electricity^t *(¢ / kW-hr)*	11.26	11.51	11.54	11.72	11.88	11.26	11.51	11.54	11.72	11.88
Rain^t *(in)*	16.24	13.85	18.54	22.33	6.52	11.13	14.42	14.08	13.36	4.38
Temperature^t *(°F)*	71.48	68.45	71.08	71.03	72.60	66.57	66.23	67.87	67.97	72.89
Farmland rental rates$_C^t$ *($ / acre)*	170.5	173.5	180.6	197.9	243.8	168.96	141.87	161.45	173.15	191.08
(Rental rate$_C^t$–SD)	(8.2)	(8.2)	(8.6)	(9.4)	(18.6)	(30.2)	(14.2)	(21.3)	(17.2)	(24.7)

Panel B	*Upper Big Blue District*			*Upper Republican District*		
	(mean)	(SD)	(range)	(mean)	(SD)	(range)
Farm control variables						
Land-size$_i$ (*acres*)	101.51	35.72	[3.0 , 349.0]	141.0	57.5	[2.0, 702.9]
Well Depth$_i$ (*ft*)	184.82	31.93	[5.0, 470.0]	248.2	74.2	[22.0, 552.0]
Transmissivity$_i$ (ft^3)	126.34	36.25	[1.66, 249.3]	106.5	3.82	[94.8, 112.4]

interpolation methods to estimate groundwater levels in the Spring and Fall of every season at every location in the dataset.[13]

Control variables — The UNL dataset also records depths of wells, and we use the same spatial interpolation method noted above to estimate the well depth for each farmer. Other, separate datasets from the UNL allow us to use similar procedures to obtain controls for transmissivity[14] and soil types. We add to these datasets by including seasonal rain and

[13]More specifically, we utilize a method called *"kriging"*; see Krige (1951), and Matheron (1963) for the seminal texts and von Stein (2012) for a modern treatment. Under reasonable assumptions, this method provides the best linear unbiased prediction of geo-spatial intermediate values that maintains geologically relevant properties, such as continuity of the water table.

[14]Transmissivity is a metric for how fast groundwater moves across the groundwater basin. Importantly, as a farmer irrigates, higher transmissivity implies the basin 'replaces' faster and alleviates the stress on water table levels during pumping, which helps stabilizes groundwater levels. Hence, farmers with higher transmissivity are less prone to receding water table levels.

temperature (sum and average from April to September, respectively) from data by the US National Oceanic and Atmospheric Administration. Finally, we include three market variables from the US Department of Agriculture: (i) average spot market price of corn from April 1st to October 1st of each season (we focus on corn because, in a typical year, roughly 70% of irrigated land is for corn), (ii) average price of electricity between April and October, and (iii) farmland rental rates which are available averaged at the county level each season (there are nine counties in the UBB and three in the URD).

Summary — In total, our final dataset consists of approximately 100,000 observations of farmer-level groundwater usage in the UBB and 50,000 observations in the URD, as well as a rich set of controls concerning resource dynamics, farmer attributes, seasonal attributes, and market conditions. We summarize the datasets from 2008–12 in Table 1, which contains a complete five year quota period for the URD.

3. Rationalizing data with dynamic games

In this section, we present our framework for studying groundwater usage as a dynamic game. We introduce a class of games that subsumes both policy-free and quota policy environments. Our model builds on Negri (1989), while also taking into account some of the key features of the UBB and the URD. Each dynamic game consists of four ingredients:

(Section 3.1.1) The revenue associated with selling crops.

(Section 3.1.2) Intra-seasonal groundwater depletion due to pumping.

(Section 3.1.3) The cost of pumping groundwater.

(Section 3.1.4) The 'dynamic' inter-seasonal aspects, i.e. how farmers view groundwater-usage 'today' affects strategic behavior and profits in the future.

Next, we introduce these four components formally. Definitions are in bold. For notation, we let $\mathbb{R}_+ = [0, +\infty)$ and $\mathbb{R}_{++} = (0, +\infty)$.

3.1. Groundwater dynamic games

Consider a set of farmers $\mathcal{N} = \{1, 2, \ldots, N\}$ producing (possibly different) crops during each season $\{1, 2, \ldots, T\} = \mathcal{T}$. Each farmer $i \in \mathcal{N}$ is characterized by a land-size $l_i \in \mathbb{R}_{++}$. For every $t \in \mathcal{T}$, farmers plant crops in spring and harvest in the late summer, and in between pump groundwater in order to increase crop yield (or total supply that can be sold after harvest). We

denote farmer i's action space in season $t \in \mathcal{T}$ as $\mathcal{W}_i = \mathbb{R}_+$. An element $w_i^t \in \mathcal{W}_i$ represents groundwater usage-per-acre (this is the unit in Figure 1), and $l_i w_i^t$ represents the total volume of groundwater pumped.

3.1.1. Revenue

Farmers make profits in various ways. They can either sell at the spot market (at possibly different prices in October, November or December) or on the futures market. Alternatively, farmers can store crops in long-term storage facilities (e.g. for raising cattle or selling later on). To capture these diverse revenue streams, we define a ***revenue function*** $\mathcal{R}_i^t : \mathcal{W}_i \to \mathbb{R}_+$, which may be both farmer- and season-dependent, thus controling for different yield and sale strategies across farmers and price changes across seasons. In order to be able to work with this function, we assume that $\mathcal{R}_i^t$ is absolutely continuous and concave in w_i^t.

3.1.2. Intra-season groundwater depletion

Groundwater usage and recharge determine the evolution of the resource. Here, we first focus on how groundwater levels change during the season as a result of farmers pumping. For farmer $i \in \mathcal{N}$ in season $t \in \mathcal{T}$, we denote $g_i^{t,Spring} \in \mathbb{R}_+$ as the distance between the water table level and wellhead in the spring. We denote $g_i^{t,Fall} \in \mathbb{R}_+$ as this distance when farmers stop pumping (i.e. at the beginning of harvest). Two dynamics determine $g_i^{t,Fall}$.

The first is due to aquifer-wide groundwater withdrawals, $W^t = \sum_{i \in \mathcal{N}} l_i (w_i^t - r^t)$, where depletion is offset by recharge from rain, $r^t \in \mathbb{R}_{++}$. We define $\mathcal{G} : \mathbb{R} \to \mathbb{R}$ as a ***global depletion function***, which we assume is (i) absolutely continuous and nondecreasing in $W^t - r^t$ and (ii) slowly changing in individual groundwater-usage (i.e. $\frac{\partial}{\partial w_j} \mathcal{G} \approx 0$ for any $j \in \mathcal{N}$).[15] We then let $g_i^{t,Spring} + \mathcal{G}(W^t - r^t)$ represent the decline/recharge of i's groundwater level that is caused by aquifer-wide groundwater withdrawals.

The second type of change is a local, heterogeneous depletion called *local drawdown* (see Section 2). Let $\mathcal{N}_i \subseteq \mathcal{N} \backslash \{i\}$ denote the set of farmers inducing local drawdown on farmer i. The groundwater-usage of these farmers is $w_{\mathcal{N}_i}^t = \sum_{j \in \mathcal{N}_i} l_j w_j^t$. We define $\mathcal{L}_i : \mathbb{R}_+ \to \mathbb{R}_+$ as a ***local drawdown function***, which is assumed to be absolutely continuous and nondecreasing in $w_{\mathcal{N}_i}^t + l_i w_i^t$. We then let $g_i^{t,Spring} + \mathcal{L}_i(w_{\mathcal{N}_i}^t + l_i w_i^t)$ represent the local decline in i's groundwater level due to local drawdown.

A farmer's groundwater level at the start of harvest is the result of local and global

[15] The latter is a somewhat innocuous assumption since a single farmer represents ~0.008% (0.03%) of the UBB (URD) farming population.

depletion:

$$\forall i, t: \qquad \underbrace{g_i^{t,Fall}(w_i^t, w_{-i}^t)}_{\text{Fall level}} = \underbrace{g_i^{t,Spring}}_{\text{Spring level}} + \underbrace{\mathcal{G}(W^t - r^t)}_{\text{Global depletion}} + \underbrace{\mathcal{L}_i(w_{N_i}^t + l_i w_i^t)}_{\text{Local drawdown}}. \tag{1}$$

3.1.3. Cost of pumping groundwater

The cost of pumping groundwater is proportional to the distance between water table and the pump wellhead, which is increasing as farmers pump. We capture the costs that accrue from a changing pumping distance through

$$C_i^t(w_i^t, w_{-i}^t) = e_i \cdot \int_0^{w_i^t} g_i^{t,Fall}(w_i^t, w_{-i}^t)\ d\xi_i \tag{2}$$

where $e_i > 0$ is the farmer-specific ***energy cost*** of pumping, which takes into account whether the farmer has a diesel or electricity pump, the age of the pump, etc. The interpretation of (2) is that, taking other farmers' groundwater usage as given, a farmer accounts for self-imposed costs associated with his/her own contribution to the depletion he/she experiences.

3.1.4. Future strategies and payoffs

Future strategies — We suppose that every farmer i is forward-looking and strategizes to maximize current *and* future payoffs in seasons $\tau \in \{t+1, t+2, \ldots, T\}$. Let Z_i^τ represent the set of possible information i could have at the beginning of season τ, which may include, for example, realized prices, realized groundwater levels $(g_i^{\tau,Spring}, g_{-i}^{\tau,Spring})$, and past information. Farmers use this information to decide not only how much groundwater to pump, but also which crops to plant, whether to buy insurance, etc. To capture such decisions, we let $\alpha_i^\tau : Z_i^\tau \to A_i^\tau$ represent a *contingency strategy*, where A_i^τ is the set of all possible such decisions. Let $\mathcal{A}_i^\tau = \{\alpha_i^\tau : Z_i^\tau \to A_i^\tau\}$ be the set of possible contingency strategies and $\alpha^\tau = (\alpha_1^\tau, \ldots, \alpha_N^\tau) \in \times_{j\in\mathcal{N}} \mathcal{A}_j^\tau = \mathcal{A}^\tau$.[16] For notation, we let $\mathcal{X}_i = \times_{\tau\in\{t+1,t+2,\ldots,T\}} \mathcal{A}_i^\tau$. A profile of such strategies is denoted as $x = (x_1, \ldots, x_N) \in \times_{j\in\mathcal{N}} \mathcal{X}_j = \mathcal{X}$.

Net-present value — Farmer i's realized payoff in season $\tau \in \{t+1, t+2, \ldots, T\}$ depends on all farmers' (realized) contingency strategies. We denote $\mathcal{P}_i^\tau : \mathcal{A}^\tau \to \mathbb{R}$ as i's payoff in season τ, which depends on the realization of information:

$$(z_1^\tau, z_2^\tau, \ldots, z_N^\tau) = z^\tau \mapsto \mathcal{P}_i^\tau(\alpha^\tau(z^\tau)).$$

[16] Note that any $\alpha_i^\tau \in \mathcal{A}_i^\tau$ may be a Markovian or non-Markovian strategy.

From i's perspective, realized information z^τ is dependent on what happens in season t. As such, we let $z^\tau \sim \mathbb{P}_i^\tau\left(\cdot \mid g_i^{t,Spring}, w_i^t, w_{-i}^t, \dots\right) \equiv \mathbb{P}_i^\tau\left(\cdot \mid \iota_i^t\right)$ where we let ι_i^t represent all relevant information i has to deduce the distribution of z^τ. We assume that $\{\mathbb{P}_i^{t+1}, \mathbb{P}_i^{t+2}, \dots, \mathbb{P}_i^T\}$ admits a well defined expectation over the realization $(z^{t+1}, z^{t+2}, \dots, z^T)$.

A farmer i's ***net-present value***, dependent on current pumping and future contingency strategies equals

$$\begin{aligned}\mathcal{P}_i^t(w_i^t, w_{-i}^t; x) &= \mathcal{R}_i^t(w_i^t) - C_i^t(w_i^t, w_{-i}^t) + \mathbb{E}\left[\sum_{\tau\in\{t+1,\dots,T\}} \delta_i^{\tau-t}\mathcal{P}_i^\tau(\alpha^\tau(z^\tau)) \;\middle|\; \iota_i^t\right] \\ &\equiv \underbrace{\mathcal{R}_i^t(w_i^t) - C_i^t(w_i^t, w_{-i}^t)}_{\text{Profit in season } t} + \underbrace{\mathcal{F}_i^t(w_i^t, w_{-i}^t; x)}_{\text{Discounted sum of future profit}}\end{aligned}$$

where $\delta_i \in (0,1)$ is i's discount factor. We refer to $\mathcal{F}_i^t(\cdot)$ as i's ***future profit function***.

We impose two assumptions on $\mathcal{F}_i^t$. First, we assume that future profit is absolutely continuous and concave in w_i^t for each w_{-i}^t (concavity here means that the marginal self-imposed cost of depletion is decreasing). This means that, for any given $(w_{-i}^t, x_{-i}) \in \mathcal{W}_{-i}\times\mathcal{X}_{-i}$, the set of possible contingency strategies is restricted to $\widehat{\mathcal{X}}_i(w_{-i}^t, x_{-i}) \subseteq \mathcal{X}_i$ such that each $\hat{x}_i \in \widehat{\mathcal{X}}_i(w_{-i}^t, x_{-i})$ renders $\mathcal{F}_i^t(w_i^t, w_{-i}^t; \hat{x}_i, x_{-i})$ absolutely continuous and concave in w_i^t. Second, we assume that i only picks *optimal* strategies within this restricted set. This means that each farmer always picks a contingency strategy such that

$$\hat{x}_i^*(w_i^t, w_{-i}^t, x_{-i}) \in \underset{\hat{x}_i\in\widehat{\mathcal{X}}_i(w_{-i}^t, x_{-i})}{\arg\max}\; \mathcal{P}_i^t(w_i^t, w_{-i}^t; \hat{x}_i, x_{-i}).$$

Assuming all players act this way, we can reduce the net-present value function in the following way:

$$\begin{aligned}\mathcal{P}_i^t(w_i^t, w_{-i}^t; \hat{x}^*) &= \mathcal{R}_i^t(w_i^t) - C_i^t(w_i^t, w_{-i}^t) + \mathcal{F}_i^t(w_i^t, w_{-i}^t; \hat{x}^*(w_i^t, w_{-i}^t)) \\ &\equiv \mathcal{P}_i^t(w_i^t, w_{-i}^t) = \mathcal{R}_i^t(w_i^t) - C_i^t(w_i^t, w_{-i}^t) + \mathcal{F}_i^t(w_i^t, w_{-i}^t).\end{aligned} \tag{3}$$

The latter equation in (3) gives us a succinct way to capture how farmers view groundwater usage in season t as impacting profits in season $\{t+1, t+2, \dots, T\}$.

3.1.5. Definition of groundwater dynamic games

Having described the model ingredients and assumptions, we can now define a *groundwater dynamic game* formally.

Definition 1. *A **groundwater (GW) dynamic game** consists of revenue functions $(\mathcal{R}_i^t)_{(i,t)\in\mathcal{N}\times\mathcal{T}}$, a global depletion function $\mathcal{G}$, local drawdown functions $(\mathcal{L}_i)_{i\in\mathcal{N}}$, energy costs $(e_i)_{i\in\mathcal{N}}$, and future profit functions $(\mathcal{F}_i^t)_{(i,t)\in\mathcal{N}\times\mathcal{T}}$, which, together, form a non-cooperative game at each time period:*

$$\{\mathcal{N}, (\mathcal{W}_i, \mathcal{P}_i^t)_{i\in\mathcal{N}}\}_{t\in\mathcal{T}}.$$

3.2. Integrating a quota policy

Since one of our datasets is from a district (the URD) where a specific quota policy exists, it is worth considering how such an environment could fit into the class of GW dynamic games. The type of quota policies considered here work as follows. Going into a quota period, the URD first estimates the total amount of groundwater usage that can be withdrawn sustainably, which determines the overall quota. Each farmer can then pump up to his/her share of that quota over the span of T-years, where $T = 5$ for the URD. If any farmer uses more (less) than his/her quota by X amount, then his/her quota is reduced (increased) by X during the *next* quota period.

To see how such a quota policy fits into a GW dynamic game, consider $t \in \{1, 2, \ldots, 5\}$. First note that, because the quota has no *immediate* impact on profits, we can relegate the strategic consequences of the quota to the future profit function. That is, the net-present value of a farmer pumping groundwater under a quota policy is given as

$$\begin{aligned} \mathcal{P}_i^t(w_i^t, w_{-i}^t) &= \mathcal{R}_i^t(w_i^t) - C_i^t(w_i^t, w_{-i}^t) + \mathcal{F}_i^t(w_i^t, w_{-i}^t \mid w_i^t, w_i^{t-1}, \ldots, w_i^1) \\ &\equiv \mathcal{R}_i^t(w_i^t) - C_i^t(w_i^t, w_{-i}^t) + \mathcal{F}_i^t(w_i^t, w_{-i}^t) \end{aligned} \tag{4}$$

where $(w_i^t, w_i^{t-1}, \ldots, w_i^1)$ added to $\mathcal{F}_i^t$ reflects the future consequences of using more/less than the quota. The second line of (4) follows by treating *past* decisions, $(w_i^{\tau-1}, \ldots, w_i^1)$, as *exogenous* variables. This means that such terms can be incorporated into $\mathcal{F}_i^t$ via the superscript t.

The main takeaway is that a quota policy can be subsumed directly by the class of GW dynamic games defined above because the assumptions underlying such games are also satisfied by environments with a quota policy. This feature shall play an important role in our subsequent analysis.

3.3. GW dynamic games can rationalize any dataset

We have taken considerable care to define our model as generally as possible and to not depend on arbitrary parameterizations. Compared with the only prior use of a revealed

preference test procedure by Banzhaf et al. (2018) in a CPR setting, we add substantial generality and realism by introducing a dynamic dimension to the problem, which in our view is central to modeling the use of groundwater.

This begs the following question: *can GW dynamic games account for real-world groundwater-usage behavior?* This question amounts to asking whether any GW dynamic game can predict the behavior we observe in the UBB and/or URD. Given the size and permisiveness of the class of GW dynamic games defined above, it is perhaps unsurprising to find that the answer is 'yes'.

We address this question by identifying a revealed preference test for GW dynamic games. This test and the results that follow form the basis of our analyses in the following section.

Dataset — Our goal is to show that GW dynamic games can explain *datasets*. A dataset concerns farmers $\mathcal{N} = \{1, 2, \ldots, N\}$ in seasons $\mathcal{T} = \{1, 2, \ldots, T\}$. For each season, each farmer's groundwater usage $(w_i^t)_{(i,t)\in\mathcal{N}\times\mathcal{T}} \in \mathbb{R}_{++}^{NT}$ is observed. Several control variables are also observed, including farmers' land-size $(l_i)_{i\in\mathcal{N}} \in \mathbb{R}_{++}^{N}$, groundwater levels at the beginning of the season $(g_i^{t,Spring})_{(i,t)\in\mathcal{N}\times\mathcal{T}} \in \mathbb{R}_{++}^{NT}$, and rainfall $(r^t)_{t\in\mathcal{T}} \in \mathbb{R}_{++}^{T}$. A ***dataset*** is the collection of such observations:

$$\mathcal{O} = \left\{w_i^t, l_i, g_i^{t,Spring}, r^t\right\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}.$$

Rationalizability — We say that a GW dynamic game can *rationalize* a dataset if the Nash equilibrium of the GW dynamic game coincides with observed behavior. We formalize this notion in the following definition.

Definition 2. *A dataset $\mathcal{O} = \{w_i^t, l_i, g_i^{t,Spring}, r^t\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ is **rationalizable** if there exists a GW dynamic game such that, for each $i \in \mathcal{N}$ and $t \in \mathcal{T}$,*

$$w_i^t \in \arg\max_{x_i\in\mathcal{W}_i} \mathcal{P}_i^t(x_i, w_{-i}^t).$$

In other words, rationalizability says that, at each observation t, each farmer i's observed groundwater usage maximizes net-present value given the groundwater usage of other farmers.

A characterization of rationalizability — It turns out that rationalizability of a dataset is characterized by a linear program, which we call the *GW program*.

Definition 3. *For a dataset $\mathcal{O}$, there exists a solution to the **GW program** if and only if there are positive numbers $\{\hat{f}_i^t, \hat{\lambda}_i^t\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ that satisfy the following three conditions:*

(CGD) Common global depletion,

$$\forall i, j, t: \quad \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring} = \hat{f}_j^t - \hat{\lambda}_j^t - g_j^{t,Spring};$$

(MGD) Monotone global depletion,

$$\forall i, s, t: \quad W^s - r^s > (=)\ W^t - r^t \quad \Longrightarrow \quad \hat{f}_i^s - \hat{\lambda}_i^s - g_i^{s,Spring} > (=)\ \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring};$$

(MLD) Monotone local depletion,

$$\forall i, s, t: \quad w_{\mathcal{N}_i}^s + l_i w_i^s > (=)\ w_{\mathcal{N}_i}^t + l_i w_i^t \quad \Longrightarrow \quad \hat{\lambda}_i^s - \hat{\lambda}_i^t > (=)\ 0.$$

The logic of the GW program works as follows. First, a dataset $\mathcal{O}$ plugs into the system of linear equations above to create the GW program. Second, the GW program works by identifying positive numbers $\{\hat{f}_i^t, \hat{\lambda}_i^t\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ that solve the system of linear equations—if such coefficients exist. It turns out that the existence of a solution to this system provides the key to determining whether or not GW dynamic games can rationalize UBB and/or URD data.

The following theorem is the first theoretical contribution of this paper, which shows that the GW program is necessary and sufficient for rationalizing a dataset with a GW dynamic game.

Theorem 1. *The following sentences about $\mathcal{O}$ are equivalent:*

(A) There exists a solution to the GW program defined by $\mathcal{O}$.

(B) $\mathcal{O}$ is rationalizable by a GW dynamic game.

In other words, Theorem 1 provides a non-parametric test that works as follows: (sufficiency) if there is a solution to the GW program, then there is at least one GW dynamic game, policy-free or with a quota policy, that can rationalize a dataset; (necessity) if there does not exist a solution to the GW program, then no GW dynamic game can rationalize the dataset of interest. This means that, by virtue of being necessary and sufficient, studying the GW program is equivalent to studying whether we can account for behavior in the UBB and/or URD using a GW dynamic game.

The proof of Theorem 1 clarifies what each part of the GW program represents. (CGD) represents the assumption that farmers are optimizing profits whilst drawing from a *common* resource (hence the = sign). (MGD) is derived from the assumption that global depletion is nondecreasing in aquifer-wide groundwater withdrawals. Finally, (MLD) represents the assumption that local drawdown is nondecreasing in local groundwater-usage.

There are two somewhat surprising features of Theorem 1. First, we find that the non-parametric class of GW dynamic games is characterized by a parametric, simple, and easy-to-implement linear system of equations. Second, we find that the *same* test can be used to study policy-free datasets and quota policy datasets. This means that we have a somewhat unique opportunity to compare strategic behavior in policy-free vs. policy regions in an unbiased manner.

Rationalizing behavior in the UBB and URD – By means of Theorem 1 and the GW program, it is relatively straightforward to show that *any* dataset can be rationalized with a GW dynamic game. Consider some arbitrary $\mathscr{O}$. For all $i, j \in \mathcal{N}$ and $s, t \in \mathcal{T}$, let $\hat{\lambda}_i^s = \hat{\lambda}_i^t = \hat{\lambda}_j^t = \lambda^*$—this means that (MLD) is always satisfied. For all $i \in \mathcal{N}$ and $t \in \mathcal{T}$, let $\hat{f}_i^t = g_i^{t,Spring}$. Consequently, $\hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring} = \lambda^*$ for every $i \in \mathcal{N}$ and $t \in \mathcal{T}$, which means that (CGD) and (MGD) are always satisfied. Taken together, a solution to the GW program always exists for any dataset. It thus follows from Theorem 1 that both the UBB and URD datasets are rationalizable via a GW dynamic game. This observation is perhaps unsurprising as it reflects the many folk theorems that exist for these types of games (Friedman, 1971; Aumann and Shapley, 1994; Dutta, 1995).

We summarize this observation in the following corrollary.

Corollary 1. *Any dataset $\mathscr{O}$ is rationalizable by a GW dynamic game.*

Although GW dynamic games as stated above do not have any testable predictions, there are two advantages of using this framework as a starting point. First, we assure ourselves of studying a sufficiently general class of models that can explain UBB and URD behavior. Second, this generality helps clarify the role of any additional assumptions we impose on GW dynamic games below for our purpose of testing strategic substitutes versus complements.

4. Revealed preference test of strategic behavior

In this section, we investigate whether farmers behave in a way that standard tragedy-of-the-commons theory would anticipate, and, if not, what type of behavior describes groundwater-usage.

We now turn to our groundwater data to identify what kinds of game-theoretic models fit the observed groundwater usages. In particular, our focus is on identifying whether games exhibiting strategic substitutability or strategic complementarity are able to rationalize our data. We will formulate two null hypotheses to represent these two classes of models.

Our first null hypothesis builds on standard CPR game theory, where one of the main takeaways is the strategic behavior implicated by the mismatch between individual and collective interests. It is in the farmers' interest to sustain groundwater. But if a farmer observes that neighbors are pumping at a sustainable rate, then this farmers has the individual incentive to over-pump—or pump unsustainably—in order to maximize profits. In other words, sustainable pumping incentivizes an individual to do the exact opposite—hence the mismatch. This is called *strategic substitutes.*

Our second hypothesis draws from several which assumes that farmers' individual interest is mismatched with farmers' collective interest. This assumption suggests that sustainable behavior incentivizes *un*sustainable behavior—or groundwater pumping decisions are strategic substitutes. The second hypothesis asserts the opposite, namely strategic complements. See the introduction for a long list of game theory, behavioral game theory, and empirical studies that put forward different mechanisms that lead to strategic complements.

Our second null hypothesis draws from several strands of literature that explores mechanisms whereby farmers' decisions are not strategic substitutes, but *strategic complements*. These mechanisms are from game theory (e.g. Green and Porter, 1984), behavioral game theory (e.g. Fischbacher et al., 2001), and empirical work (e.g., Ostrom, 1990)—see the introduction for a more extensive list. These mechanisms all share the common feature that sustainable pumping incentivizes sustainable pumping, i.e., decisions are strategic complements.

In sum, our hypotheses are as follows:

First null hypothesis, $\mathscr{H}_0^{SUBS}$: Farmers' groundwater-usage decisions are strategic substitutes.

Second null hypothesis, $\mathscr{H}_0^{COMP}$: Farmers' groundwater-usage decisions are strategic complements.

Alternative hypothesis, $\mathscr{H}_A$: Farmers' groundwater-usage decisions are neither strategic substitutes nor complements.

In this section, we pose our null hypotheses as revealed preference tests. We do so in two steps. First, we incorporate strategic substitutes and complements into the class of GW dynamic games defined above. Second, we identify linear programs that are necessary and sufficient for checking whether there exists *any* GW dynamic game with strategic substitutes/complements that can rationalize our data. We conclude by presenting empirical results. In the end, our main finding is a resounding rejection of our first null hypothesis—i.e.

a rejection of what standard CPR game theory would predict—and considerable evidence in favor or reciprocity-like behavior as a means of explaining groundwater-usage in the UBB and URD.

4.1. Strategic substitutes vs. complements

In order to pose our null hypotheses as revealed preference tests, we first must clarify what strategic substitutes and complements means. Following Bulow et al. (1985), strategic substitutes and complements are defined according to marginal cost. We say i's and $\mathcal{N}_i$'s pumping decisions are strategic substitutes (complements) if a decrease in $w^t_{\mathcal{N}_i}$ *increases (decreases)* i's marginal cost of pumping. The comparative statics consequence is that, if $\mathcal{N}_i$ decreases groundwater-usage, then i has the incentive to free-ride (reciprocate) by increasing (decreasing) groundwater-usage (we formalize this comparative statics observation in fn. 17).

Equivalently, strategic substitutes (complements) can be defined in terms of ***submodularity*** (***supermodularity***). We say that $\mathcal{P}^t_i$ is submodular (supermodular) in $(w^t_i, w^t_{\mathcal{N}_i})$ if *ceteris paribus* $\overline{w}^t_{\mathcal{N}_i} > \underline{w}^t_{\mathcal{N}_i}$ implies that, for each $\overline{w}^t_i > \underline{w}^t_i$,

$$\overbrace{\mathcal{P}^t_i(\overline{w}^t_i, \overline{w}^t_{\mathcal{N}_i}) - \mathcal{P}^t_i(\underline{w}^t_i,\ \overline{w}^t_{\mathcal{N}_i})}^{\text{Marginal gain/cost of pumping more GW at } \overline{w}^t_{\mathcal{N}_i}} \quad \leq (\geq) \quad \overbrace{\mathcal{P}^t_i(\overline{w}^t_i, \underline{w}^t_{\mathcal{N}_i}) - \mathcal{P}^t_i(\underline{w}^t_i, \underline{w}^t_{\mathcal{N}_i})}^{\text{Marginal gain/cost of pumping more GW at } \underline{w}^t_{\mathcal{N}_i}}.$$

Submodularity (supermodularity) means that decreasing $w^t_{\mathcal{N}_i}$ increases (decreases) i's marginal cost of pumping.[17]

We incorporate strategic substitutes and complements in GW dynamic games in the following way.

Definition 4. *A tuple* $\{\mathcal{N}, (\mathcal{W}_i, \mathcal{P}^t_i)_{i\in\mathcal{N}}\}_{t\in\mathcal{T}}$ *is a* ***GW dynamic game with strategic substitutes (complements)*** *if, for every* $i \in \mathcal{N}$ *and* $t \in \mathcal{T}$*:*

(i) $\mathcal{R}^t_i(w^t_i) \equiv \mathcal{R}_i(w^t_i)$ *and* $\mathcal{F}^t_i(w^t_i, w^t_{-i}) \equiv \mathcal{F}_i(w^t_i, w^t_{\mathcal{N}_i})$*;*

(ii) $\mathcal{P}^t_i$ *is submodular (supermodular) in* $(w^t_i, w^t_{\mathcal{N}_i})$*.*

As discussed above, the second assumption distinguishes between strategic substitutes and complements.

[17] If $\mathcal{P}^t_i$ is submodular (supermodular) in $(w^t_i, w^t_{\mathcal{N}_i})$, it follows from (e.g.) Quah and Strulovici (2009, Proposition 2) that $\overline{w}^t_{\mathcal{N}_i} > \underline{w}^t_{\mathcal{N}_i} \implies \arg\max_{x_i\in\mathcal{W}_i} \mathcal{P}^t_i(x_i, \overline{w}^t_{\mathcal{N}_i})$ is less (greater) than $\arg\max_{x_i\in\mathcal{W}_i} \mathcal{P}^t_i(x_i, \underline{w}^t_{\mathcal{N}_i})$ by the strong set order (for more details see, e.g., Topkis, 1978, and Milgrom and Shannon, 1994).

The first assumption is needed to derive meaningful revealed preference tests. The first part of this assumption asserts that revenue and future profits functions are the same across seasons. This is obviously a strong condition, and its consequences are clear in the final results—however, we can somewhat attenuate this condition with our test procedure described below. The second part of this assumption, $\mathcal{F}_i^t(w_i^t, w_{-i}^t) \equiv \mathcal{F}_i(w_i^t, w_{\mathcal{N}_i}^t)$, asserts that neighbors' behavior is sufficient to formulate future profits. This can be viewed as a condition that emphasizes the importance of locality for groundwater-usage decisions (and, it turns out, this condition is particularly important for deriving the revealed preference tests below).

4.2. A test based on revealed preferences

Our test of the null hypotheses is based on Definition 4. We are interested in whether strategic substitutes—$\mathscr{H}_0^{SUBS}$—and/or strategic complements—$\mathscr{H}_0^{COMP}$—can characterize real-world behavior. The test procedure is similar to Section 3.3. We find evidence in favor of (against) $\mathscr{H}_0^{SUBS}$ if, for a given dataset $\mathcal{O}$, we can (cannot) find a GW dynamic game with strategic *substitutes* that can *rationalize* $\mathcal{O}$ (see Definition 2 for the definition of 'rationalize'). We find evidence in favor of (against) $\mathscr{H}_0^{COMP}$ if, for a given dataset $\mathcal{O}$, we can (cannot) find a GW dynamic game with strategic *complements* that can *rationalize* $\mathcal{O}$.

Similar to Theorem 1, we operationalize our test procedure by means of a linear program, which builds directly on the GW program (see Definition 3).

Definition 5. *For a dataset $\mathcal{O}$, consider the following equations:*

(SUBS) Strategic substitutes,

$$\forall i, s, t: \quad w_i^s < w_i^t \text{ and } w_{\mathcal{N}_i}^s < w_{\mathcal{N}_i}^t \quad \Longrightarrow \quad (\hat{f}_i^s - \hat{\lambda}_i^s) - (\hat{f}_i^t - \hat{\lambda}_i^t) \leq 0;$$

(COMP) Strategic complements,

$$\forall i, s, t: \quad w_i^s < w_i^t \text{ and } w_{\mathcal{N}_i}^s > w_{\mathcal{N}_i}^t \quad \Longrightarrow \quad (\hat{f}_i^s - \hat{\lambda}_i^s) - (\hat{f}_i^t - \hat{\lambda}_i^t) \leq 0.$$

There exists a solution to the ***GW program with strategic substitutes*** *if and only if there are nonnegative numbers $\{\hat{f}_i^t, \hat{\lambda}_i^t\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ that satisfy (CGD, MGD, MLD) and (SUBS).*

There exists a solution to the ***GW program with strategic complements*** *if and only if there are nonnegative numbers $\{\hat{f}_i^t, \hat{\lambda}_i^t\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ that satisfy (CGD, MGD, MLD) and (SUBS).*

The following theorem is the second main theoretical contribution of this paper, which shows that the GW program with strategic substitutes (complements) is a necessary and

sufficient for rationalizing a dataset with a GW dynamic game with strategic substitutes (complements).

Theorem 2. *The following sentences about $\mathcal{O}$ are equivalent:*

(C) There exists a solution to the GW program with strategic substitutes (complements) defined by $\mathcal{O}$.

(D) $\mathcal{O}$ is rationalizable by a GW dynamic game with strategic substitutes (complements).

Similar to above, Theorem 2 provides a straightforward way of testing whether farmers' behavior exhibit strategic substitutes vs. complements. This test is nonparametric, simple, and easy to implement as a system of linear equations. In addition, we can employ this test on both UBB and URD data, which provides us with an unbiased way of comparing behavior in a policy-free vs. policy region.

4.3. Test setup

In implementing the GW program, testing all UBB and URD data at the same time with a somewhat unforgiving revealed preference test is perhaps not the most informative test. We are trying to account for nearly 15,000 farmers and 100,000 decisions. If a single farmer is acting 'irrationally' or a single decision is 'irrational', then the test will fail and we loose information about how the remaining $N-1$ farmers are behaving. This is one of several issues we can overcome by taking a more nuanced testing approach.

Following Carvajal et al. (2013) and Banzhaf et al. (2018), we run the GW program on subsets of UBB and URD data, with each subset consisting of $\{2,3,4,5,6,8\}$ farmers over $\{2,3,4,5,6,8\}$ consecutive seasons. It follows from Theorem 2 that the GW program with strategic substitutes (complements) is necessary and sufficient for any subset of the data to be rationalized by a GW dynamic game with strategic substitutes (complements). We also restrict attention to $\{2,3,4,5,6,8\}$ farmers who are mutual neighbors.[18] This is because our goal is to test farmers who are likely to view each others' actions as strategically interdependent—this argument is more compelling for farmers in a one mile proximity rather than, e.g. a thirty mile proximity.

As noted above, there are several advantages to testing subsets of our data, rather than the entire dataset. First, we exclude the possibility that the GW program is rejected because

[18] Operationally, we do so by randomly selecting i, recovering his/her closest $\{1,2,3,4,5,7\}$ neighbors, and then running the GW program on these farmers.

a few farmers are not responding optimally to others. Second, running the test on subsets of the data mitigates errors that might result from unobserved variable bias. It is very likely that idiosyncratic shocks impact farmers' behavior during the season, such as the mechanical failure of a center pivot irrigator, which may render behavior as seemingly irrational. Third, we rule out the possibility of spurious rejection rates due to farmers learning and innovating to improve their farming practices between seasons. Such learning can change a farmer's revenue function and/or cost function over time, which is a process that is not captured by the class of GW dynamic games presented above. Looking at smaller samples with short time scales addresses this issue. Finally, as is made clear below, we demonstrate the ability of the test to reject behavior on a small number of observations notwithstanding the permissiveness of our non-parametric framework.

Unlike Carvajal et al. (2013) and Banzhaf et al. (2018), the size of our dataset renders it computationally impossible to sample over the 4.7 trillion farmer pairs, 3.7×10^{18} farmer quadruplets, let alone each consecutive time sequence. We thus randomly sample subsets of our data, run the GW program on each such subset, and record the percentage of cases with a solution to the corresponding GW program—we call this the *acceptance rate*. In doing so, we can leverage the law of large numbers and randomly sample until we reach a convergence in the rejection rate (in practice, running the GW program on 1,000 random sub-samples is sufficient).

4.4. Results

Upper Big Blue District — Following Hardin (1968), we might anticipate that groundwater-usage behavior are strategic substitutes in the UBB because farmers operate in an unrestricted and ungoverned manner. Without such restrictions, the incentives to avoid the tragedy of the commons are, in theory, absent.

We report results using the UBB dataset in Table 2. Each reported number is an *acceptance rate*, and it is interpreted as the probability of sampling a subset of farmers whose behavior is rationalizable by a GW dynamic game with (Panel A) strategic substitutes or (Panel B) strategic complements. Two patterns in Table 2 are perhaps unsurprising. First, acceptance rates are decreasing in the number of farmers. This is unsurprising because, if a dataset consisting of (e.g.) five farmers over three years passes the GW program test, then any subset of this dataset involving four or less farmers over a shorter time window will also pass the GW program test. Second, acceptance rates are decreasing with the time window. This is unsurprising because of the season-independence assumption made in the definition of GW dynamic games with strategic substitutes/complements.

Table 2: Upper Big Blue District — Acceptance rates of the GW program with (*Panel A*) strategic substitutes and (*Panel B*) strategic complements (sub-sample = 1,000).

Panel A – Strategic substitutes		*Number of farmers* ($\mathcal{N}$)					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	0.63	0.56	0.52	0.44	0.39	0.30
	$T = 3$	0.23	0.14	0.12	0.08	0.06	0.05
	$T = 4$	0.06	0.03	0.01	0.01	0	0
	$T = 5$	0.01	0.01	0	0	0	0
	$T = 6$	0	0	0	0	0	0
	$T = 8$	0	0	0	0	0	0
Panel B – Strategic complements		*Number of farmers* ($\mathcal{N}$)					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	0.90	0.87	0.84	0.82	0.75	0.74
	$T = 3$	0.73	0.70	0.61	0.57	0.51	0.4
	$T = 4$	0.66	0.50	0.40	0.35	0.28	0.24
	$T = 5$	0.52	0.33	0.24	0.19	0.15	0.20
	$T = 6$	0.37	0.23	0.15	0.10	0.06	0.02
	$T = 8$	0.12	0.04	0.01	0	0	0

The results in Table 2 point to strong evidence against strategic substitutes and in favor of strategic complements. We find that the GW program with strategic substitutes exhibit low acceptance rates: as soon as the number of farmers exceeds four (which is less than 0.05% of the UBB population) and for sufficiently long time windows, the data systematically reject the possibility of finding a GW dynamic game with strategic substitutes that can rationalize behavior in the UBB. This is in stark contrast to the strong pattern of high acceptance rates we find with the GW program with strategic complements. For example, for $T = 2$ and $N = 8$, we find that 44% more data can be explained with GW dynamic games with strategic complements than with strategic substitutes.

Table 2 reports *absolute* acceptance rates. It turns out that looking at our results from the perspective of *relative* acceptance rates further highlights the explanatory power of GW dynamic games with strategic complements. We explain the relative measure below.

Selten and Krischker (1983) introduced a *measure for predictive success* that accounts for prediction accuracy *and* precision. As articulated by Selten (1991), a class of games is *accurate* if it can predict real-world behavior successfully—this is what Table 2 reports.

Table 3: Upper Big Blue District — Selten and Krischker's (1983) measure of predictive success of the GW program with (*Panel A*) strategic substitutes and (*Panel B*) strategic complements (sub-sample = 1,000).

Panel A – Strategic substitutes		*Number of farmers* ($\mathcal{N}$)					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	−0.17	−0.15	−0.15	−0.18	−0.18	−0.14
	$T = 3$	−0.3	−0.26	−0.17	−0.19	−0.12	−0.04
	$T = 4$	−0.26	−0.15	−0.08	−0.05	−0.05	−0.02
	$T = 5$	−0.13	−0.05	−0.03	−0.01	−0.01	0
	$T = 6$	−0.06	−0.03	−0.01	0	0	0
	$T = 8$	−0.01	0	0	0	0	0
Panel B – Strategic complements		*Number of farmers* ($\mathcal{N}$)					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	0.12	0.25	0.29	0.36	0.38	0.49
	$T = 3$	0.33	0.44	0.46	0.48	0.45	0.39
	$T = 4$	0.46	0.42	0.37	0.34	0.28	0.24
	$T = 5$	0.45	0.31	0.24	0.19	0.15	0.1
	$T = 6$	0.34	0.22	0.15	0.1	0.06	0.02
	$T = 8$	0.12	0.04	0.01	0	0	0

However, Selten (1991) argued that accuracy should be understood with respect to *precision*, namely the relative size of the set of predicted outcomes. The higher the precision, the smaller the set of predicted outcomes. If two classes of games are equally accurate, if the first is more precise than the second, then the first should be preferred.

Selten and Krischker (1983) proposed a simple, straightforward way of comparing classes of games that accounts for accuracy and prediction. If A is the accuracy and P is the precision, then Selten and Krischker suggest that classes of games should be compared via its accuracy minus the precision, or

$$m(A, P) = A - P \tag{5}$$

(see Selten, 1991, who axiomized $m(A, P)$ up to positive affine transformations and Beatty and Crawford, 2011, for an alternative proof). If $m(A, P) = 1 - 0 = 1$, then observed behavior is the unique prediction. If $m(A, P) = 1 - 1 = 0$, then the prediction's accuracy is equally offset by its lack of precision. For example, GW dynamic games from Corollary 1 have a predictive success measure of $m(A, P) = 1 - 1 = 0$ because it is perfectly accurate ($A = 1$) by

virtue of being able to predict any behavior ($P = 1$).[19]

In our case, A represents the acceptance rate reported in Table 2. We estimate the precision associated with GW dynamic games by assuming that $\mathcal{W}_i = [0, 40\text{in}]$ $\forall i \in \mathcal{N}$ (40in is the most groundwater-usage observed in our datasets). We randomly sample $(w_i^t)_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ and farmers, and then run the GW program with strategic substitutes/complements; P is then the acceptance rate using this randomly sampled data.

We report Selten and Krischker's (1983) measure of predictive success in Table 3. As is visually apparent, the evidence shows that GW dynamic games with strategic complements is more successful at predicting behavior than its strategic substitutes analog. With the GW program with strategic substitutes, we find that $m(A, P) \leq 0$ for each (N, T) case; this means that the theory is systematically predicting the wrong region of $\times_{i\in\mathcal{N}}\mathcal{W}_i$. With the GW program with strategic complements, we find that $m(A, P) \geq 0$ for each (N, T) case (we find that $m(A, P) \geq 0.12$ for $T \leq 3$). Taken together, the evidence suggests that GW dynamic games with strategic complements is a more accurate and precise theory for explaining UBB behavior than GW dynamic games with strategic substitutes.

Upper Republican District — We report results using the URD dataset in Tables C.8 and C.9, where a similar pattern emerges as above. As shown in Table C.8, we find that GW dynamic games with strategic complements explains more data than its strategic substitutes analog. However, from the perspective of Table C.8 alone, this difference is not as stark as the UBB case. While strategic complements rationalizes $25-44\%$ more UBB data than strategic substitutes (restricting attention to $T = 2, 3$), we find that only 7–16% more URD data is explained by strategic complements than strategic substitutes.

We report the measure of predictive success with URD data in Table C.9. This measure sheds more light on the predictive success strategic complements. Similar to the UBB, for the GW program with strategic substitutes, we find that $m(A, P) \leq 0$ for each (N, T) case; this means that the theory systematically predicts the wrong region of $\times_{i\in\mathcal{N}}\mathcal{W}_i$. For the GW program with strategic complements, we find that $m(A, P) \geq 0$ for each (N, T) case (we find that $m(A, P) \geq 0.10$ for $T \leq 3$). From the perspective of predictive success, we again find that GW dynamic games with strategic complements is a more accurate and precise theory for explaining URD behavior than its strategic substitutes analog.

[19] Beatty and Crawford (2011) were the first to propose this measure as a way of shedding light on the explanatory power of revealed preference tests.

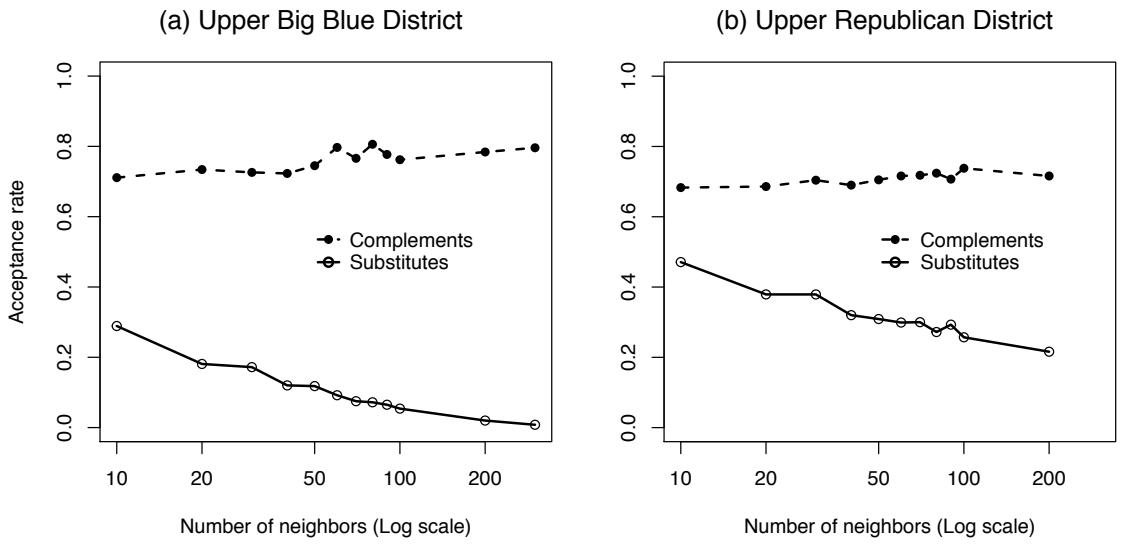


Figure 3: Acceptance rate of the GW program with strategic substitutes/complements with $T = 2$ as a function of N.

4.5. Robustness check I: Larger N

A follow-up question is asking what acceptance rates look like for $N > 8$. Considering up to eight neighbors indeed demonstrates that strategic complements is a better predictor of behavior than strategic substitutes and that our revealed preference tests can reject even small sample sizes. *But does this pattern continue to hold for larger N?* If not, then trying to extrapolate our findings to UBB- and URD-wide behavior might be subject to concern. We address this question by estimating acceptance rates for the GW program with strategic substitutes/complements for $T = 2$ and $N = \{10, 20, \ldots, 100, 200, 300\}$ (we stopped at 300 because of computational limitations) for both the UBB and URD. The question here amounts to asking whether the long-term trend of acceptance rates go up or down for the strategic substitutes/complements models—if both acceptance rates go down, then our results might not map to UBB- and URD-wide behavior more generally.

We report our findings in Figure 3. These results reinforce the findings in Tables 2–C.9: especially for large N, strategic complements is a better predictor of behavior than strategic substitutes. In both the UBB and URD, acceptance rates for the GW program with strategic complements *increases* with increasing N. The opposite is true for the strategic substitutes, where acceptance rates *decrease* with increasing N. The difference in explanatory power is most obvious in the UBB for $N = 300$: in this case, GW dynamic games with strategic complements explains nearly 80% more data than GW dynamic games with strategic

substitutes.

4.6. Robustness check II: Network structure

In Appendix C.1, we test the robustness of our results in two ways.

The first concerns network structure. Above, we study farmers' revealed preference on a *single* network structure based on local drawdown. We argued that this is a reasonable starting point, as local drawdown is a well-known and widely discussed dynamic among farmers. But it is unclear whether the local drawdown network is the true, or even appropriate network to be studying strategic interactions, which begs the question: *are our revealed preference test results robust to different network structures*?

The second concerns viewing our revealed preferences results at the population vs. individual level. We base conclusions on strategic substitutes vs. complements for the entire URD and URD population on a handful of numbers. This begs the question: *are all farmers' pumping decisions explained by strategic complements, or are decisions better characterized by a mix of substitutes and complements?* In other words, we are interested in disaggregating the results above from the population level to the individual level.

We address both robustness questions simultaneously in Appendix C.1. In brief, we can define an acceptance rate per farmer by randomly sampling network structures and testing each farmer's passing rate of the various GW programs individually (see the appendix for more details). We thus recover a distribution of acceptance rates, rather than a single value for the entire UBB/URD.

We plot the results in Figure C.5, which support the results above in three ways. First, we find that each farmer's individual-level acceptance rate of the GW program with strategic complements is higher than the strategic substitutes counterpart. This pattern is especially true for the UBB, where strategic complements explains behavior 50% to 90% more often than strategic substitutes. Second, the farmer-level acceptance rates cluster around those reported above (the clustering is stronger for the UBB than the URD; see the figure). This means that the group-level statistics we used do a reasonable job at representing farmer-level behavior. Finally, the evidence suggests that strategic complements does a better job at explaining behavior in the UBB than the URD. In the former, farmer-level acceptance rates cluster around 0.7 and 0.1 for strategic complements and substitutes, respectively. In the latter, farmer-level acceptance rates cluster around 0.7 and 0.35 for strategic complements and substitutes, respectively.

5. Estimating strategic interactions

In this section, we take another approach to test whether strategic behavior can be explained in terms of strategic substitutes or strategic complements. The test we used above had the advantage of being nonparametric. However, the test was restricted to the classes of dynamic games defined above, which means that we could only test for/against strategic substitutes and complements that emerged from such game-theoretic assumptions. This would exclude the possibility of testing for more behavioral drivers of groundwater-usage, such as interactive preferences (Levine, 1998; Nax et al., 2015) and social norms (Ostrom, 1990). The goal of this section is to test behavior in way that explicitly avoids game-theoretic assumptions by, instead, making parametric (i.e. regression) assumptions.

Using a non-game-theoretic and parametric approach, our goal is to address two questions. (i) *Are strategic interactions important in common-pool resource settings?* This is a motivational question that we could not address with the revealed preference approach above. But with appropriately designed regressions, we can pit strategic interactions against other drivers of behavior, such as market and climate trends. (ii) *If strategic interactions are important, does strategic complements characterize groundwater-usage?* That is, we are looking for secondary evidence that supports the findings in the previous section.

In the end, we find a resounding 'yes' to both questions: our evidence suggests that strategic complements play an equally important role, if not more, than any other exogenous driver of behavior.

This section is structured as follows. We begin by presenting our regression framework and clarifying what substitutes and complements mean in this context. Following, we discuss our estimation strategy. We then present the hypothesis test setup and results. We conclude by relating our findings here with the revealed preference test results in the previous section.

5.1. Data

In our analysis below, we proceed only with the UBB data and not with the URD data because the URD quota policy poses an empirical issue. If a farmer in the URD exceeds his/her quota, then his/her future quota will be decreaed. On the other hand, the quota policy allows farmers to 'save' unused groundwater quota for future years. As such, each URD farmer's groundwater pumping from 1982–2000 can influence can influence his/her behavior that we observe from 2001–14. But because URD data does not include decisions before 2001, we have an unobserved variable issue. Fortunately, this issue is not present in the UBB data.

5.2. Regression setup

Strategic network — The first step in estimating strategic interactions is clarifying *with whom* a farmer strategically interacts. To do so, the longitude-latitude information in the dataset comes into play. As in Section 3, let $\mathcal{N}_i \subseteq \mathcal{N}\backslash\{i\}$ be farmer i's neighbors. We combine neighborhoods to define a ***strategic network*** $\boldsymbol{A} \in \mathbb{R}_+^{N\times N}$ as a N-by-N matrix where each row $\boldsymbol{A}_i = (\boldsymbol{A}_{ij})_{j\in\mathcal{N}}$ represents farmer i's neighborhood such that $\boldsymbol{A}_{ij} > 0$ if and only if $j \in \mathcal{N}_i$, and $\boldsymbol{A}_{ij} = 0$ otherwise. The matrix is row-normalized.

Regression model — Suppose that the strategic network, $\boldsymbol{A}$, is known *a priori* (we relax this assumption later on). We assume that farmer i's behavior in season $t \in \mathcal{T}$ is explained by[20]

$$w_i^t = \alpha_i + \beta_{\mathcal{N}} \sum_{j\in\mathcal{N}} \boldsymbol{A}_{ij} w_j^t + \boldsymbol{x}_i^t \boldsymbol{\psi} + \boldsymbol{y}^t \boldsymbol{\phi} + \boldsymbol{d}_i \boldsymbol{\gamma} + \tilde{u}_i^t. \tag{6}$$

The second term, $\beta_{\mathcal{N}} \sum_{j\in\mathcal{N}} \boldsymbol{A}_{ij} w_j^t$, corresponds to i's behavior as determined by strategic interactions. $\boldsymbol{x}_i^t$ denotes a $1 \times K$ vector of individual-level time-varying exogenous controls included in the dataset—including groundwater levels and interaction terms, such as $\text{rain}^t \times \text{land-size}_i$—and $\boldsymbol{\psi}$ is the corresponding $K \times 1$ parameter vector. $\boldsymbol{y}^t$ denotes a $1 \times L$ vector of UBB-wide time-varying exogenous controls included in the dataset—including crop prices, electricity prices, rain, temperature, and land rental rates—and $\boldsymbol{\phi}$ is the corresponding $L \times 1$ parameter vector. $\boldsymbol{d}_i$ is a $1 \times R$ vector of farmer-level time-invariant controls—including land-size, well depth, transmmissivity, and soil type—and $\boldsymbol{\gamma}$ is its corresponding $R \times 1$ parameter vector. α_i is a farmer specific fixed-effect constant. Because we are interested in $\beta_{\mathcal{N}}$, we can use a first-difference estimator, which means that we estimate

$$w_i^t - w_i^{t-1} = \beta_{\mathcal{N}} \sum_{j\in\mathcal{N}} \boldsymbol{A}_{ij}(w_j^t - w_j^{t-1}) + (\boldsymbol{x}_i^t - \boldsymbol{x}_i^{t-1})\boldsymbol{\psi} + (\boldsymbol{y}^t - \boldsymbol{y}^{t-1})\boldsymbol{\phi} + u_i^t, \tag{7}$$

thereby avoid estimating $(\alpha_i)_{i\in\mathcal{N}}$ and $\boldsymbol{\gamma}$.

The final term, $\tilde{u}_i^t$, captures overall disturbance and makes it possible to control for unobserved and possibly confounding factors. Perhaps most importantly, we must control for disturbances that can occur *spatially*; e.g., hail storms can damage a localized region of the UBB, which is unobserved in the dataset but may appear to take the form of positive/negative strategic interactions between farmers. We can model such unobserved events by exploiting

[20]When implementing this regression, we work with $\log w_i^t$ rather than w_i^t since $w_i^t < 0$ is not possible. We remove log to keep notation lighter.

the fact that it changes behavior in a spatially tractable way.

We proceed with a commonly made assumption that such disturbances are related via a first-order spatial autocorrelation expression, which is given as

$$u_i^t = \rho \sum_{j \in \mathcal{N}} \boldsymbol{B}_{ij} u_j^t + \nu_i^t. \tag{8}$$

where ρ is the degree of spatial correlation, $\boldsymbol{B} \in \mathbb{R}_+^{N \times N}$ is a weighting matrix that governs the spatial relation between i and j ($\boldsymbol{B}$ may be different than $\boldsymbol{A}$), and ν_i^t captures the individual-time innovations. The expression in (8) is also known as a Cliff and Ord (1973a,b) process (see also Kapoor et al., 2007, Baltagi, 2008, and Mutl and Pfaffermayr, 2011). The interpretation of (8) is that, if j receives an unobserved shock u_j^t, then his/her neighbors defined by $\boldsymbol{B}$ experience a shock ρu_j^t.

Our goal is to estimate (7,8).

Hypotheses — Based on the regression setup above, our null and alternative hypotheses are given as follows.

> *Null hypothesis, $\mathscr{H}_0'$: Framers' groundwater-usage decisions are strategic substitutes or do not exhibit any systematic relations.*
>
> *Alternative hypothesis, $\mathscr{H}_A'$: Framers' groundwater-usage decisions are strategic complements.*

The null hypothesis here is slightly different than that considered in the previous section. Here, we give the benefit of the doubt to strategic substitutes by allowing for the possibility that, in our regression framework, strategic substitutes are not identifiable.

How do $\mathscr{H}_0'$ and $\mathscr{H}_A'$ translate to the regression setup above? Supposing we estimate (7,8) robustly, we can utilize $\beta_{\mathcal{N}} \pm 2\text{SE}$ to delineate between our null and alternative hypotheses. For an estimated $\beta_{\mathcal{N}} \pm 2\text{SE}$, $\mathscr{H}_0'$ holds if and only if $\beta_{\mathcal{N}} - 2\text{SE} \leq 0$ (otherwise, $\beta_{\mathcal{N}} - 2\text{SE} > 0 \iff \mathscr{H}_A'$). Before discussing the p-value associated with testing $\mathscr{H}_0'$ vs. $\mathscr{H}_A'$, we must first discuss how we estimate (7,8).

5.3. Identification

As we briefly noted in the introduction, such a regression model cannot be estimated without somehow addressing issues related to *endogeneity*. In particular, if $water_j^t$ is used to explain $water_i^t$, which is used to explain $water_j^t$, then there exists endogeneity that may

bias the estimation of coefficients in uninterpretable ways—as j influences him/herself by virtue of influencing i. This identification issue is known as the *reflection problem* (Manski, 1993; Moffit, 2001), and resolving it requires a suitably selected set of instrumental variables (IVs).[21]

We address this issue by utilizing an IV setup proposed by Mutl and Pfaffermayr (2011) for fixed-effect regression designs (which builds on Kapoor et al., 2007 and Baltagi, 2008). Specifcially, Mutl and Pfaffermayr (2011) assume that, from any farmer i's perspective, his/her neighbors' control variables do not influence w_i^t *directly*, but only indirectly by virtue of influencing $\sum_{j\in\mathcal{N}} \boldsymbol{A}_{ij} w_j^t$ *directly*. This means that $\boldsymbol{A}$ and i's neighbors' control variables can be used as the basis of IVs. Mutl and Pfaffermayr (2011) go on to show that $\boldsymbol{B}$ can be used to render such IVs uncorrelated with the error terms, which means that $\beta_{\mathcal{N}}$ is identified in our setting.

Following Mutl and Pfaffermayr (2011), $\boldsymbol{A}$, $\boldsymbol{B}$ and the control variables in the dataset can be organized to form IVs for estimating (7,8). We refer the interested reader to Appendix B for a derivation of the IVs as well as a discussion on the assumptions required to estimate (7,8) robustly.

5.4. Testing procedure

The estimator procedure proposed by Mutl and Pfaffermayr (2011) requires an assumption that is somewhat problematic in our setting, namely *that matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are observed.* Our setting has a different setup. The UBB dataset has information about the geographical location of farmers, and it is possible to leverage this information to define a network that represents interactions based on local drawdown. Yet, while we have reason to believe that defining 'local' interaction networks based on this information is a reasonable starting point, such a network may not necessarily represent the true strategic interaction networks describing reality.

It turns out that resolving this issue also allows us to derive a p-value for testing our hypotheses. We proceed as follows. We define a set of networks $\mathscr{N}$ large enough to encompass at least one $\boldsymbol{A}^*, \boldsymbol{B}^* \in \mathscr{N}$ that represent the true network of strategic interactions and spatially correlated disturbances, respectively. We then randomly sample from this set of networks and estimate (6,8), which means that we recover $\beta_{\mathcal{N}} \pm 2\text{SE}$ from each randomly

[21] See Blume et al. (2010) for an overview of the reflection problems and various approaches to resolving it. In spatial/network games à la Jackson (2008), using IVs is one of the most systematically explored means of resolving endogeneity; see, e.g., Bramoulle et al. (2009). We utilize a technique proposed by Mutl and Pfaffermayr (2011) because it allows us to incorporate spatially correlated errors.

sampled network. Perhaps somewhat surprisingly, it is shown below that the 'sampling' of $\beta_{\mathcal{N}}$s in this way follows a normal distribution (statistically speaking). This allows us to calculate the probability of identifying a network such that $\beta_{\mathcal{N}} - 2\text{SE} \leq 0$, which is given as

$$P\left[\boldsymbol{A}, \boldsymbol{B} \in \mathcal{N} : \beta_{\mathcal{N}} - 2\text{SE} \leq 0\right]. \tag{9}$$

Recall that, for an estimated $\beta_{\mathcal{N}} \pm 2\text{SE}$, $\mathscr{H}_0'$ holds if and only if $\beta_{\mathcal{N}} - 2\text{SE} \leq 0$. This means that (9)—which is the probability of sampling such that $\mathscr{H}_0'$ is true—serves as a p-value for testing $\mathscr{H}_0'$ versus $\mathscr{H}_A'$.

This testing procedure is valid only if the *true* networks $\boldsymbol{A}^*$ and $\boldsymbol{B}^*$ are inside of $\mathcal{N}$. More formally, we say that $\mathcal{N} \subseteq \mathbb{R}_+^{N\times N}$ is the ***set of feasible networks*** if it contains every $\boldsymbol{A} \in \mathbb{R}_+^{N\times N}$ such that: (i) $\boldsymbol{A}_{ii} = 0$, or no farmer strategically affects him/herself, (ii) $\boldsymbol{A}_{ij} > 0$ only if i and j are less than 5 miles apart, depending on the case, (that is, only neighbors induce strategic interaction effects), (iii) $\boldsymbol{A}_{ij} = \boldsymbol{A}_{ik}$ for any j and k with non-zero weights, or neighbors strategically affect i equally, and (iv) $\boldsymbol{A}$ is row normalized. Our testing procedure requires the following assumption.

Assumption 1. *Let $\mathcal{N}$ represent the set of feasible networks. Then there exists at least one pair of networks $\boldsymbol{A}^*, \boldsymbol{B}^* \in \mathcal{N}$ that represents the true network of strategic interactions and spatially correlated disturbances in the UBB, respectively.*

We assume that neighbors come from less than 5 miles for two reasons. First, we incorporate the 'local drawdown network', which characterizes interactions that occur between farmers less than 1 mile. Second, we wish to extend the network space in order to incorporate other possible avenues for strategic interactions. It could be that the true network of interactions takes place on a larger scale than one mile if, e.g., mechanisms such as social pressure are at play—this is especially true taking into account that some farmers live in towns that are several miles away from their farm land. Therefore, we increase the scope of possible networked strategic interactions to increase the scope of possible mechanisms that could be tested.

In Appendix C, we run the same statistical test supposing that networked strategic interactions take place between farmers less than $\{1, 3, 10, 20\}$ miles away as a robustness check. We find that the results are nearly the same as those presented below.

5.5. Results

We randomly sample 10,000 pairs of $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{N}$ and estimate (7,8) for each pair accordingly. This allows us to estimate (9), namely the probability of sampling a network that

Table 4: Micro-level regressions on UBB data. Number of observations = 97,160. Number of randomly sampled regressions = 10,000. Regressions are combined using the Rubin's rule (Rubin, 1987), and p-values are calculatd using the Barnard-Rubin corrected degrees of freedom (Barnard and Rubin, 1999).

	Dependent variable: Log ($Water_i^t$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Strategic interaction effects						
Log(Neighbors' $water_i^t$)	0.956^{***} (0.021)	0.390^{***} (0.027)	0.307 (0.352)	0.237^{***} (0.032)	0.948^{***} (0.022)	0.395^{***} (0.027)
Groundwater controls						
Spring GW_i^t	-0.03^{***} (0.006)	-0.051^{***} (0.006)			-0.034^{***} (0.006)	-0.057^{***} (0.006)
Spring GW_i^t − Fall GW_i^{t-1}	0.019^{***} (0.003)	0.044^{***} (0.003)			0.023^{***} (0.003)	0.051^{***} (0.003)
UBB-level time-dependent controls						
Spot market $price^t$		0.018^{***} (0.003)		0.034^{***} (0.003)		0.016^{***} (0.003)
$Electricity^t$		-0.115^{***} (0.009)		-0.155^{***} (0.01)		-0.112^{***} (0.009)
Land rental $rates_{C_i}^t$		0.107^{***} (0.006)		0.126^{***} (0.006)		0.107^{***} (0.005)
$Rain^t$		-0.164^{***} (0.008)		-0.199^{***} (0.009)		-0.163^{***} (0.008)
$Temperature^t$		-0.105^{***} (0.007)		-0.145^{***} (0.009)		-0.102^{***} (0.007)
(Temp.)×$(Rain)^t$		-0.055^{***} (0.005)		-0.061^{***} (0.005)		-0.055^{***} (0.005)
Farmer- and time-dependent controls						
(Land-size)×$(Rain)_i^t$			-0.017^{***} (0.002)	-0.016^{***} (0.002)	-0.017^{***} (0.002)	-0.017^{***} (0.002)
(Land-size)×$(Rain)_i^t$			-0.012^{***} (0.002)	-0.012^{***} (0.002)	-0.012^{***} (0.002)	-0.012^{***} (0.002)
(Well depth)×$(Rain)_i^t$			-0.007^{***} (0.002)	-0.007^{***} (0.002)	-0.008^{***} (0.002)	-0.01^{***} (0.002)
(Well depth)×$(Temp.)_i^t$			0.005^{**} (0.002)	0.006^{**} (0.002)	0.005^{**} (0.002)	0.005^{**} (0.002)
(Land-size)×(Rain)×$(Temp.)_i^t$			-0.016^{***} (0.002)	-0.015^{***} (0.002)	-0.016^{***} (0.002)	-0.016^{***} (0.002)
(Well depth)×(Rain)×$(Temp.)_i^t$			-0.014^{***} (0.002)	-0.014^{***} (0.002)	-0.016^{***} (0.002)	-0.018^{***} (0.002)

Note: *p<0.05; **p<0.01; ***p<0.001

supports our null hypothesis that farmers' decisions are strategic substitutes. This is our main finding and is presented in Section 5.5.2.

But before we can discuss these results, it is crucial to understand whether we have 'good' IVs. Mutl and Pfaffermayr (2011) prove that the IVs—which we derive in Appendix B—are orthogonal to the disturbances terms. While this is important, the IVs must still be sufficiently correlated with $w^t_{\mathcal{N}_i}$ for the purposes of identification. A number of issues can arise if we do not have 'good' IVs: if we have weak IVs or if the results are sensitive to which IVs are used, then any findings we report are potentially moot. We thus show evidence that our IVs are valid before discussing the main findings.

5.5.1. Robustness of IVs

One way to check IVs is testing whether results are *robust* to different IVs. With a number of different IVs, if each report a different strategic interaction term $\beta_{\mathcal{N}}$, then we are left with a problem because it suggests that our IVs are either (i) weak or (ii) unable to identify $w^t_{\mathcal{N}_i}$ robustly. Yet, if different IVs each yield roughly the same $\beta_{\mathcal{N}}$, then we have good reason to believe that our IVs are indeed identifying strategic interactions.

We take this approach as a robustness check of our IV estimation procedure. We estimate (7,8) using six different IV setups, and we report the findings in Table 4. The first two focus on using groundwater levels and UBB-wide variables, such as rainfall and crop prices. The second two models focus on using individual-level and UBB-wide variables, such as (land-size$_i$)×(rainfall^t). The fifth model utilizes a mix of IVs from the previous four. The final model utilizes all control variables included in our UBB dataset. To estimate parameters, we randomly sample 10,000 networks $\boldsymbol{A}, \boldsymbol{B} \in \mathscr{N}$ and combine the regression estimations using Rubin's rule (Rubin, 1987). This means that: (i) the coefficients are combined using a simple average, (ii) standard errors are combined in a way that accounts for between- and within-variance of the estimated coefficients, and (iii) p-values are computed using the Barnard-Rubin corrected degrees of freedom (Barnard and Rubin, 1999). All coefficients are normalized, so that, for all the control variables, β =% increase in groundwater usage per +1SD increase in the variable.

Two patterns in Table 4 give positive evidence that our UBB control variables are able to identify strategic interaction effects. First, in models (2,4,6) where we include market and climate trends, we recover roughly the same $\beta_{\mathcal{N}}$. Interestingly, model (3) shows that interacting farmer-level controls with climate variable also results in the same $\beta_{\mathcal{N}}$ as models (2,4,6) but with very high SEs. The interpretation is that the IVs in model (3) were consistent but very weak. Second, it is obvious when IVs are weak/fail in Table 4. Models (1) and

Test of negative ($\mathscr{H}'_0$) vs. positive ($\mathscr{H}'_A$) strategic interaction effects

	Hypothesis test results for $\mathscr{H}'_0$ vs. $\mathscr{H}'_A$		
	$\mathcal{N}(\mu, \sigma^2)$	KS p-value	p-value for $\mathscr{H}'_0$ vs. $\mathscr{H}'_A$
Neighs $\leq$ 5 miles	(0.351, 0.015)	0.69	<0.0001

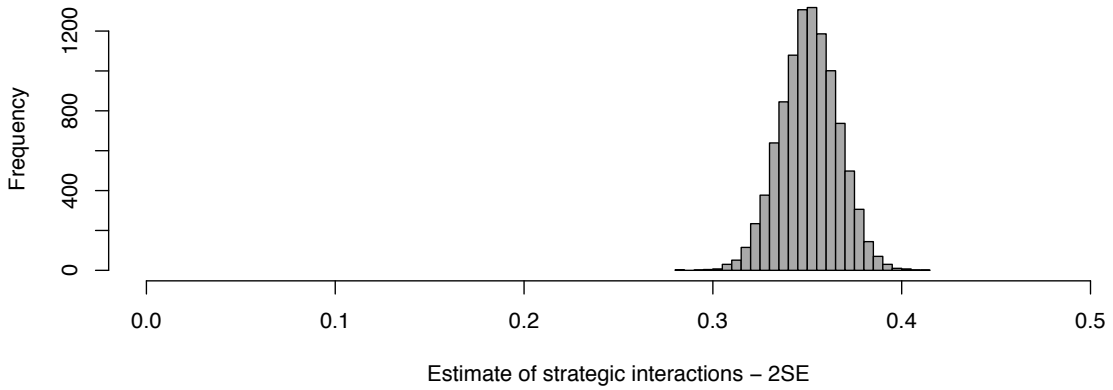


Figure 4: *(Top Table)* Results for KS-test of the distribution of (second column) $\beta_{\mathcal{N}} - 2\text{SE}$ and (third column) $\mathscr{H}'_0$ vs. $\mathscr{H}'_A$. *(Bottom Figure)* $\beta_{\mathcal{N}}$ for randomly sampled network structures. Interpretation of $\beta_{\mathcal{N}}$: $+1w^t_{\mathcal{N}_i}$ corresponds to $+(100\% \times \beta_{\mathcal{N}})$ increase in w^t_i.

(5) show that when market and climate variables are excluded, we recover estimates where $\beta_{\mathcal{N}} \approx 1$. This is problematic because spatial regression models are generally not identifiable if $\beta_{\mathcal{N}} \geq 1$, and recovering $\beta_{\mathcal{N}} \approx 1$ suggests that identification issues are present.

The main takeaway is that our estimation procedure is robust to which IVs we use, as long as we use 'good' IVs. As such, we proceed with model (6) because it yielded the lowest SE for $\beta_{\mathcal{N}}$. Our main findings do not change if we instead use model (2) or (4).

5.5.2. Main findings

Our main findings, presented in Figure 4, are based on the 10,000 regressions that underlie model (6) in Table 4. Figure 4 has two parts. First, in the lower figure, we plot every ($\beta_{\mathcal{N}} - 2\text{SE}$) estimated for each randomly sampled $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{N}$; in practice, there are 10,000 estimates. The visual evidence is unambiguous: we find that every estimated $\beta_{\mathcal{N}} - 2\text{SE} > 0.2 > 0$. This already points to strong and robust evidence against $\mathscr{H}'_0$ and in favor of strategic complements.

Our second set of findings in the upper Table formalize that which is visually apparent in the histogram below. We find that the distribution of recovered $\beta_{\mathcal{N}} - 2\text{SE}$ follow a normal distribution—we report the fitted normal distribution in the fist column, $\mathcal{N}(\mu, \sigma^2)$, and the

Kolmogorov-Smirnov test p-value in the second column ($p = 0.70$). With evidence of $\beta_{\mathcal{N}} - 2\text{SE}$ following a normal distribution, we can thus estimate the probability of drawing any $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{N}$ such that $\beta_{\mathcal{N}} - 2\text{SE} \leq 0$, which serves as the p-value for testing $\mathscr{H}_0'$ vs. $\mathscr{H}_A'$ (see equation (9)). We report this value in the third column. As is visually evident in the histogram, the probability of finding networks such that $\beta_{\mathcal{N}} - 2\text{SE} \leq 0$ is statistically negligible. This means that we reject the null hypothesis with unambiguously low p-values. Our results instead point to strong and robust evidence in favor our alternative hypothesis, that is, positive strategic interactions characterize groundwater-usage decisions.

We next address our first question posed at the beginning of this section: are strategic interactions the appropriate framework to be studying groundwater-usage behavior? We are interested in comparing these effects for two reasons. The first reason is motivational: perhaps the UBB is best-modeled as a single-agent decision-making process, where behavior is driven by the market and not strategic interactions? The second reason is more related to policymaking. A key mission of the UBB and URD is to equip farmers with the tools to pump groundwater at a profitable yet sustainable rate. As such, understanding whether groundwater usage is more influenced by the market versus strategic interactions helps NRDs more generally to make the most effective decisions.

It is clear from model (6) in Table C.7 that strategic interactions are the main driver of behavior. We find that crop price movements have a minor a effect on groundwater-usage. Rainfall is the second strongest predictor of behavior, where we find that a 1 SD increase in rainfall (= 5.8 inches) is associated with a 16% decrease in groundwater usage. Farmland rental rates is also strongly associated with groundwater-usage (which complements Kirwan, 2009). That said, strategic interactions is the strongest driver of behavior. We find that $\beta_{\mathcal{N}}$ is greater than 2× the absolute value of any other effect, and, on average, a 50% decrease in neighbors' pumping is associated with a 13.7% decrease in a farmer's pumping. In sum, we find evidence that strategic complements not only characterizes behavior better than strategic substitutes, but it also characterizes behavior better than models that explicitly avoid strategic interactions.

Robustness check — The testing procedure above assumed that are less than five miles away. In Appendix C.2, we relax this assumption by running the same testing procedure and allowing that neighbors can be $\{1, 3, 5, 10, 20\}$ miles away. We report p-values for each case in Table C.6 and combine the 10,000 regressions for each case via Rubin's rule in Table C.7. We find that, qualitatively, our results remain unchanged: $\beta_{\mathcal{N}} - 2\text{SE} > 0$ for almost every $\boldsymbol{A}, \boldsymbol{A} \in \mathcal{N}$ we randomly sample. The similar to above, we unambiguously reject our null

hypothesis with strong evidence that positive strategic interactions characterize groundwater pumping.

5.6. Predicting revealed preferences

We end by exploring connections between this section and the revealed preference results from Section 4. The main question we consider is: *are the positive strategic interactions in Figure 4 associated with the same mechanism(s) that underlie the low (high) acceptance rates of the GW program with strategic substitutes (complements)?* This association is not obvious. The revealed preference setup was rooted in game-theoretic reasoning, and the type of mechanisms that underlie our results are likely related to non-cooperative collusion (Green and Porter, 1984; Rotemberg and Saloner, 1986) and social norms (Young, 1993; Young and Burke, 2001; Burke and Young, 2010). By estimating interactions, however, we appeal to a much broader class of mechanisms that may not be captured by GW dynamic games with strategic complements. Whether the two tests have rejected strategic substitutes for the same reason remains to be shown.

We can test this relationship empirically using the following setup. Consider a randomly sampled network $\boldsymbol{A} \in \mathscr{N}$.

- We can use $\boldsymbol{A} = \boldsymbol{B}$ to estimate interactions β_N.
- We can also use $\boldsymbol{A}$ in a revealed preference setup. Similar to Section 4, we randomly sample farmers and also draw their neighbors according to $\boldsymbol{A}$. We test each farmer and his/her neighbors via the GW program with strategic substitutes/complements fixing $T = 2$. This means that we recover $\{1, 0\}$ each farmer where 0 = failing the GW program and 1 = passing the GW program.

In practice, we randomly sample 2,000 networks and 100 farmers form each network, which amounts to 200,000 observations. These observations are indexed according to network ($\boldsymbol{A}$), farmer (i), and consecutive time periods (τ).

We test whether β_N can predict the probability of drawing farmers that can pass the GW program with strategic substitutes/complements. Let $S_{(\boldsymbol{A})}$ ($C_{(\boldsymbol{A})}$) denote the probability of passing the GW program with strategic substitutes (complements) for network m. We can use β_N to explain $S_{(\boldsymbol{A})}$ ($C_{(\boldsymbol{A})}$) using a logistic regression setup:

$$logit\left(S_{(\boldsymbol{A})}\right) = \gamma_1 \beta_{N(\boldsymbol{A})} + \gamma_2 \left|\mathcal{N}_i\right|_{(\boldsymbol{A},i)} + \varepsilon_{(\boldsymbol{A},i,\tau)}$$

$$logit\left(C_{(\boldsymbol{A})}\right) = \gamma_3 \beta_{N(\boldsymbol{A})} + \gamma_4 \left|\mathcal{N}_i\right|_{(\boldsymbol{A},i)} + \varepsilon_{(\boldsymbol{A},i,\tau)}$$

Table 5: Predicting revealed preferences with estimated strategic interactions. Coefficients are reported in terms of log odd ratio.

	Dependent variable: Acceptance rate of the GW program			
	Substitutes		Complements	
	(S.1)	(S.2)	(C.1)	(C.2)
$\beta_{\mathcal{N}}$	-2.088^{**}	-2.004^{**}	0.552^{*}	0.520^{*}
	(0.872)	(0.874)	(0.303)	(0.304)
# of neighbors		-0.021^{***}		0.005^{***}
		(0.001)		(0.0002)
(constant)	-2.777^{***}	-0.860^{**}	0.775^{***}	0.330^{**}
	(0.385)	(0.389)	(0.134)	(0.136)
Observations	200,000	200,000	200,000	200,000
Log Likelihood	−22,764	−22,033	−115,685	−115,416
AIC	45,532	44,071	231,373	230,838
Note:				*p<0.1; **p<0.05; ***p<0.01

where $|\mathcal{N}_i|_{(\boldsymbol{A},i)}$ controls for how pass/fail rates are influenced by having more or less neighbors, and $\varepsilon_{(\boldsymbol{A},i,\tau)}$ accounts for unobserved errors.

We report results in Table 5 (coefficients are reported in terms of log odds ratio). Columns (S.1) and (S.2) use the pass/fail results of the GW program with strategic substitutes as the dependent variable. Columns (C.1) and (C.2) use the pass/fail results of the GW program with strategic complements as the dependent variable.

The key variable in Table 5 is $\beta_{\mathcal{N}}$, which relates estimated strategic interactions with pass/fail rates of the GW program. The evidence is as anticipated: higher $\beta_{\mathcal{N}}$ corresponds with (i) lower acceptance rates of the GW program with strategic substitutes and (ii) higher acceptance rates of the GW program with strategic complements. That is, the sign of $\beta_{\mathcal{N}}$ is correct. But taking a closer look at the magnitude of this effect sheds light on a somewhat different perspective. For a farmer with eight neighbors, we estimate that increasing strategic interactions from zero to $\beta_{\mathcal{N}} = 0.5$ (nearly the highest we estimated): (i) *decreases* the acceptance rate of the GW program with strategic substitutes by 17% and (ii) *increases* the acceptance rate of the GW program with strategic complements by 6%. In other words, spanning the entire space of observed $\beta_{\mathcal{N}}$ amounts to a marginal change in GW program acceptance rates.

It is true that these are not groundbreaking coefficients, which is itself an insightful finding. The interpretation is that the mechanism(s) driving the high (low) acceptance rate of the GW

program with strategic substitutes (complements) is only *loosely* related to the mechanism(s) driving the positive strategic interactions. We take this as evidence that a multiplicity of mechanisms are at play: somewhat overlapping mechanisms drive the revealed preference test results and the regression results. But, based on our findings, these mechanisms have the unifying feature that strategic complements, not substitutes, characterize groundwater-usage.

6. Outlook

This paper takes a step by putting to the test two disparate perspectives that persist in the CPR literature. Broadly speaking, the two can be contrasted as theory versus evidence, as was done in a recent *Science* policy forum (December 2018) entitled *"Tragedies Revisited"* where scholars from both sides were invited to contribute. To illustrate the divide, we cite from two contributions in this special issue:

Representing the anti-theory side, Boyd and Richerson (2018) voice sharp criticism:

> *"The enduring influence of Hardin's essay testifies to the power of a clear argument [...] This argument is clear and powerful, but wrong. Many village-scale human societies have organized hundreds of people to [...] solve commons problems, regulated not by formal coercive institutions but by informal, culturally evolved moral norms."*

In defense of theory, Jackson (2018) says:

> *"Over the past five decades, we have come to a deep understanding of commons problems and how to solve them [...] Game theory and market design have helped us understand how to provide appropriate incentives."*

These two views stand at odds. The former claims that game theory is wrong, and the solutions to commons problems lie in informal and evolved social norms, while the latter heralds the contributions game theory has made to resolving commons problems.

In our view, both statements are wrong. Game theory is neither wrong—as is evidenced by the many CPRs that are being depleted—nor right—as is evidenced by the many CPRs that are successfully governed without formal mechanisms (Ostrom, 1990). In this paper, we take the middle ground and ask how game theory must be corrected in order to reconcile the behavioral patterns that have become clear after 40+ years of fieldwork and empirical research.

Based on a unique dataset of detailed groundwater usage in the American Midwest, we firstly pursued a revealed preference test which forcefully rejected the benchmark game theory

model. We could have stopped here and taken this result as motivation to revisit, not the behavioral micro-foundations of the theory in terms of rationality and behavior *per se*, but the underlying *game* of the CPR model itself. It is possible that the canonical CPR model we presented here, despite being substantially more general than previous models, is still too simple and therefore under-specified to capture the salient drivers of behavior. We see three dimensions along which we could enrich model. First, we could obtain alternative models that could be consistent with our data by enriching the strategy space of farmers. For example, we could incorporate inter-seasonal dynamics of crop choice, the intra-seasonal dynamics of irrigation, the use of fertilizers, crop-share agreements, the use of options, etc., that result in strategic complements between farmers. Second, we could model the incomplete and asymmetric nature of information explicitly (Fershtman and Pakes, 2012), which again could change predictions in such a way that our data could be rationalized. Finally, we could elaborate the stochastic nature of the game, which would open up more complex strategies such as non-cooperative collusion that could offer a means of rationalizing our data (Green and Porter, 1984; Rotemberg and Saloner, 1986).

We could call this line of enriching the game as the "game specification correction project"—that is, taking evidence against a simple model as motivation to enrich the model—without immediately challenging the underlying behavioral assumptions.[22] This is obviously a valid approach, as we have not falsified the behavioral assumptions *per se*, but only rejected the combined hypotheses of the proposed model and the behavioral assumptions. For the moment, however, this would be a dead end in our opinion, as suitable tests for those richer models have not yet been proposed.

The dead end of this approach brings us to our exploratory analyses concerning the underlying behavioral assumptions. This analysis points to strong and robust evidence in favor of positive interaction effects, which can be viewed as evidence of a flawed behavioral model. Taking such a perspective, then should we pursue a "behavioral assumptions correction project" instead?[23] Such a project would use these findings to build alternative behavioral foundations for formulating new CPR models, where the structural components of the pre-existent game theory are preserved (rather than enriched as outlined above) but the behavioral assumptions are exchanged. Such models would incorporate behavioral assumptions that

[22] Here, we loosely adapt the expression "subjective expected utility correction project" coined by Gigerenzer and Selten (2001), which is used to describe the social preferences approach.

[23] This would include both the "subjective expected utility correction project" and other approaches allowing for bounded rationality.

would mirror other streams of behavioral economics.

Exploring this second route of revisiting, not the CPR game itself, but the behavioral foundations of CPR usage, the immediate challenge would be to formulate a model that would produce positive interaction effects. Indeed, there exists a well-developed experimental literature that has uncovered various behavioral patterns that might do this. One candidate for driving positive interaction effects is *conditional cooperation* (Fischbacher et al., 2001). Controlled laboratory and online experiments have found extensive evidence of conditional cooperation, however, the existing evidence in field work on CPRs is conflicting (Rustagi et al., 2010; Stoop et al., 2012) and suggest that such a simple account may not be sufficient. Another candidate is the presence of *social preferences* à la Rabin (1993), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Fehr and Schmidt (2006). Alternatively, social preference models where pro-/anti-social attitudes are interactive (Levine, 1998; Nax et al., 2015) and/or interdependent (Sobel, 2005) could be developed in the CPR context, which would likely produce positive interaction effects in resource usage too (via the positive interactions between preferences). The third and final candidate we would like to list is the presence of locally evolving social norms (Young, 2015), whose enforcements may require a number of elements currently not encoded in our dataset related to punishment of groundwater abusers (Fehr and Gaechter, 2000, 2002), social pressuring (Gaechter and Fehr, 1999; Masclet, 2003), and collective identities (Eckel and Grossman, 2005; Charness et al., 2007).

In sum, we hope the contributions in this paper will: (i) on the one hand, help the empirical literature side incorporate the "game specification correction project" into future analyses rather than considering game theory moot based on clearly over-simplified models, and, (ii) on the other hand, help the game theory literature incorporate the "behavioral correction project" to build behavioral models that reflect real-world empirical patterns.

These thoughts bring us back to Elinor Ostrom, the daily activities of the NRDs, and the policy implications of our findings more generally. Game theory has had a definitive impact on the governance of CPRs. Policy recommendations frequently rely on game theory as a means of arguing for/against resource regulatory measures, which is evident in, e.g., Huang and Smith (2014) who propose an alternative to the Individual Transferable Quota policy that is widely applied in CPR management today (Stavins, 2011). However, re-considering game-theoretic foundations also calls for a re-consideration of downstream policy recommendations: *"When the results of a theory seem to flow specifically from a special crucial assumption, then if the assumption is dubious, the results are suspect,"* (Solow, 1956).

As such, fresh look that this paper brings to the strategic foundations of CPRs suggests, as a next step, also taking a fresh look at the foundations of CPR institutional principles and mechanism design.

A. Proofs

A.1. Brief summary of nonsmooth optimization

Before we prove Theorems 1 and 2, we review some concepts related to nonsmooth optimization, in particular a necessary (and sufficient condition) for a local (global) maximum of a nonsmooth function (these notes are taken from Clarke, 1983).

Let $f : \mathbb{R}^K \to \mathbb{R}$ be an absolutely continuous function. The ***generalized directional derivative*** of f at $x \in \mathbb{R}^K$ in the direction of $d \in \mathbb{R}^K$ is given as

$$f^\circ(x; d) = \limsup_{y \to x, t \downarrow 0} \frac{f(y + td) - f(y)}{t}.$$

The generalized directional derivative always exists for an absolutely continuous function.

Following Clarke (1983), we define the ***subdifferential*** of f at x as

$$\partial f(x) = \{\xi \in \mathbb{R}^K : f^\circ(x; d) \geq \xi^T d \quad \forall d \in \mathbb{R}^K\}.$$

Each vector $\xi \in \partial f(x)$ is called a ***subgradient*** of f at x.

The following theorem is an analog of the first-order condition from standard smooth optimization theory. Broadly speaking, it says that x is a local maximizer of f only if 0 is a subgradient of f at x. If f is concave, then $0 \in \partial f(x)$ is both a necessary and sufficient condition for x to be a global solution to $\max_{x' \in \mathbb{R}^K} f(x')$.

Theorem 3 (Clarke, 1983, Proposition 2.3.2, p. 38). *Let $f : \mathbb{R}^K \to \mathbb{R}$ be an absolutely continuous function.*

(i) If f attains a local maximum at x then $0 \in \partial f(x)$.

(ii) If f is concave, then f attains a global maximum at x if and only if $0 \in \partial f(x)$.

A.2. Proof of Theorem 1

Let $\partial \mathcal{R}_i^t(w_i^t)$ denote the subdifferential of $\mathcal{R}_i^t$ at w_i^t. The subdifferentials $\partial \mathcal{F}_i^t(\cdot, w_{-i}^t)$ and $\partial \mathcal{P}_i^t(\cdot, w_{-i}^t)$ with respect to w_i^t are defined accordingly. These subdifferentials are non-empty sets since all functions are assumed to be absolutely continuous.

A.2.1. (B) $\implies$ (A)

We prove the necessity of (CGD), (MGD), and (MLD) separately.

Necessity of (CGD) — Consider farmer i in season t. Because $\mathcal{P}_i^t$ is absolutely continuous and $w_i^t \in \mathcal{W}_i = \mathbb{R}_+$ is an interior solution, it follows from Theorem 3 that w_i^t solves $\max_{x_i \in \mathcal{W}_i} \mathcal{P}_i^t(x_i, w_{-i}^t)$ only if $0 \in \partial \mathcal{P}_i^t(w_i^t, w_{-i}^t)$. That is, $w_i^t \in \arg\max_{x_i \in \mathcal{W}_i} \mathcal{P}_i^t(x_i, w_{-i}^t)$ only if there exists $\tilde{\rho}_i^t \in \partial \mathcal{R}_i^t(w_i^t)$ and $\tilde{f}_i^t \in \partial \mathcal{F}_i^t(w_i^t, w_{-i}^t)$ such that

$$\tilde{\rho}_i^t - e_i \cdot \left(g_i^{t,Spring} + \mathcal{G}(W^t - r^t) + \mathcal{L}_i(w_{\mathcal{N}_i}^t + w_i^t) \right) + \tilde{f}_i^t = 0. \tag{A.1}$$

Let $\hat{\lambda}_i^t = \mathcal{L}_i(w_{\mathcal{N}_i}^t + w_i^t)$ and $\hat{f}_i^t = \left(\tilde{\rho}_i^t + \tilde{f}_i^t \right) / e_i$. Then re-arranging terms in (A.1) yields

$$\mathcal{G}(W^t - r^t) = \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring}. \tag{A.2}$$

Note that $\mathcal{G}(W^t - r^t)$ is common for all farmers. Consequently,

$$\forall i, j, t: \qquad \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring} = \hat{f}_j^t - \hat{\lambda}_j^t - g_j^{t,Spring}.$$

Necessity of (MGD) — By assumption, $\mathcal{G}$ is absolutely continuous and nondecreasing in $\{W^t - r^t\}_{t \in \mathcal{T}}$. Then for any $z', z'' \in \{W^t - r^t\}_{t \in \mathcal{T}}$, it follows that $z' > (=)\ z'' \implies \mathcal{G}(z') > (=)\ \mathcal{G}(z'')$. Translating this to farmer $i \in \mathcal{N}$ in seasons $s, t \in \mathcal{T}$ amounts to

$$\begin{aligned} & W^s - r^s > (=)\ W^t - r^t \\ \implies \quad & \mathcal{G}(W^s - r^s) > (=)\ \mathcal{G}(W^t - r^t) \\ \iff \quad & \hat{f}_i^s - \hat{\lambda}_i^s - g_i^{s,Spring} > (=)\ \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring} \end{aligned}$$

where $\iff$ holds from (A.2). Since $W^s - r^s > (=)\ W^t - r^t \implies \hat{f}_i^s - \hat{\lambda}_i^s - g_i^{s,Spring} > (=)\ \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring}$, it follows that (MGD) is necessary.

Necessity of (MLD) — By assumption, $\mathcal{L}_i$ is absolutely continuous and nondecreasing in $\{w_{\mathcal{N}_i}^t + w_i^t\}_{t \in \mathcal{T}}$. This means that, for any $s, t \in \mathcal{T}$,

$$\begin{aligned} & w_{\mathcal{N}_i}^s + w_i^s > (=)\ w_{\mathcal{N}_i}^t + w_i^t \\ \implies \quad & \mathcal{L}_i(w_{\mathcal{N}_i}^s + w_i^s) > (=)\ \mathcal{L}_i(w_{\mathcal{N}_i}^t + w_i^t) \\ \iff \quad & \hat{\lambda}_i^s - \hat{\lambda}_i^t > (=)\ 0 \end{aligned}$$

where $\iff$ holds by the definition of $\hat{\lambda}_i^s$ and $\hat{\lambda}_i^s$. Since $w_{\mathcal{N}_i}^s + w_i^s > (=)\ w_{\mathcal{N}_i}^t + w_i^t \implies hat\lambda_i^s - \hat{\lambda}_i^t, > (=)\ 0$, it follows that (MLD) is necessary. *Q.E.D.*

A.2.2. (A) $\implies$ (B)

Here, we prove that the GW program is sufficient for the existence of a GW dynamic game that can rationalize $\mathscr{O}$.

Suppose that $\{\hat{f}_i^t, \hat{\lambda}_i^t\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ are positive numbers that solve the GW program. We construct each function underlying the GW dynamic game separately.

Revenue functions — We define all revenue functions as constant (which means they are all concave). That is, $\forall i \in \mathcal{N}$ and $t \in \mathcal{T}$, we let $\mathcal{R}_i^t(w_i) = R\ \forall w_i \in \mathcal{W}_i$.

Local drawdown functions — We begin by defining local drawdown functions. For any farmer i, let $\{w_{\mathcal{N}_i}^t + w_i^t\}_{t\in\mathcal{T}}$ be in increasing order. Then it follows from (MLD) that $w_{\mathcal{N}_i}^t + w_i^t > (=)\ w_{\mathcal{N}_i}^{t+1} + w_i^{t+1} \implies \hat{\lambda}_i^t > (=)\ \hat{\lambda}_i^{t+1}$. We define $\mathcal{L}_i : \mathbb{R}_+ \to \mathbb{R}_+$ as a piecewise linear function such that $\mathcal{L}_i(w_{\mathcal{N}_i}^t + w_i^t) = \hat{\lambda}_i^t\ \forall t \in \mathcal{T}$. As such, (MLD) ensures that $\mathcal{L}_i$ is an absolutely continuous and nondecreasing function of $w_{\mathcal{N}_i} + w_i$.

Global depletion function — Fix $i \in \mathcal{N}$ (which i is used is inconsequential because of (CDG)). Let $\{W^t - r^t\}_{t\in\mathcal{T}}$ be in increasing order. It then follows from (MGD) that $W^t - r^t > (=)\ W^{t+1} - r^{t+1} \implies \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring} > (=)\ \hat{f}_i^{t+1} - \hat{\lambda}_i^{t+1} - g_i^{t+1,Spring}$. We define $\mathcal{G} : \mathbb{R} \to \mathbb{R}$ as a piecewise linear function such that $\mathcal{G}(W^t - r^t) = \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring}\ \forall t \in \mathcal{T}$. As such, (MGD) ensures that $\mathcal{G} : \mathbb{R} \to \mathbb{R}$ defined as such is an absolutely continuous and nondecreasing function of $W - r$.

Energy cost — We let $e_i = 1\ \forall i \in \mathcal{N}$.

Future profit functions — Fix $i \in \mathcal{N}$ and $t \in \mathcal{T}$. Let $h_i^t : \mathcal{W} \to \mathbb{R}_+$ be a function of (w_i, w_{-i}) with the following properties: (i) h_i^t is continuously differentiable in all arguments, (ii) $h_i^t(\cdot, w_{-i})$ is strictly decreasing in w_i for every $w_{-i} \in \mathcal{W}_{-i}$, and (iii) $h_i^t(w_i^t, w_{-i}^t) = \hat{f}_i^t$. We define a future profit function $\mathcal{F}_i^t : \mathcal{W} \to \mathbb{R}$ as follows:

$$\mathcal{F}_i^t(w_i, w_{-i}) = \int_0^{w_i} h_i^t(\xi_i, w_{-i})\, d\xi_i.$$

It follows by construction that: (i) $\mathcal{F}_i$ is concave in w_i because $h_i^t(\cdot, w_{-i})$ is strictly decreasing in w_i for each w_{-i} and (ii) $\hat{f}_i^t \in \partial\mathcal{F}_i^t(w_i^t, w_{-i}^t)$. Each $(\mathcal{F}_i^t)_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ is defined accordingly.

Optimizing net-present value — All that remains is to show that w_i^t optimizes i's net-present value in season t. By construction, $\mathcal{P}_i^t$ is concave in w_i^t since: $\mathcal{R}_i^t$ is concave, $-C_i^t$ is concave in w_i^t as $\mathcal{G}$ and $\mathcal{L}_i$ are nondecreasing in their respective arguments, and $\mathcal{F}_i^t$ is concave in w_i^t

for each w^t_{-i}. Consequently, it follows from Theorem 3(ii) that w^t_i solves $\max_{x_i \in \mathcal{W}_i} \mathcal{P}^t_i(x_i, w^t_{-i})$ if and only if $0 \in \partial \mathcal{P}^t_i(w^t_i, w^t_{-i})$. This holds if

$$\begin{aligned} 0 - g^{t,Spring}_i - \mathcal{G}(W^t - r^t) - \mathcal{L}_i(w^t_{\mathcal{N}_i} + w^t_i) + \hat{f}^t_i = 0 \\ \iff \qquad \hat{f}^t_i - \hat{\lambda}^t_i - g^{t,Spring}_i = \mathcal{G}(W^t - r^t), \end{aligned} \tag{A.3}$$

where the latter equation in (A.3) is ensured by (CGD). Hence, w^t_i solves $\max_{x_i \in \mathcal{W}_i} \mathcal{P}^t_i(x_i, w^t_{-i})$ by Theorem 3. *Q.E.D.*

A.3. Proof of Theorem 2

A.3.1. (D) $\implies$ (C)

Here, we prove the necessity of the GW program with strategic substitutes/complements. Necessity of (CGD), (MGD), and (MLD) follow directly from the proof of Theorem 1. Hence, all that remains is to show the necessity of (SUBS)/(COMP).

Let $\{\tilde{\rho}^t_i\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ and $\{\tilde{f}^t_i\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ be from the proof of Theorem 1 (specifically Section A.2.1). Define $\rho_i(\cdot)$ such that (i) $\rho_i(w_i) \in \partial \mathcal{R}_i(w_i)\ \forall w_i \in \mathcal{W}_i$ and (ii) $\rho_i(w^t_i) = \tilde{\rho}^t_i$ for every $t \in \mathcal{T}$. Similarly, define $f_i(\cdot)$ such that (i) $f_i(w_i, w_{\mathcal{N}_i}) \in \partial \mathcal{F}_i(w_i, w_{\mathcal{N}_i})\ \forall (w_i, w_{\mathcal{N}_i}) \in \mathbb{R}^2_+$ and (ii) $f_i(w^t_i, w^t_{\mathcal{N}_i}) = \tilde{f}^t_i\ \forall t \in \mathcal{T}$.

Consider farmer $i \in \mathcal{N}$ in seasons $s, t \in \mathcal{T}$. Suppose that $w^s_{\mathcal{N}_i} > w^t_{\mathcal{N}_i}$ ($w^s_{\mathcal{N}_i} < w^t_{\mathcal{N}_i}$) and $w^s_i > w^t_i$. Consider $\frac{\partial}{\partial w_i}\mathcal{P}^t_i$:

$$\frac{\partial}{\partial w_i}\mathcal{P}^t_i = \rho_i(w^t_i) - e_i \cdot (g^{t,Spring}_i + \mathcal{G}(W^t - r^t) + \overbrace{\mathcal{L}_i(w^t_{\mathcal{N}_i} + w^t_i)}^{=\hat{\lambda}^t_i}) + f_i(w^t_i, w^t_{\mathcal{N}_i}) \tag{A.4}$$

$$\geq \rho_i(w^s_i) - e_i \cdot \left(g^{t,Spring}_i + \mathcal{G}(W^t - r^t) + \mathcal{L}_i(w^t_{\mathcal{N}_i} + w^t_i)\right) + f_i(w^s_i, w^t_{\mathcal{N}_i}) \tag{A.5}$$

$$\geq \rho_i(w^s_i) - e_i \cdot \left(g^{t,Spring}_i + \mathcal{G}(W^t - r^t) + \mathcal{L}_i(w^s_{\mathcal{N}_i} + w^t_i)\right) + f_i(w^s_i, w^s_{\mathcal{N}_i}) \tag{A.6}$$

$$\geq \rho_i(w^s_i) - e_i \cdot (g^{t,Spring}_i + \mathcal{G}(W^t - r^t) + \underbrace{\mathcal{L}_i(w^s_{\mathcal{N}_i} + w^s_i)}_{\hat{\lambda}^s_i}) + f_i(w^s_i, w^s_{\mathcal{N}_i}) \tag{A.7}$$

were (A.5) follows because $\mathcal{R}_i$ and $\mathcal{F}_i$ are concave in w_i and $w^s_i > w^t_i$, (A.6) follows because $\mathcal{P}^t_i$ is submodular (supermodular) in $(w_i, w_{\mathcal{N}_i})$ and $\mathcal{G}$ is slowly changing in $w_{\mathcal{N}_i}$, and (A.7) follows because $\mathcal{L}_i$ is nondecreasing and $w^s_i > w^t_i$. Because $\hat{f}^t_i = \left(\rho_i(w^t_i) + f_i(w^t_i, w^t_{\mathcal{N}_i})\right)/e_i$ and

$\hat{f}_i^s = \left(\rho_i(w_i^s) + f_i(w_i^s, w_{\mathcal{N}_i}^s)\right)/e_i$, it follows that (A.4) $\geq$ (A.7) if and only if

$$\begin{aligned} \hat{f}_i^t - g_i^{t,Spring} - \mathcal{G}(W^t - r^t) - \hat{\lambda}_i^t \geq \hat{f}_i^s - g_i^{t,Spring} - \mathcal{G}(W^t - r^t) - \hat{\lambda}_i^s \\ \iff \quad 0 \geq \left(\hat{f}_i^s - \hat{\lambda}_i^s\right) - \left(\hat{f}_i^t - \hat{\lambda}_i^t\right), \end{aligned}$$

which means that (SUBS)/(COMP) is necessary. *Q.E.D.*

A.3.2. (C) $\implies$ (D)

Here, we prove the sufficiency of the GW program with strategic substitutes (complements). Let $\{\hat{f}_i^t, \hat{\lambda}_i^t\}_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ be positive numbers that solve the GW program with strategic substitutes (complements).

Revenue functions — We define all revenue functions as constant (which means they are all concave). That is, $\forall i \in \mathcal{N}$, we let $\mathcal{R}_i(w_i) = R\ \forall w_i \in \mathcal{W}_i$.

Local drawdown functions — We let $\mathcal{L}_i(x) = 0\ \forall x \in \mathbb{R}_+$.

Global depletion functions — Fix $i \in \mathcal{N}$ (which i is used is inconsequential because of (CDG)). Let $\{W^t - r^t\}_{t\in\mathcal{T}}$ be in increasing order. It then follows from (MGD) that $W^t - r^t >$ (=) $W^{t+1} - r^{t+1} \implies \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring} >$ (=) $\hat{f}_i^{t+1} - \hat{\lambda}_i^{t+1} - g_i^{t+1,Spring}$. We define $\mathcal{G} : \mathbb{R} \to \mathbb{R}$ as a piecewise linear function such that $\mathcal{G}(W^t - r^t) = \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring}\ \forall t \in \mathcal{T}$. As such, (MGD) ensures that $\mathcal{G} : \mathbb{R} \to \mathbb{R}$ defined as such is an absolutely continuous and nondecreasing function of $W - r$.

Energy costs — We let $e_i = 1$ for every $(i, t) \in \mathcal{N} \times \mathcal{T}$.

Future profit functions — Fix $i \in \mathcal{N}$. Let $h_i : \mathbb{R}_+^2 \to \mathbb{R}_+$ be a function of $(w_i, w_{\mathcal{N}_i})$ with the following properties: (i) h_i is absolutely continuous in all arguments, (ii) $h_i(\cdot, w_{\mathcal{N}_i})$ is strictly decreasing in w_i for every $w_{\mathcal{N}_i} \in \mathbb{R}_+$, (iii) $h_i(w_i, \cdot)$ is strictly decreasing/increasing in $w_{\mathcal{N}_i}$ for every $w_i \in \mathbb{R}_+$, and (iv) $h_i(w_i^t, w_{\mathcal{N}_i}^t) = \hat{f}_i^t - \hat{\lambda}_i^t$ for each $t \in \mathcal{T}$. Note that (iii) and (iv) can be ensured by (SUBS)/(COMP).

We define a future profit function $\mathcal{F}_i : \mathbb{R}_+^2 \to \mathbb{R}$ as follows:

$$\mathcal{F}_i(w_i, w_{\mathcal{N}_i}) = \int_0^{w_i} h_i(\xi_i, w_{\mathcal{N}_i})\, d\xi_i.$$

It follows by construction that: (i) $\mathcal{F}_i$ is concave in w_i because $h_i(\cdot, w_{\mathcal{N}_i})$ is strictly decreasing in each $w_{\mathcal{N}_i}$ and (ii) $\hat{f}_i^t - \hat{\lambda}_i^t \in \partial\mathcal{F}_i^t(w_i^t, w_{-i}^t)\ \forall t \in \mathcal{T}$. Each $(\mathcal{F}_i)_{i\in\mathcal{N}}$ is defined accordingly.

Submodularity (supermodularity) of $\mathcal{P}_i^t$ — We must check that, given the definitions above, $\mathcal{P}_i^t$ is indeed submodular (supermodular) in $\mathcal{P}_i^t$. The partial derivative of $\mathcal{P}_i^t$ is given as

$$\frac{\partial}{\partial w_i}\mathcal{P}_i^t(w_i, w_{-i}) = 0 - e_i \cdot \left(g_i^{t,Spring} + \mathcal{G}(W - r^t)\right) + h_i(w_i, w_{\mathcal{N}_i}) \tag{A.8}$$

(which exists a.e. since $\mathcal{P}_i^t$ is absolutely continuous in w_i). Note that $\mathcal{G}$ is slowly changing in $w_{\mathcal{N}_i}$ by assumption. Consequently, because h_i is decreasing (increasing) in $w_{\mathcal{N}_i}$, $\frac{\partial}{\partial w_i}\mathcal{P}_i^t$ is decreasing (increasing) in $w_{\mathcal{N}_i}$. This means that $\mathcal{P}_i^t$ is submodular (supermodular) in $(w_i, w_{\mathcal{N}_i})$.

Optimizing net-present value — Consider i's net-present value. By construction, $\mathcal{P}_i^t$ is concave in w_i^t since: $\mathcal{R}_i$ is concave, $-C_i^t$ is concave in w_i^t as $\mathcal{G}(\cdot)$ is a nondecreasing function, and $\mathcal{F}_i(\cdot, w_{\mathcal{N}_i}^t)$ is concave in w_i^t for each $w_{\mathcal{N}_i}^t$. Consequently, it follows from Theorem 3(ii) that w_i^t solves $\max_{x_i \in \mathcal{W}_i} \mathcal{P}_i(x_i, w_{-i}^t)$ if and only if $0 \in \partial\mathcal{P}_i^t(w_i^t, w_{-i}^t)$. This holds if

$$\begin{aligned} 0 - \left(g_i^{t,Spring} + \mathcal{G}(W^t - r^t) + 0\right) + \overbrace{\hat{f}_i^t - \hat{\lambda}_i^t}^{\in \mathcal{F}_i(w_i^t, w_{-i}^t)} &= 0 \\ \iff \qquad \hat{f}_i^t - \hat{\lambda}_i^t - g_i^{t,Spring} &= \mathcal{G}(W^t - r^t), \end{aligned} \tag{A.9}$$

where the latter equation in (A.9) is ensured by (CGD). Hence, w_i^t solves $\max_{x_i \in \mathcal{W}_i} \mathcal{P}_i(x_i, w_{-i}^t)$ by Theorem 3. *Q.E.D.*

B. Instrumental variables

In order to estimate (6,8) robustly, we rely on the following assumptions (for notation: $\boldsymbol{I}_N$ is the N-by-N identity matrix).

Assumption 2. *In the regression model (6,8):*

> *(A1.1) The individual-time innovations $(\nu_i^t)_{(i,t)\in\mathcal{N}\times\mathcal{T}}$ are i.i.d. with finite absolute $4 + \delta_\nu$ moments for some $\delta_\nu > 0$. Furthermore, $\mathbb{E}[\nu_i^t] = 0$ and $\mathbb{E}[(\nu_i^t)^2] = \sigma_\nu^2 > 0$.*
>
> *(A1.2) The matrices $(\boldsymbol{I}_N - \beta_N \boldsymbol{A})$ and $(\boldsymbol{I}_N - \rho\boldsymbol{B})$ are non-singular.*

Assumption (A1.1) ensures that individual innovations are well-behaved. Assumption (A1.2) is a mild condition in our setting and ensures that the matrix of strategic interactions is also well-behaved for estimation purposes. See Kapoor et al. (2007), Baltagi (2008), Bramoulle et al. (2009), and Mutl and Pfaffermayr (2011) who also make such assumptions.

B.1. Derivation of IVs

Mutl and Pfaffermayr (2011) derived the ideal set of IVs for estimating (7,8). To present and discuss these IVs, we require additional notation. Stack groundwater usage in the following way:

$$\boldsymbol{w} = \left(w_1^1, w_2^1, \ldots, w_N^1, w_1^2, \ldots, w_{N-1}^T, w_N^T\right)$$

We can write the regression model compactly by stacking the explanatory and control variables accordingly:

$$\begin{aligned} \boldsymbol{w} &= \boldsymbol{I}_{NT}\boldsymbol{\alpha} + \beta_N \mathbb{A}\boldsymbol{w} + \boldsymbol{X}\boldsymbol{\Psi} + \mathbb{D}\boldsymbol{\gamma} + \boldsymbol{u} \\ \boldsymbol{u} &= \rho\mathbb{B}\boldsymbol{u} + \boldsymbol{\nu} \end{aligned} \tag{B.1}$$

where $\boldsymbol{I}_{NT}$ is an identify matrix of size $(N \cdot T) \times (N \cdot T)$. Denoting $\boldsymbol{\iota}_T$ as a vector of ones of length T and $\otimes$ as the Kronecker product: $\mathbb{A} = \boldsymbol{I}_T \otimes \boldsymbol{A}$ is the stacked strategic network, $\mathbb{D} = \boldsymbol{\iota}_T \otimes \boldsymbol{D}$ is the stacked individual controls matrix, and $\mathbb{B} = \boldsymbol{I}_T \otimes \boldsymbol{B}$ is the stacked disturbance matrix. The remaining variables—namely $\boldsymbol{\alpha}$, $\boldsymbol{X}$, $\boldsymbol{\gamma}$, $\boldsymbol{u}$, and $\boldsymbol{\nu}$—are ordered as $\boldsymbol{w}$. Note that $\boldsymbol{X}$ collects both $\boldsymbol{x}_i^t$ and $\boldsymbol{y}^t$ terms, and $\boldsymbol{\Psi}$ collects (ψ, ϕ).

Next, we introduce a standard *within transformation matrix* to normalize (B.1) (see, e.g., Mundlak, 1978, and Baltagi, 2008). To this end, let $\boldsymbol{J}_T$ be a $T \times T$ matrix of unit elements. We define $\boldsymbol{Q}_0 = \left(\boldsymbol{I}_T - \frac{1}{T}\boldsymbol{J}_T\right) \otimes \boldsymbol{I}_N$ as the standard *within transformation matrix* such that, for any matrix $\mathbb{A}$ of size $(N \cdot T) \times (N \cdot T)$, $\boldsymbol{Q}_0\mathbb{A}$ results in a matrix such that the mean of every row is zero:

$$(\boldsymbol{Q}_0\mathbb{A})_{kl} = \mathbb{A}_{kl} - \frac{1}{N \cdot T}\sum_{l'=1}^{N \cdot T} \mathbb{A}_{kl'}.$$

It follows that

$$\begin{aligned} \boldsymbol{Q}_0\boldsymbol{w} &= \beta_N \boldsymbol{Q}_0\mathbb{A}\boldsymbol{w} + \boldsymbol{Q}_0\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Q}_0\boldsymbol{u} \\ \boldsymbol{Q}_0\boldsymbol{u} &= \rho\boldsymbol{Q}_0\mathbb{B}\boldsymbol{u} + \boldsymbol{Q}_0\boldsymbol{\nu} \end{aligned} \tag{B.2}$$

where $\boldsymbol{Q}_0$ removes time-invariant elements including $\boldsymbol{\alpha}$ and $\mathbb{D}$. This equation is what we ultimately estimate, since it retains the key variable of interest, β_N, with fewer terms to handle.

Mutl and Pfaffermayr (2011) show that the ideal set of IVs for estimating (B.2) is given by $\boldsymbol{Q}_0\boldsymbol{G}_0$ where $\boldsymbol{G}_0$ contains a subset of linearly independent columns of

$$\left[\boldsymbol{X}, \mathbb{A}\boldsymbol{X}, \mathbb{A}^2\boldsymbol{X}, \ldots, \mathbb{B}\boldsymbol{X}, \mathbb{B}\mathbb{A}\boldsymbol{X}, \mathbb{B}\mathbb{A}^2\boldsymbol{X}, \ldots\right]. \tag{B.3}$$

The interpretation is as follows. Consider i's neighbors, $\mathcal{N}_i$. $\boldsymbol{X}$ represents i's individual attributes; $\mathbb{A}\boldsymbol{X}$ represents i's neighbors' attributes; $\mathbb{A}^2\boldsymbol{X}$ represents i's neighbors' neighbors' attributes; etc. Each are exogenous from the perspective of farmer i. In addition, because behaviors are assumed to be networked: $\boldsymbol{X}$ explains i's behavior, which in turn influences $\mathcal{N}_i$'s behavior; $\mathbb{A}\boldsymbol{X}$ explains $\mathcal{N}_i$'s behavior; $\mathbb{A}^2\boldsymbol{X}$ explains $\mathcal{N}_i$'s neighbors' behavior, which in turn influences $\mathcal{N}_i$'s behavior, etc. Put another way, the network itself allows us to use farmers' attributes as exogenous IVs to explain behavior in different parts of the network.

With Assumption 2 and an additional asymptotic assumption common in the spatial panel data literature,[24] the estimation procedure proposed by Mutl and Pfaffermayr (2011) is consistent and asymptotically normal (we refer the interested to details therein). We follow this procedure to formulate our results in the main text.

C. Robustness checks

C.1. Revealed preferences on different network structures

In this appendix section, we subject our revealed preference test results from Section 4 to two robustness checks.

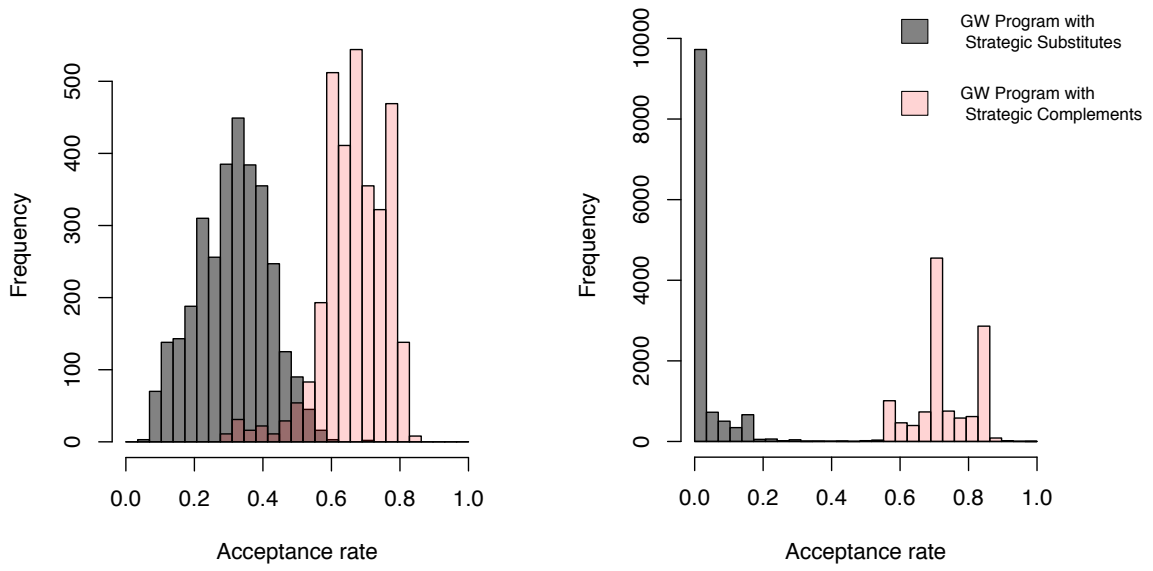
The first concerns network structure, where we ask: *are our revealed preference test results from Section 4 robust to different network structures*? The second concerns viewing our revealed preferences results at the population vs. individual level, where we ask: *are all farmers' pumping decisions explained by strategic complements, or are decisions better characterized by a mix of substitutes and complements?* In other words, we are interested in disaggregating our main findings from Section 4 from the population perspective to the individual-level perspective.

We address both robustness questions by taking an approach similar to Section 5. Informally, we define a set of possible networks which we assume contains at least one network that represents true strategic interaction in the UBB/URD. We then randomly sample from this set, and we utilize these networks to calculate how often farmer $i + \mathcal{N}_i$ pass the GW program with strategic substitutes/complements. With enough samples, we then have an estimation of the probability of finding a network structure such that groundwater pumping decisions by $i + \mathcal{N}_i$ are strategic substitutes/complements.

We now define this measure more formally. As defined in Section 5, we say that $\mathscr{N} \subseteq \mathbb{R}_+^{NT}$ is the ***set of feasible networks*** if it contains every $\boldsymbol{A} \in \mathbb{R}_+^{NT}$ such that: (i) $\boldsymbol{A}_{ii} = 0$, or no

[24] In particular, we require Assumption 3.3 from Mutl and Pfaffermayr (2011). Other versions of this assumption can be found in Kelejian and Prucha (1998, 1999) and Kapoor et al. (2007).

Farmer-level acceptance rates of the GW program with strategic substitutes/complements



(a) Upper Republican District

(b) Upper Big Blue District

Figure C.5: Histogram distributions of farmer-level passing rates of the GW program with strategic substitutes/complements in the (a) Upper Republican District and (b) Upper Big Blue District. The farmer-level acceptance rate = Prob[randomly sampling a network from $\mathscr{N}$ such that $i + \mathcal{N}_i$ pass the respective GW program]; see Section 5.4 for the definition of $\mathscr{N}$.

farmer strategically affects him/herself, (ii) $\boldsymbol{A}_{ij} > 0$ only if i and j are less than 5 miles apart, depending on the case, (that is, only neighbors induce strategic interaction effects), (iii) $\boldsymbol{A}_{ij} = \boldsymbol{A}_{ik}$ for any j and k with non-zero weights, or neighbors strategically affect i equally, and (iv) $\boldsymbol{A}$ is row normalized. In other words, $\mathscr{N}$ is the set of all possible networks such that neighbors are less than five miles away. In what follows, we assume that at least one network in $\mathscr{N}$ represents true strategic interaction in the UBB/URD.

For each farmer $i \in \mathcal{N}$, we define ρ_i^{SUBS} in the following way:

$$\rho_i^{SUBS} = \text{Prob}_i\left[\boldsymbol{A} \in \mathscr{N} \quad : \quad \begin{array}{l}\text{farmers } \{i\} \cup \{j \in \mathcal{N} : \boldsymbol{A}_{ij} = 1\} \text{ pass the} \\ \text{GW program with strategic } \textit{substitutes}\end{array}\right].$$

In other words, ρ_i^{SUBS} is the probability of finding a network structure such that $i + \mathcal{N}_i$ pass the GW program with strategic substitutes. We can define a measure for i passing the GW

program with strategic complements similarly:

$$\rho_i^{COMP} = \text{Prob}_i\left[\boldsymbol{A} \in \mathscr{N} \quad : \quad \begin{array}{l}\text{farmers } \{i\} \cup \{j \in \mathcal{N} : \boldsymbol{A}_{ij} = 1\} \text{ pass the} \\ \text{GW program with strategic } \textit{complements}\end{array}\right].$$

We estimate ρ_i^{SUBS} and ρ_i^{COMP} for each farmer in the UBB and URD by sampling $1,000$ networks.

We report results in Figure C.5, which point to two findings that support those in the main text. First, we find that strategic complements characterizes groundwater pumping better than strategic substitutes for every farmer in the UBB (Figure C.5(b)) and most farmers in the URD (Figure C.5(c)). This observation is particularly strong for the UBB: farmers' acceptance rates of the GW program with strategic complements cluster around 0.7, whereas acceptance rates for the strategic substitutes analog clusters around zero. Second, we find that strategic complements explains behavior in the UBB better than the URD. The latter seems to suggest, at this farmer-level perspective, that individual-level behavior is better characterized by a mix of strategic substitutes and complements. But in the UBB, this ambiguity is absent. These findings are consistent with the stronger trends we observe in the UBB in Tables 2 and 3 vs. the trends we observe in the URD in Tables C.8 and C.9.

C.2. Strategic interactions on different network structures

In Section 5, we estimated strategic interaction by assuming that neighbors were less than 5 miles away. Our reasoning behind the '5 miles' assumption to subsume the local drawdown network from Section 4 while also permitting interactions beyond those captured by game-theoretic assumptions.

In this section, we explore what happens when we weaken (strengthen) the '5 mile' assumption by increasing (decreasing) this distance. We suppose that neighbors may come from less than $\{1, 3, 10, 20\}$ miles away. $\{1, 3\}$ are distances that get closer to the local drawdown networks. $\{10, 20\}$ considerably increase the space of possible network configurations, thereby increasing the likelihood of drawing networks that are consistent with strategic substitutes. For each case, we utilize the same test procedure as outlined in Section 5.

We report results in Table C.6 (we re-report the 5 miles case for the sake of comparison). This table shows that the results reported in the main text are robust to different ways of defining neighborhoods. Firstly, as in the 5 mile case, we evidence that the distribution of $\beta_{\mathcal{N}} - 2\text{SE} < 0$ follows a normal distribution. Secondly, for each case $\{1, 3, 10, 20\}$, we find that the probability of drawing networks that result in negative strategic interactions (or

Table C.6: Hypothesis test results for $\mathscr{H}_0'$ vs. $\mathscr{H}_A'$.

	Hypothesis test results for $\mathscr{H}_0'$ vs. $\mathscr{H}_A'$		
	$\mathcal{N}(\mu, \sigma^2)$	KS p-value	p-value for $\mathscr{H}_0'$ vs. $\mathscr{H}_A'$
Neighs $\leq$ 1 mile	$(0.072, 0.009)$	0.70	<0.001
Neighs $\leq$ 3 miles	$(0.234, 0.014)$	0.70	<0.001
Neighs $\leq$ 5 miles	$(0.351, 0.015)$	0.70	<0.001
Neighs $\leq$ 10 miles	$(0.434, 0.013)$	0.69	<0.001
Neighs $\leq$ 20 miles	$(0.29, 0.009)$	0.69	<0.001

$\beta_{\mathcal{N}} - 2\text{SE} < 0$) is statistically negligible. This means that we reject the null hypothesis $\mathscr{H}_0'$ based on strong evidence for *positive* strategic interactions.

In Table C.7, we also report the aggregated results of each regression. As in the 5 mile case, we find that strategic interactions is the strongest predictor of groundwater-usage.

Table C.7: Micro-level regressions on UBB data. Number of observations = 97,160. $\{1, 3, 5, 10, 20\}$ mile(s) corresponds to neighborhood sizes as in Figure 4. Coefficients of control variables are normalized such that β =% *water*$_i^t$ / +1*SD variable*.

	Dependent variable: Log (Water$_i^t$)				
	(1 mile)	(3 miles)	(5 miles)	(10 miles)	(20 miles)
Strategic interaction effects					
Log(Neighbors' water$_i^t$)	0.095***	0.271***	0.395***	0.483***	0.339***
	(0.015)	(0.023)	(0.027)	(0.027)	(0.026)
Groundwater controls					
Spring GW$_i^t$	−0.063***	−0.059***	−0.057***	−0.055***	−0.053***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Spring GW$_i^t$ − Fall GW$_i^{t-1}$	0.059***	0.055***	0.051***	0.047***	0.045***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
UBB-wide time-level controls					
Spot market price^t	0.027***	0.02***	0.016***	0.014***	0.024***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Electricity^t	−0.172***	−0.137***	−0.112***	−0.095***	−0.126***
	(0.007)	(0.008)	(0.009)	(0.009)	(0.012)
Land rental rates$_i^t$	0.16***	0.129***	0.107***	0.09***	0.102***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)
Rain^t	−0.242***	−0.196***	−0.163***	−0.141***	−0.182***
	(0.005)	(0.007)	(0.008)	(0.008)	(0.008)
Temperature^t	−0.156***	−0.124***	−0.102***	−0.088***	−0.122***
	(0.006)	(0.007)	(0.007)	(0.008)	(0.01)
Farmer-level and time-level controls					
(Temp.)×(Rain)^t	−0.078***	−0.065***	−0.055***	−0.049***	−0.069***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.007)
(Land-size)×(Rain)$_i^t$	−0.017***	−0.017***	−0.017***	−0.017***	−0.017***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
(Land-size)×(Temp.)$_i^t$	−0.012***	−0.012***	−0.012***	−0.012***	−0.012***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
(Well depth)×(Rain)$_i^t$	−0.011***	−0.011***	−0.01***	−0.009***	−0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
(Well depth)×(Temp.)$_i^t$	0.006**	0.006**	0.005**	0.006**	0.006**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
(Land-size)×(Rain)×(Temp.)$_i^t$	−0.016***	−0.016***	−0.016***	−0.016***	−0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
(Well depth)×(Rain)×(Temp.)$_i^t$	−0.019***	−0.018***	−0.018***	−0.017***	−0.017***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Note: *p<0.05; **p<0.01; ***p<0.001

Table C.8: Upper Republican District — Acceptance rates of the GW program with (*Panel A*) strategic substitutes and (*Panel B*) strategic complements (sub-sample = 1,000).

Panel A – Substitutes		*Number of farmers ($\mathcal{N}$)*					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	0.79	0.76	0.67	0.64	0.59	0.54
	$T = 3$	0.58	0.47	0.43	0.33	0.28	0.23
	$T = 4$	0.41	0.30	0.21	0.15	0.10	0.05
	$T = 5$	0.27	0.19	0.15	0.09	0.06	0.02
	$T = 6$	0.25	0.14	0.09	0.05	0.03	0
	$T = 8$	0.14	0.07	0.02	0.01	0	0
Panel B – Complements		*Number of farmers ($\mathcal{N}$)*					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	0.90	0.85	0.80	0.74	0.70	0.62
	$T = 3$	0.72	0.63	0.54	0.46	0.38	0.30
	$T = 4$	0.61	0.42	0.31	0.21	0.14	0.09
	$T = 5$	0.48	0.26	0.14	0.07	0.05	0.02
	$T = 6$	0.31	0.17	0.06	0.03	0.02	0.01
	$T = 8$	0.18	0.07	0.01	0.01	0	0

Table C.9: Upper Republican District — Selten and Krischker's (1983) measure of predictive success of the GW program with (*Panel A*) strategic substitutes and (*Panel B*) strategic complements (sub-sample = 1,000).

Panel A – Strategic substitutes		*Number of farmers ($\mathcal{N}$)*					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	−0.09	−0.09	−0.11	−0.06	−0.04	0.03
	$T = 3$	−0.14	−0.16	−0.08	−0.05	−0.02	0.06
	$T = 4$	−0.15	−0.11	−0.07	−0.05	0	0
	$T = 5$	−0.19	−0.09	−0.01	0.01	0.02	0.01
	$T = 6$	−0.07	−0.04	−0.01	0.01	0.01	0
	$T = 8$	−0.05	−0.01	0	0.01	0.01	0
Panel B – Strategic complements		*Number of farmers ($\mathcal{N}$)*					
		2	3	4	5	6	8
Years ($\mathcal{T}$)	$T = 2$	0.1	0.22	0.25	0.28	0.33	0.34
	$T = 3$	0.16	0.29	0.34	0.33	0.3	0.27
	$T = 4$	0.25	0.28	0.25	0.18	0.13	0.09
	$T = 5$	0.25	0.18	0.12	0.06	0.05	0.02
	$T = 6$	0.16	0.13	0.05	0.03	0.02	0.01
	$T = 8$	0.13	0.06	0.01	0.01	0	0

References

W. Aeschbach-Hertig and T. Gleeson. Regional strategies for the accelerating global problem of groundwater depletion. *Nature Geoscience*, 5(12):853–861, 2012.

R. J. Aumann and L. S. Shapley. Long-term competition—A game-theoretic analysis. In N. Megiddo, editor, *Essays in Game Theory*, pages 1–15. Springer, New York, NY, 1994.

B. Baltagi. *Econometric Analysis of Panel Data*. John Wiley & Sons, Chichester, 2008.

S. Banzhaf, Y. Liu, F. Asche, and M. Smith. Non-parametric tests of the tragedy of the commons. Georgia State University: Working paper, 2018.

J. Barnard and D. B. Rubin. Small-sample degrees of freedom with multiple imputation. *Biometrika*, 86(4):948–955, 1999.

T. K. Beatty and I. A. Crawford. How demanding is the revealed preference approach to demand? *American Economic Review*, 101(6):2782–95, 2011.

L. Blume, W. Brock, S. Burlauf, and Y. Ioannides. Identification of social interactions. In J. B. A. Bisin and M. O. Jackson, editors, *The Handbook of Social Economics*, pages 853–964. North-Holland, Amsterdam, 2010.

G. E. Bolton and A. Ockenfels. ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193, 2000.

R. Boyd and P. J. Richerson. Collective actions, cultural norms. *Science*, 362(6420), 2018.

Y. Bramoulle, H. Djebbari, and B. Fortin. Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55, 2009.

J. I. Bulow, J. D. Geanakoplos, and P. D. Klemperer. Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political Economy*, 93(3):488–511, 1985.

M. A. Burke and H. P. Young. Social norms. In A. Bisin, J. Benhabib, and M. Jackson, editors, *The Handbook of Social Economics*. North-Holland, Amsterdam, 2010.

J. Carpenter and E. Seki. Do social preferences increase productivity? Field experimental evidence from fishermen in Toyama bay. *Economic Inquiry*, 49(2):612–630, 2011.

A. Carvajal, R. Deb, J. Fenske, and J. K.-H. Quah. Revealed preference tests of the Cournot model. *Econometrica*, 81(6):2351–2379, 2013.

P. M. Cashman and M. Preene. *Groundwater Lowering in Construction: A Practical Guide*. Spon Press, London, 2001.

G. Charness and M. Rabin. Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3):817–869, 2002.

G. Charness, L. Rigotti, and A. Rustichini. Individual behavior and group membership. *American Economic Review*, 97(4):1340–1352, 2007.

F. Clarke. *Optimization and Nonsmooth Analysis*. Wiley-Interscience, New York, 1983.

S. Clemhout and H. Y. Wan. Dynamic common property resources and environmental problems. *Journal of Optimization Theory and Applications*, 46(4):471–481, 1985.

S. Clemhout and H. Y. Wan. The nonuniqueness of Markovian strategy equilibrium: The case of continuous time models for nonrenewable resources. In T. Basar and A. Haurie, editors, *Advances in Dynamic Games and Applications*, pages 339–355. Birkhauser, Boston, 1993.

A. Cliff and J. Ord. *Spatial Autocorrelation*. Pion, London, 1973a.

A. Cliff and J. Ord. *Spatial Processes, Models and Applications*. Pion, London, 1973b.

K. Coman. Some unsettled problems of irrigation. *American Economic Review*, 1(1):1–19, 1911.

H. H. Cooper and C. E. Jacob. A generalized graphical method for evaluating formation constants and summarizing well-field history. *Transactions American Geophysical Union*, 27(4):526–534, 1946.

E. J. Dockner and N. Van Long. International pollution control: Cooperative versus noncooperative strategies. *Journal of Environmental Economics and Management*, 25(1):13–29, 1993.

P. K. Dutta. A folk theorem for stochastic games. *Journal of Economic Theory*, 66(1):1–32, 1995.

P. K. Dutta and R. K. Sundaram. The tragedy of the commons? *Economic Theory*, 3(3): 413–426, 1993.

C. C. Eckel and P. J. Grossman. Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3):371–392, 2005.

J. Famiglietti. The global groundwater crisis. *Nature Climate Change*, 4:945–949, 2014.

E. Fehr and S. Gaechter. Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994, 2000.

E. Fehr and S. Gaechter. Altruistic punishment in humans. *Nature*, 415(6868):137, 2002.

E. Fehr and A. Leibbrandt. A field study on cooperativeness and impatience in the tragedy of the commons. *Journal of Public Economics*, 95(9-10):1144–1155, 2011.

E. Fehr and K. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868, 1999.

E. Fehr and K. Schmidt. The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In S. Kolm and J. Ythier, editors, *Handbook of the Economics of Giving, Altruism and Reciprocity*, volume 1, pages 615–691. Princeton University Press, Princeton, 2006.

C. Fershtman and A. Pakes. Dynamic games with asymmetric information: A framework for empirical work. *Quarterly Journal of Economics*, 127(4):1611–1661, 2012.

U. Fischbacher, S. Gaechter, and E. Fehr. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3):397–404, 2001.

S. Foster and P. Chilton. Groundwater: The processes and global significance of aquifer degradation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358 (1440):1957–1972, 2003.

A. Friedman. The economics of the common pool: Property rights in exhaustible resources. *UCLA Law Review*, 18:855–887, 1971.

S. Gaechter and E. Fehr. Collective action as a social exchange. *Journal of Economic Behavior & Organization*, 39(4):341–369, 1999.

G. Gaudet, M. Moreaux, and S. W. Salant. Intertemporal depletion of resource sites by spatially distributed users. *American Economic Review*, 91(4):1149–1159, 2001.

G. Gigerenzer and R. Selten. Rethinking rationality. In G. Gigerenzer and R. Selten, editors, *Bounded Rationality: The Adaptive Toolbox*. MIT Press, Cambridge, US, 2001.

M. Giordano. Global groundwater? Issues and solutions. *Annual Review of Environment and Resources*, 34:153–178, 2009.

T. Gleeson, Y. Wada, M. F. Bierkens, and L. P. van Beek. Water balance of global aquifers revealed by groundwater footprint. *Nature*, 488(7410):197–200, 2012.

P. Gleick. Water management: Soft water paths. *Nature*, 418(6896):373–373, 2002.

P. Gleick. Global freshwater resources: Soft-path solutions for the 21st century. *Science*, 302 (5650):1524–1528, 2003.

P. Gleick and M. Palaniappan. Peak water limits to freshwater withdrawal and use. *Proceedings of the National Academy of Sciences*, 107(25):11155–11162, 2010.

H. S. Gordon. The economic theory of a common-property resource: The fishery. *Journal of Political Economy*, 62:124–142, 1954.

E. J. Green and R. H. Porter. Noncooperative collusion under imperfect price information. *Econometrica*, 52(1):87–100, 1984.

T. R. Green, M. Taniguchi, H. Kooi, J. J. Gurdak, D. M. Allen, K. M. Hiscock, H. Treidel, and A. Aureli. Beneath the surface of global change: Impacts of climate change on groundwater. *Journal of Hydrology*, 405(3-4):532–560, 2011.

E. D. Gutentag, F. J. Heimes, N. C. Krothe, R. R. Luckey, and J. B. Weeks. *Geohydrology of the High Plains Aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas and Wyoming*. Professional Paper 1400-B, U.S. Geological Survey, Alexandria, Virginia, 1984.

E. M. Haacker, A. D. Kendall, and D. W. Hyndman. Water level declines in the High Plains Aquifer: Predevelopment to resource senescence. *Groundwater*, 54(2):231–242, 2016.

G. Hardin. Tragedy of the commons. *Science*, 162(3859):1243–1248, 1968.

L. Huang and M. D. Smith. The dynamic efficiency costs of common-pool resource exploitation. *American Economic Review*, 104(12):4071–4103, 2014.

International Resource Panel (IRP). Groundwater Resources Outlook 2019: Natural resources for the future we want. Technical report, United Nations Environment Programme, Nairobi, Kenya, 2019.

M. O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, 2008.

M. O. Jackson. Couple issues to address conflict. *Science*, 362(6420), 2018.

A. Jaskiewicz and A. Nowak. Stochastic games of resource extraction. *Automatica*, 54: 310–316, 2015.

R. Jensen. The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *Quarterly Journal of Economics*, 122(3):879–924, 2007.

M. Kapoor, H. H. Kelejian, and I. R. Prucha. Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1):97–130, 2007.

H. Kelejian and I. Prucha. A generalized spatial two state least squares procedure for estimating a spatialautoregressive model with autoregressive disturbances. *Journal of Real Estate and Finance Economics*, 17(1):99–121, 1998.

H. Kelejian and I. Prucha. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40(2):509–533, 1999.

B. E. Kirwan. The incidence of US agricultural subsidies on farmland rental rates. *Journal of Political Economy*, 117(1):138–164, 2009.

L. F. Konikow. Contribution of global groundwater depletion since 1900 to sea-level rise.

Geophysical Research Letters, 38(17), 2011.

C. Koulovatianos. Strategic exploitation of a common-property resource under rational learning about its reproduction. *Dynamic Games and Applications*, 5(1):94–119, 2015.

D. G. Krige. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6):119–139, 1951.

A. Lara. Rationality and complexity in the work of Elinor Ostrom. *International Journal of the Commons*, 9(2):573–594, 2015.

D. Levhari and L. J. Mirman. The great fish war: An example using a dynamic Cournot-Nash solution. *Bell Journal of Economics*, 11(1):322–334, 1980.

D. K. Levine. Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1(3):593–622, 1998.

W. F. Lloyd. *Two Lectures on the Checks to Population*. Oxford University Press, Oxford, 1833.

C. Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–542, 1993.

D. Masclet. Ostracism in work teams: A public good experiment. *International Journal of Manpower*, 24(7):867–887, 2003.

G. Matheron. Principles of geostatistics. *Economic Geology*, 58(8):1246–1266, 1963.

V. L. McGuire, M. R. Johnson, R. L. S. J. S. Stanton, S. K. Sebree, and I. M. Verstraeten. *Water in Storage and Approaches to Groundwater Management, High Plains Aquifer, 2000*. U.S. Geological Survey, Reston, Virginia, 2003.

P. Milgrom and C. Shannon. Monotone comparative statics. *Econometrica*, 62(1):157–180, 1994.

L. J. Mirman and T. To. Strategic resource extraction, capital accumulation, and overlapping generations. *Journal of Environmental Economics and Management*, 50(2):378–386, 2005.

R. Moffit. Policy interventions, low-level equilibria, and social interactions. In H. Young and S. Durlauf, editors, *Social Dynamics*, pages 6–17. MIT Press, Cambridge, 2001.

Y. Mundlak. On the pooling of time series and cross section data. *Econometrica*, 46(1): 69–85, 1978.

J. Mutl and M. Pfaffermayr. The Hausman test in a Cliff and Ord panel model. *The Econometrics Journal*, 14(1):48–76, 2011.

H. H. Nax, R. O. Murphy, and K. A. Ackermann. Interactive preferences. *Economics Letters*, 135:133–136, 2015.

Nebraska Department of Agriculture. *Nebraska Agriculture: Agriculture Facts Brochure*. 2014. URL `http://www.nda.nebraska.gov/publications/ne_ag_facts_brochure.pdf`.

D. H. Negri. The common property aquifer as a differential game. *Water Resources Research*, 25(1):9–15, 1989.

J. Newton. Shared intentions: The evolution of collaboration. *Games and Economic Behavior*, 104:517–534, 2017.

C. N. Noussair, D. van Soest, and J. Stoop. Cooperation in a dynamic fishing game: A framed field experiment. *American Economic Review*, 105(5):408–13, 2015.

E. Ostrom. *Public Entrepreneurship: A Case Study in Groundwater Basin Management*. PhD thesis, University of California, Los Angeles, 1965.
E. Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, Cambridge, 1990.
E. Ostrom. A behavioral approach to the rational choice theory of collective action. *American Political Science Review*, 92(1):1–22, 1998.
E. Ostrom. Collective action and the evolution of social norms. *Journal of Economic Perspective*, 14(3):137–158, 2000.
E. Ostrom. Common-pool resources and institutions: Toward a revised theory. In B. Gardner and G. Rausser, editors, *Handbook of Agricultural Economics*. Elsevier Science, 2002.
E. Ostrom. How types of goods and property rights jointly affect collective action. *Journal of Theoretical Politics*, 15(3):239–270, 2003.
E. Ostrom. *Understanding Institutional Diversity*. Princeton University Press, Princeton, 2005.
E. Ostrom. Institutional rational choice: An assessment of the institutional analysis and development framework. In P. Sabatier, editor, *Theories of the Policy Process*, pages 21–64. Westview Press, Boulder, 2007.
E. Ostrom. Design principles of robust property rights institutions: What have we learned? In G. Ingram and Y.-H. Hong, editors, *Property Rights and Land Policies*, pages 25–51. Lincoln Institute of Land Policy, Cambridge, 2009.
E. Ostrom. Reflections on “Some unsettled problems of irrigation". *American Economic Review*, 101(1):49–63, 2011.
A. R. Poteete and E. Ostrom. Fifteen years of empirical research on collective action in natural resource management: Struggling to build large-N databases based on qualitative research. *World Development*, 36(1):176–195, 2008.
J. K.-H. Quah and B. Strulovici. Comparative statics, informativeness, and the interval dominance order. *Econometrica*, 77(6):1949–1992, 2009.
M. Rabin. Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302, 1993.
J. F. Reinganum and N. L. Stokey. Oligopoly extraction of a common property natural resource: The importance of the period of commitment in dynamic games. *International Economic Review*, pages 161–173, 1985.
J. J. Rotemberg and G. Saloner. A supergame theoretic model of price wars. *American Economic Review*, 76:390–407, 1986.
D. Rubin. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, New York, 1987.
D. Rustagi, S. Engel, and M. Kosfeld. Conditional cooperation and costly monitoring explain success in forest commons management. *Science*, 330(6006):961–965, 2010.
A. Scott. The fishery: The objectives of sole ownership. *Journal of Political Economy*, 63 (2):116–124, 1955.
R. Selten. Properties of a measure of predictive success. *Mathematical Social Sciences*, 21: 153–167, 1991.
R. Selten and S. Krischker. Comparison of two theories for characteristic function experi-

ments. In R. Tietz, editor, *Aspiration Levels in Bargaining and Economic Decision Making*, pages 259–264. Springer, Berlin and Heidelberg, 1983.

L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10): 1095–1100, 1953.

V. L. Smith. On models of commercial fishing. *Journal of Political Economy*, 77(2):181–198, 1969.

J. Sobel. Interdependent preferences and reciprocity. *Journal of Economic Literature*, 43(2): 392–436, 2005.

R. M. Solow. A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1):65–94, 1956.

G. Sorger. A dynamic common property resource problem with amenity value and extraction costs. *International Journal of Economic Theory*, 1(1):3–19, 2005.

C. P. Spalding and R. Khaleel. An evaluation of analytical solutions to estimate drawdowns and stream depletions by wells. *Water Resources Research*, 27(4):597–609, 1991.

R. N. Stavins. The problem of the commons: Still unsettled after 100 years. *American Economic Review*, 101(1):81–108, 2011.

K. Stephenson. Groundwater management in nebraska: Governing the commons through local resource districts. *Natural Resources Journal*, 36:761–778, 1996.

D. R. Steward and A. J. Allen. Peak groundwater depletion in the High Plains Aquifer, projections from 1930 to 2110. *Agricultural Water Management*, 170:36–48, 2016.

J. Stoop, C. N. Noussair, and D. Van Soest. From the lab to the field: Cooperation among fishermen. *Journal of Political Economy*, 120(6):1027–1056, 2012.

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

C. V. Theis. The effect of a well on the flow of a nearby stream. *Transactions American Geophysical Union*, 22(3):734–738, 1941.

D. M. Topkis. Minimizing a submodular function on a lattice. *Operations Research*, 26(2): 305–321, 1978.

L. E. Torres-Guevara and A. Schlueter. External validity of artefactual field experiments: A study on cooperation, impatience and sustainability in an artisanal fishery in Colombia. *Ecological Economics*, 128:187–201, 2016.

M. L. von Stein. *Interpolation of spatial data: Some theory for kriging*. Springer, 2012.

M. Voors, E. Bulte, A. Kontoleon, J. A. List, and T. Turley. Using artefactual field experiments to learn about the incentives for sustainable forest use in developing economies. *American Economic Review*, 101(3):329–33, 2011.

Y. Wada, L. P. Beek, F. C. Sperna Weiland, B. F. Chao, Y.-H. Wu, and M. F. Bierkens. Past and future contribution of global groundwater depletion to sea-level rise. *Geophysical Research Letters*, 39(9), 2012.

F. Wirl. Do multiple Nash equilibria in Markov strategies mitigate the tragedy of the commons? *Journal of Economic Dynamics and Control*, 31(11):3723–3740, 2007.

F. Wirl. Tragedy of the commons in a stochastic game of a stock externality. *Journal of Public Economic Theory*, 10(1):99–124, 2008.

H. P. Young. The evolution of conventions. *Econometrica*, 61(1):57–84, 1993.

H. P. Young. The evolution of social norms. *Annual Review of Economics*, 7(1):359–387, 2015.

H. P. Young and M. A. Burke. Competition and custom in economic contracts: A case study of Illinois agriculture. *American Economic Review*, 91:559–573, 2001.