5-2019

Three Essays on Land Property Rights, Water Trade, and Regional Development

Muyang Ge
Utah State University

THREE ESSAYS ON LAND PROPERTY RIGHTS, WATER TRADE, AND REGIONAL DEVELOPMENT

by

Muyang Ge

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Economics

Approved:

Eric C. Edwards, Ph.D. Co-major Professor	Ryan Bosworth, Ph.D. Co-major Professor
Reza Oladi, Ph.D. Committee Member	Kynda Curtis, Ph.D. Committee Member
Joanna Endter-Wada, Ph.D. Committee Member	Richard S. Inouye, Ph.D. Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2019

ABSTRACT

Three Essays on Land Property Rights, Water Trade, and Regional Development

by

Muyang Ge, Doctor of Philosophy

Utah State University, 2019

Major Professor: Eric C. Edwards, Ph.D. and Ryan Bosworth, Ph.D.
Department: Applied Economics

This dissertation explores how property rights to a natural resource affect economic decisions for investment or sale, and how these decisions may, in turn, impact other areas of the economy. The first essay focuses on how incomplete land ownership on Indian Reservations in the United States affects landowner incentives to engage in agricultural production. Utilizing a regression discontinuity design, we find that incomplete land ownership, where tribal lands are held in trust by the US government, creates significant barriers to the acquisition of capital for agricultural investment, including investment in efficient irrigation systems. As a result, we show less high-value agriculture occurs on these lands. The second essay explores how the transfer of water in arid regions via water right sales affects local labor markets and environmental outcomes. We develop a general-equilibrium representation of a hydrologic-ecological-economic system to understand the labor market and environmental effects of water trade. To explore the problem empirically, we examine the water transfer from the Imperial Irrigation District to the City of San Diego. Using a synthetic counterfactual approach, we find a decline in the number of low- and high-skill jobs in Imperial County corresponding to the water transfer, as well as a decrease in overall crop production, as predicted by the theoretical model. The loss of jobs and environmental

benefits as a result of transfers suggests why local communities often oppose water transfers. In the third essay, we seek to understand how shale-gas drilling has affected organic food production. Using an instrumental variable estimate of a survival function, as well as a joint model with time-dependent covariates, we obtain causal estimates of the effect of shale development externalities on organic farming certification in Colorado. Organic farms near gas wells see a small but significant increase in the probability of reducing organic production. The results suggest that real or perceived contamination concerns from gas wells impact the producer choice to engage in organic production.

(184 pages)

PUBLIC ABSTRACT

Three Essays on Land Property Rights, Water Trade, and Regional Development

Muyang Ge

This dissertation explores how property rights to a natural resource affect economic decisions for investment or sale, and how these decisions may in turn impact other areas of the economy. The first essay focuses on how incomplete land ownership on Indian Reservations in the United States affects landowner incentives to engage in agricultural production. The second essay explores how the transfer of water in arid regions via water right sales affects local labor markets and environmental outcomes. The third essay seeks to understand how shale-gas drilling has affected organic food production. This dissertation provides several policy implications. First, the findings suggest that the key to improving lagging agricultural development on American Indian land is to improve tribal farmers' access to capital, so they can invest in agricultural systems (including irrigation) at the level of their neighbors enjoying fee-simple title. Second, while a potentially effective solution to reduce costly water shortfalls among high-value urban users, water sales from agricultural to urban users appear to simultaneously decrease employment and environmental quality in the water exporting region. Third, Drilling activities appear to discourage organic farming in Colorado. While farmers with mineral ownership benefit, identifying the direct causes of lost organic certification can inform policy that regulates negative externalities on organic farms caused by drilling.

This dissertation is dedicated to my Dad and in loving memory of my Mom with my deepest love for always loving and supporting me.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest appreciation and thanks to my and committee co-chair, Dr. Eric C. Edwards for providing me encouragement and support throughout my study, for his patience, motivation, and immense knowledge. His guidance helped me shaping my research and writing of this dissertation. I could not have imagined having a better advisor and mentor for my Ph.D. study. Without his precious support and advice, it would not be possible to finish my dissertation so quickly and smoothly. Besides my advisor, I would like to thank the rest of my dissertation committee: Dr. Ryan Bosworth, Dr. Reza Oladi, Dr. Kynda Curtis and Dr. Joanna Endter-Wada, for their comments and suggestions that helped me to polish my dissertation. In particular, I want to give my great thankfulness to Dr. Ryan Bosworth, for his continuous service as my co-chair in my fourth year. My sincere thanks also go to Dr. Sherzod Akhundjanov, who provided me valuable advice related to the econometric part of my first essay.

I also want to express my thanks to the Climate Adaptation Science (CAS) program for providing me the opportunity to intern with the USGS Southwest Biological Science Center. I am grateful to know many wonderful students in the CAS program. Our interdisciplinary team, Emily Esplin, Natalie Gillard, Liana Prudencio, Ryan Choi, and Jeffery Haight, you have been great friends with whom I have a fantastic working experience over the past two years. I also want to thank my mentor, Dr. Courtney Flint and Dr. Patrick Belmont for their insightful comments on our team project.

Many of my friends helped me with my Ph.D. study. I would like to thank all my friends in China and the U.S. for their support. In particular, I want to express gratitude to my friend Haomin Zhang for always providing technical support of data analysis that was critical for my research. I also want to thank the wonderful students I met in the Applied Economic Department, Ahsan Kibria, Ramjee Acharya, and Tatiana Drugova, for their help and companionship in the past four years.

Lastly, I want to give my greatest thankfulness to my lifetime mentor, my Dad. His support and guidance make me mentally stronger and encourages me to become a better person in my life.

Muyang Ge

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AF	acre-foot
AMS	Agricultural Marketing Service
ARMS	Agricultural Resource Management Survey
BEA	Bureau of Economic Analysis
BIA	Bureau of Indian Affairs
CCAC	California County Agricultural Commissioners
CDL	Cropland Data Layer
CVWD	Coachella Valley County Water District
EIA	Energy Information Administration
EPA	Environmental Protection Agency
GIS	Geographic Information System
IID	Imperial Irrigation District
ILTF	Indian Land Tenure Foundation
IV	instrumental variable
LEHD	Longitudinal Employer-Household Dynamics
MSPE	Mean of the Squared Deviations
MWD	Metropolitan Water District
NOP	National Organic Program
OID	Organic Integrity Database
PI	Productivity Index
PLSS	Public Land Survey System
QSA	Quantification Settlement Agreement
RD	Regression Discontinuity
SDCWA	San Diego County Water Authority
SGID	State Geographic Information Database
RMSPE	Root Mean Square Percentage Error
SRTM	Shuttle Radar Topographic Mission

TERASs	Tribal Energy Resource Agreement
USDA	United State Department of Agriculture
USFS	United State Forest Service
USGS	United State Geological Survey
WRL	Water Related Land

CHAPTER 1

INTRODUCTION

Natural resources such as land, water, oil, and gas are often characterized by incomplete ownership and subject to common-pool losses. For instance, land held in common may affect incentives for investment; water in a common-pool might be over-extracted; and oil and gas extraction may result in environmental externalities. The focus of this dissertation is to charactize these three examples of incomplete property rights in natural resource extraction, and then estimate the economic effect of the current property right institution. The dissertation is organized into three essays. Each essay first characterizes the institutional setup that creates incomplete ownership and links this context to observable economic outcomes. Then, each essay establishes a credible set of counterfactual outcomes for comparison. Each essay makes its contribution to the literature by using econometric techniques novel to the application at hand.

Recent studies have discussed the underlying causes for limited Native American economic development by studying the relationship between insecure property rights and poverty on American Indian land (Anderson and Lueck 1992; Cornell and Kalt 2000; Anderson and Parker 2008). The first essay extends the literature to an analysis of agricultural irrigation. This essay uses the case of the Uintah and Ouray Indian Reservation in eastern Utah to explore how tribal trust land ownership affects agricultural development on reservation land. A spatial Regression Discontinuity (RD) approach is used in the empirical analysis to identify the causal effect of weak tribal institutions on agricultural investment. The empirical framework of this paper can be divided into two sections. In the first section, the sharp RD approach with the 1905 historical allotment boundary is applied to explore whether the insecure tribal trust land ownership will affect the agricultural development inside the historical allotment boundary. In the second section, the 2017 tribal landownership is used to apply the fuzzy RD approach. The spatial RD approach has been widely applied

to many institutional settings but to our knowledge it has not been used to examine trust land ownership. Moreover, this paper adds to current RD literature by examining both sharp and fuzzy RD together. We develop a new dataset by linking agricultural irrigation choice, land ownership and historical land allocation. The data construction procedure in this paper provides an alternative solution to micro-level dataset construction for research topics with difficulties in obtaining micro-level dataset.

Another essential resource, water, plays an important role in agricultural production. The second essay examines the economic and environmental impact of a water transfer agreement, the Quantification Settlement Agreement, in California, which began in 2004. Water stress in the arid region is increasing due to increasing urban water demand. Water reallocation between different regions and sectors has become one of the solutions to address the need to meet urban water demand. Few studies have explored the effect of cross-sectoral water transfer in different regions, but effects are varied (Brooks and Harris 2008; Cai 2008; Juana, Strzepek, and Kirsten 2011; Wimmer et al. 2015). This study extends current literature to a case study of the United States' largest ever ag-to-urban water transfer.

A water transfer between agricultural and urban areas benefits the parties who are directly involved, but the impact on other parties is hotly debated (Howe, Lazo, and Weber 1990; Holcombe and Sobel 2001; Hadjigeorgalis and Lillywhite 2004; Mann and Wüstemann 2008; Grafton et al. 2011). We develop a general equilibrium representation of a hydrologic-ecological-economic system to explore the theoretical effect of trading water on agricultural production and employment in the water exporting region. In our empirical framework, we apply the synthetic control methodology to test the predictions of our theoretical model. This is the first study that links general equilibrium with empirical results in examining the efficiency of a regional water agreement between agricultural and urban sector. Our theoretical and empirical analyses together suggest that a decline in water availability may cause both reductions in employment and environmental damage. However, the increased value of water will be captured by parties directly involved in the transfer. In this paper, we look at the increased return in terms of direct payment from water trading in the water

exporting region.

In addition to agricultural inputs, such as land and water, environmental externalities may affect agricultural production. The rising concern about hydraulic fracturing has been documented recently (Allred et al. 2015; Vidic et al. 2013; McKenzie et al. 2012; Rakitan 2018; Muehlenbachs, Spiller, and Timmins 2015), but surprisingly few studies have focused on the effect of shale development on agriculture (Weber, Brown, and Pender 2013; Hitaj, Boslett, and Weber 2017; Farah 2017). The third essay adds to current literature by studying the environmental externality caused by oil and gas extraction in the state of Colorado.

We created a novel geospatial dataset with organic farm locations and certification length in the United States. By plotting the organic farm and fracking well locations, Colorado is selected as the target state to explore whether hydraulic fracturing affects maintaining organic certification. Two empirical models were used in this essay: Instrumental Variable Estimation in a Survival Analysis and a Joint Model with an endogenous time-dependent variable. An Instrumental Variable Estimation in a Survival Analysis context is used to solve the endogeneity problem caused by a lack of correspondence between oil and gas deposits and suitable agricultural land. Then, the exposure to fracking wells is treated as a time-dependent variable using the Joint Model to more accurately explore the impact of fracking on maintaining organic certification. To our knowledge, this study is the first study using survival analysis methodology to explore the impact of hydraulic fracturing on organic farming.

The dissertation is organized as follows. Chapter 2, 3, and 4 display three essays as discussed in the introduction. Each essay has a separate introduction and conclusion. Chapter 5 summarizes the dissertation and discusses the possible policy implications of the research contained in the dissertation. References for each essay are collected together and provided after Chapter 5.

CHAPTER 2

LAND OWNERSHIP AND IRRIGATION ON AMERICAN INDIAN RESERVATIONS

2.1 Abstract

American Indian reservations are often characterized by low income and high rates of poverty relative to adjacent non-reservation land. To understand the role institutions governing land ownership play in these outcomes, we examine agricultural land use and irrigation on parcels on and adjacent to the Uintah-Ouray Indian Reservation in eastern Utah. Land within the reservation is held in trust by the federal government and has significant restrictions on its use and development. We predict that this land will see lower investment in irrigation and therefore lower agricultural productivity. We use the exogenous allocation boundaries of a 1905 land allotment as a natural experiment, employing both a sharp and a fuzzy regression discontinuity (RD) design to explore how land ownership has affected agricultural land use, irrigation levels, and irrigation investment. Our results suggest that the original allocations provided land of similar quality across the border. Despite this, tribal lands are around 18 percentage points less likely to be irrigated today, and conditional on being irrigated, tribal land has a 31 percentage point lower rate of capital-intensive sprinkler irrigation. Tribal land is also less likely to grow high-value crops. These results suggest that trust ownership creates significant barriers to the acquisition of capital for agricultural investment, and helps explain lagging agricultural development on reservations.

2.2 Introduction

The link between insecure property rights and poverty on American Indian reservations has drawn significant attention in recent years. The median household income for American Indian communities in 2016 was $38,502 while the estimate of the U.S. as a whole was

\$55,322.[1] This divergence is even more pronounced in terms of agricultural production. In 2007, the average American Indian farm saw sales of \$40,331, less than 1/3 of the US average.[2] Previous studies have traced the underlying causes for limited tribal development to weak institutions as a result of both tribal and federal policies (Anderson and Lueck 1992; Cornell and Kalt 2000; Anderson and Parker 2008). We extend this literature to an analysis of agricultural irrigation. With 75% of land in Indian country dedicated to agriculture, understanding how institutions affect the productivity is key to improving economic development on reservations (Shoemaker 2006, p.11).

In this paper, we use the case of the Uintah and Ouray (Uintah) Reservation in eastern Utah to explore how institutions have affected the pattern of agricultural development. The Uintah Reservation is the second largest by area in the United States and, like many reservations, its current area has been reduced significantly over time. Important to this paper, the tribe was ultimately allotted a few contiguous blocks of land in 1905 via the Dawes Act, with the remaining portions of the reservation opened to white settlement. Within this allocation, some land was claimed as fee-simple by tribal members while the unclaimed land reverted to tribal control as federal trust land in 1937. Fee-simple owners have complete property rights and can freely sell or lease the land. In contrast, tribal land sales are restricted and require the review of both the tribal government and the Bureau of Indian Affairs (BIA). Throughout the paper we define tribal land as any land or interest in land owned by a tribe or tribes, title to which is held in trust by the United States or is subject to a restriction against alienation under the laws of the United States.[3] The lack of land use flexibility and the inability of lenders to enforce contracts on reservations results in a lack of access to commercial credit, limiting the opportunities to borrow money for capital-intensive improvements (Anderson and Lueck 1992).

In this study, we apply a spatial regression discontinuity (RD) approach to identify the effect of tribal ownership on agricultural development. Specifically, we utilize the straight-line boundary of the 1905 allocation, both directly and as an instrument on current land

[1] U.S. Census Bureau, 2012-2016 American Community Survey 5-Year Estimates.
[2] Data is from Census of Agriculture 2012
[3] Definition is from Tribal Energy Resource Agreement (TERAs).

ownership, to identify the effect of trust ownership on irrigation and irrigation investment. Land ownership changed discretely at the straight-line boundary in 1905 at the time of allocation. On one side, all the lands were under tribal trust, while on the other side, all the lands were fee-simple. The spatial RD approach has been widely applied to a variety of institutional settings (see, for instance, Bayer, Ferreira, and McMillan 2007; Dell 2010; Grout, Jaeger, and Plantinga 2011; Dachis, Duranton, and Turner 2011; Dell 2015; Card and Giuliano 2016; Pan, Smith, and Sulaiman 2018), but to our knowledge has not been used to examine trust land ownership.

We develop a new dataset by linking agricultural irrigation choice, land ownership data, and historic land allocation. We implement a sharp RD approach with local polynomial regression to examine the impacts of current agricultural choices across the 1905 allotment boundary. However, since 1905 some land has changed hands, so the assignment of the treatment today may be based on additional variables that are unobserved. Selection into treatment is dependent on both observable and unobservable factors, and we therefore expect the boundary of 2017 land ownership to be a "fuzzy" rather than "sharp" discontinuity. To address this issue we utilize a sharp RD design on the 1905 boundary excluding all lands which have switched ownership, and then implement a fuzzy RD design. This approach treats the 1905 boundary as an instrument for current land ownership and rescales the observed effect of the discontinuity based on the probability of receiving treatment using a nonparametric local linear (polynomial) estimator.

We find that tribal lands have irrigation rates around eighteen percentage points lower under the instrumented 2017 tribal land ownership. Further, tribal lands see significantly less investment in capital-intensive irrigation systems, with irrigated tribal land seeing 31 percentage point lower rates of sprinkler irrigation. Tribal land is also less likely to grow high-value crops. On the lands that did not change hands, the sharp RD results show that tribal lands have 13 percentage points less investment in sprinkler irrigation systems, and are less likely to grow high-value crops. These results suggest that tribal trust ownership inhibits agricultural production and irrigation investment on the reservation. While there

is anecdotal evidence the difficulties are related to insecure land tenure, we also discuss several alterative explanations for the underinvestment in tribal agriculture.

The paper proceeds as follows: section two provides background on tribal land allocation and the Uintah Reservation. Section three describes an economic framework for the effect of insecure land tenure and provides predictions. Section four provides details on the empirical design and econometric approach. The econometric results are provided in section five and section six concludes.

2.3 Background

2.3.1 Reservation Land Ownership

American Indian Reservations were formed from territory controlled by the United States government to provide an area of settlement for previously autonomous tribes. Initially, reservation allocations were to tribes, but as land pressure increased, the US Congress in 1887 passed the Dawes Act which allowed the government to allocate land within the reservations to individuals. Reservation areas had portions reserved for allocation to tribal members and the remaining land was opened for white settlement. In the allocated areas, individual tribal members could make a claim to own land individually. The 1934 Indian Reorganization Act again changed the rules and the unclaimed allotment areas reverted to tribal control. The Act resulted in the three categories of land ownership we see on reservations today: fee-simple, land which is privately owned; tribal trust, land allocated to tribes under the Dawes Act but which was never claimed by tribal members and reverted to tribal control; and individual trust, which is allocated land that was claimed by individuals but for which the process of transitioning to fee-simple was never completed.

Trust land has significantly different constraints on land trade and alienation, relative to fee-simple. While the owner of the private land can freely sell or lease the land, tribal trust land is owned by the federal government and managed jointly by tribal governmental organizations and the Bureau of Indian Affairs (BIA). BIA maintains ownership records and manages almost any transaction involving trust land. Trust property cannot be transferred,

alienated, or leased without the approval of the BIA. These approvals typically require long appraisal and documentation processes. In 2003, the Indian Land Tenure Foundation (ILTF) conducted a community survey to measure the view of Indian peoples on land ownership and managemraent. It found perceptions of systematic barriers in the use of property rights related to land and natural resources, especially the slowness of BIA actions. Specifically, that the federal bureaucracy is unable to provide legal certainty or act quickly and is insensitive to traditional ways and knowledge.[4] Anderson and Lueck (1992) found that trust land constraints imposed by the federal government significantly reduced the value of agricultural output on reservation land.

Individual or tribal trust land may be mortgaged with the consent of the landowners and the BIA. However, many private commercial lending difficulties exist on trust lands. First, individuals seldom own direct title and therefore do not have collateral. Second, it is nearly impossible to get title insurance on Indian trust land because only a few title insurance companies are qualified to offer it. Loans secured by trust land still require BIA approval, and there is no uniform approval process for different BIA offices.[5]

2.3.2 Indian Agriculture

The potential for jurisdictional uncertainty creates complexity and reduces access to credit for Indian farmers and ranchers. Even though the tribe functions as a sovereign entity according to the governing by-laws, the U.S. Secretary of Interior has final authority over many tribal actions. Agricultural land leases are an example. Agricultural leases may be negotiated directly with the landowner, often the tribal government, but they are still subject to BIA approval. Tribal leases are subject to the National Environmental Policy Act, which applies to federal agencies but not private fee-simple sales or leases (Shoemaker 2006, p.13). Leases are codified as having a maximum duration of 10 years, unless substantial

[4] Indian Land Tenure Foundation (ILTF). 2003. "Community Survey: Importance of Land and Value of Property Rights." URL:https://iltf.org/wpcontent/uploads/2016/11/community_survey_2003.pdf

[5] Information is summarized from U.S. Department of Treasury. 2006. Guide to Mortgage Lending in Indian Country.

investment is required, in which case 25-year leases are possible.[6]

Indian farmers have faced difficulties and discrimination in accessing USDA loans. Evidence suggests that the USDA systematically discriminated against Indian farmers by denying them credit they routinely offered to white farmers under the USDA Farm Loan Program. A class-action lawsuit encompassing the period 1981-1999 (Keepseagle v. Vilsack) was settled in 2010 with a $760 million payment to affected Indian farmers. USDA has traditionally been the largest single lender to Indian farmers and ranchers (Shoemaker 2006, p.22). Discrimination in access to credit is one potential explanation for lagging agricultural development. Tribes have also argued that crop insurance products offered by USDA are not well-suited for the agricultural practices of tribal farmers and that tribal farms may not qualify for federal disaster assistance.[7]

Another potential limit to the development of irrigated agriculture on tribal land is problematic access to federal irrigation projects. Reservations are primarily located in arid regions, and the BIA operates 16 irrigation projects. In 2006, the General Accounting Office criticized the operation of these projects due to deferred maintenance, a lack of managerial expertise in water systems, and uncertainty over financial sustainability. Because irrigation management is not a priority for BIA, the report concludes that it might be beneficial if an agency like the Bureau of Reclamation, which provides water for non-tribal farmers, managed these projects(GAO 2006, p.28).

2.3.3 Uintah and Ouray Reservation

The Uintah reservation was established for the native people of eastern Utah as a combined reservation in 1886 (Cuch 2000, p.196). The passage of the Dawes Act in 1887 started the process by which significant portions of the reservation were reallocated to private individuals. Six years later Congress passed another Indian Appropriations Act and set a timeline for the BIA to acquire an agreement with the tribe on their land allotment.

[6] This information is summarized from 25 U.S. Code § 3715 - Leasing of Indian agricultural lands. URL: https://www.law.cornell.edu/uscode/text/25/3715

[7] https://www.tribalselfgov.org/wp-content/uploads/2018/03/Farm-Bill.pdf;http://www.ncai.org/NFBC_Policy_Recommendations.pdf

The reservation was allotted in 1905 and entry by settlers onto the unreserved and unallotted lands occurred after that time. Under the allotment policy, adult members of the Uintah tribe received allotment lands between 40 and 640 acres, depending on the suitability of the land for farming. This property was subject to a protected status that forbade it being sold by the individual for twenty-five years, at the end of which time the owner would be recognized as an American citizen (McPherson 2000, p.22).

In 1906 the federal government authorized construction of the Uintah Indian Irrigation Project, which provided water to both Indian and non-Indian farmers in the area. Within fifteen years of the allotment, tribal members had sold or leased 30,000 acres of Uintah land, much of which was then irrigated by non-Indian farmers (Cuch 2000, p.207). In 1937, under the 1934 Indian Reorganization Act, all tribal lands that had not been privatized reverted to Uintah control. Today, this land is held in tribal trust and the U.S. Secretary of the Interior must approve many Uintah tribal actions, which hinders the tribe's ability to create economic growth (Cuch 2000, p.222). "Even though the Ute Tribe is one of the major economic contributors to Uinta Basin and the state, the tribe experiences the lingering problems associated with having been proclaimed sovereign yet not being treated as such by county, state, and federal entities. This creates disputes between the tribe and these bodies of government over issues such as jurisdiction, double taxation, rights-of-way, and water rights (Cuch 2000, p.221)."

Today, the Uintah reservation is the second-largest US Indian reservation in land area. Figure 2.1 shows the allocation of land within the reservation. Federal lands located around the northern and western boundaries of the Uintah and Ouray Indian reservation are primarily national forest in the Uintah Mountains. In the agricultural areas, tribal trust and private fee-simple land are the primary ownership types. Uintah tribal bylaws limit land leases to a period of five years, although exceptions may be made for irrigable land.[8]

The area around the Uintah Reservation is arid, with the agricultural areas receiving approximately 270mm of precipitation per year. There are thirty-two different crops grown

[8] Constitution and By-Laws of the Ute Indian Tribe of the Uintah and Ouray Reservation Article VI(1)(c).

in the area, but the majority of acreage is in alfalfa. The average irrigation rate within two miles of the tribal boundary is around 39.7% on fee-simple land versus 22.8% on tribal land. Within the two-mile window, only 7.3% of irrigated tribal land uses sprinkler irrigation, compared to 31.2% of private irrigated land. We now turn to an analytic framework to demonstrate how insecure property right institutions could cause tribal lands to differ in their investment in irrigation.

2.4 Economic Framework and Predictions

Previous research on property rights and investment (Demsetz 1974; Besley 1995; Anderson and Parker 2008) suggests there are multiple channels through which land property rights affect agricultural investment. We adapt Besley's model to our case.

Consider a farmer who invests c amount of capital in his/her farm. The revenue function of investment can be written as $R(c,x)$ where x represents land property rights now and in the future; x increases as the land property rights become stronger. $R(\cdot)$ is assumed to be an increasing function of c and x, and concave in c. $C(c,x)$ represents the cost of investment and it is an increasing function of c and non-increasing function of x. The optimal investment choice is then given by:

$$\max_{c} I(c,x) = R(c,x) - C(c,x) \tag{2.1}$$

The first order condition for the choice of capital investment, c, is:

$$I_1(c,x) = 0 \tag{2.2}$$

Taking the total derivative of the first order condition in equation 2.2, we get:

$$\frac{\partial c}{\partial x} = -\frac{I_{12}(c,x)}{I_{11}(c,x)} \tag{2.3}$$

Because of the concavity of the investment function, $I(\cdot)$, the maximum point exists if $I_{11} < 0$. Importantly, if $I_{12} > 0$, it implies a positive relationship between agricultural

investment and land property rights. We will discuss how land property rights could affect agricultural investment through three different channels.

The first channel is freedom from expropriation (Demsetz 1974; Alchian and Demsetz 1973). That is, a farmer does not have incentive to invest in his/her land if it could easily be seized by others. Suppose the probability of losing farmland in the future is $p(x)$, where $p(x)$ is between zero and one, and decreases as property rights increase. The direct return from farming is defined as $R_p(c)$. Then, the maximization of the expected return for the farmer is:

$$\max_c R(c,x) = (1-p(x)) \times R_p(c) + p(x) \times 0 \tag{2.4}$$

$R_{12}(c,x)$ can be calculated by taking the derivative of $R_1(c,x)$ with respect to x:

$$R_1(c,x) = R_p^{'}(c) - p(x) \times R_p^{'}(c) \tag{2.5}$$

$$R_{12}(c,x) = -p^{'}(x) \times R_p^{'}(c) > 0 \tag{2.6}$$

Since $I_{12}(c,x) = R_{12}(c,x) - C_{12}(c,x)$ and $C(c,x)$ is a non-increasing function of x by assumption, it is straightforward to conclude that $I_{12}(c,x) > 0$.

The second channel is using land as collateral: secure land property rights reduce the interest rate. Lower interest rates increase land investment because the interest rate is equal to the required marginal productivity of capital investment (Feder and Feeny 1991; Besley 1995). Suppose a farmer would like to borrow money from a lender to invest in a sprinkler system. We assume the initial wealth of the farmer is 0. The money borrowed from the lender is defined as b. The lender charges an interest rate of $r(x)$. We assume that interest rate is negatively correlated with the land property rights, $\frac{\partial r(x)}{\partial x}$. The probability of earning the return is q. The physical return from the new sprinkler system is $R_p(c)$, $R_p^{'}(\cdot) > 0$ and $R_p^{''}(\cdot) < 0$. The utility function $u(\cdot)$ is a smooth, concave and increasing function. Thus,

the farmer's expected utility function can be written as:

$$(2.7)\qquad I(c,x) = \max_{b,c} u(b-c) + qu(R_p(c) - r(x) \times b) + (1-q) \times 0$$

The first order condition with respect to the choice variables b, c can be specified as:

$$(2.8)\qquad -u^{'}(b-c) + qu^{'}(R_p(c) - r(x) \times b) \times R_p^{'}(c) = 0$$

$$(2.9)\qquad u^{'}(b-c) + qu^{'}(R_p(c) - r(x) \times b) \times -r(x) = 0$$

It is straightforward to show that:

$$(2.10)\qquad R_p^{'}(c) = r(x)$$

The first order condition for the choice of c, after the envelope theorem is used for the choice of b, can be written as:

$$(2.11)\qquad I_1(c,x) = R_p^{'}(c) - r(x)$$

Solving equation 2.10 and 2.11 simultaneously, we obtain:

$$(2.12)\qquad I_{12}(c,x) = -\frac{\partial r(x)}{\partial x}$$

Equation 2.10 implies that at the maximum utility, the marginal productivity of capital invested on an Indian farmland is equal to the interest rate charged from a lender. Since we assume a negative relationship between land property rights and interest rate, $\frac{\partial r(x)}{\partial x}$, we can conclude that $I_{12}(c,x) > 0$.

The third channel comes from the intuition that better transfer rights reduce the land transfer cost and increase investment incentives. We assume that the trading cost is dependent on a farmer's transfer rights. Suppose the sale price of the land is p. If the farmer sells the land, the best offer available is w, which has the density function of $g(w)$, $w \in [\underline{w}, \overline{w}]$. If the Indian farmer decides to use the land, his/her payoff is δc, where c is his/her return

to investment and δ is the marginal product of capital, which has the density function of $f(\delta)$, $\delta \in [\underline{\delta}, \overline{\delta}]$. The trading cost, defined as $\pi(x)c$, is a decreasing function of x, and $\pi'(x)$ is less than zero. Then, the optimal land price under a Nash bargaining solution is:

$$\max_p (p - \pi c - \delta c)(wc - p) \tag{2.13}$$

Solving equation 2.13, we get the optimal land price, $p^* = \frac{\pi+\delta+w}{2}c$. Hence, the farmer's payoff from selling his/her farmland is $p^* - \pi c = \frac{\delta+w-\pi}{2}c$, and that from not selling the farmland is δc. Consequently, the farmer's expected return is:

$$R(c, x) = cE(\max \frac{\delta + w - \pi}{2}, \delta) \tag{2.14}$$

Differentiating equation 2.14 with respect to c, we obtain:

$$\begin{aligned} R_1(c, x) &= E(\max \frac{\delta + w - \pi}{2}, \delta) \\ &= \int_{\underline{w}}^{\overline{w}} [\int_{\underline{\delta}}^{w-\pi(x)} \frac{\delta + w - \pi(x)}{2} f(\delta) d\delta + \int_{w-\pi(x)}^{\overline{\delta}} \delta f(\delta) d\delta] g(w) dw \end{aligned} \tag{2.15}$$

Further differentiating equation 2.16 with respect to x yields:

$$R_{12}(c, x) = -[\int_{\underline{w}}^{\overline{w}} F(w - \pi(x)) g(w) dw] \pi'(x) \tag{2.16}$$

Because land property rights are negatively correlated with land transfer cost, that is $\pi'(x) > 0$, we get $R_{12}(c, x) > 0$ and $I_{12}(c, x) > 0$.

Uncertainty from expropriation, insufficient collateral, and high land transfer costs contribute to insecure land property rights which, in turn, suppress agricultural investment. In the subsequent empirical analysis, we focus on whether this prediction holds for investment in irrigation capital. Capital is required to construct irrigation works, purchase pumps, pipes, and other equipment, as well as to prepare a field to receive water. Both flood and sprinkler irrigation requires capital expenditure, although the investment cost of flood irrigation is significantly lower than sprinkler systems, such as center pivot systems (Dumler,

Rogers, and O'Brien 2007). Importantly, a more-efficient sprinkler system increases yields and allows for more acres to be irrigated (Dumler, Rogers, and O'Brien 2007). Irrigation, and particularly sprinkler irrigation, increases a farmer's ability to grow high-value crops. Therefore, on two otherwise identical parcels, we expect: (1) less investment in irrigation technology on tribal land; (2) conditional on irrigation, we expect less investment in sprinkler irrigation on tribal land; and (3) we expect lower value crops to be grown on tribal land. The next section lays out our empirical methodology for testing these predictions.

2.5 Empirical Framework

2.5.1 Data Construction

Variables on land use, land ownership, land quality and climate are constructed for the Uintah-Ouray Indian Reservation. Table 2.1 shows summary statistics and data construction formulae. Our unit of observation is the parcel from cadastral survey records housed by the Bureau of Land Management (BLM) and supplemented with local records and geographic control coordinates obtained from states, counties, and the United States Geological Survey (USGS) and the United States Forest Service (USFS). Parcels are generally around 40 acres. The survey typically divides land into 6-mile-square townships and townships are subdivided into 36 one-mile-square sections. Sections can be further subdivided into quarter sections, quarter-quarter sections, or irregular government lots.[9] We include the township as a control variable to make sure that we only compare the adjacent parcels. Land ownership type is assigned to each parcel using Geographic Information System (GIS) measurement. The land ownership data comes from the State Geographic Information Database (SGID). This data set contains current surface land ownership administration and designation categories as of 2017. The 2017 tribal land boundary is extracted from this data set. The 1905 allotment boundary is digitized from the Uintah Indian Reservation Disposition map created in 1905. This disposition map contains historical land allotment details at the parcel level for the Uintah reservation. Distance to the boundary is calculated

[9] https://nationalmap.gov/small_scale/a_plss.html

as the shortest distance from the border of each parcel to the 1905 and 2017 boundaries using GIS. We then link the public land survey system (PLSS) quarter, quarter section (parcel) land ownership in 1905 and 2017 to a soil productivity index (PI) grid.

Soil Productivity Data

We obtain the soil PI grid raster map from Iowa State University Geospatial Laboratory for Soil Information. The PI is an ordinal measure of soil productivity, which ranges from 0 (least productive) to 19 (most productive), based on soil taxonomy information (Schaetzl, Krist Jr, and Miller 2012). Since the index is ordinal and some parcels contain two different soil productivity indices, we cannot calculate the mean soil productivity of each parcel as a continuous variable. Following Schaetzl, Krist Jr, and Miller (2012), we assign different soil productivity ranks to each PLSS parcel to ensure each parcel has a unique soil productivity rank. If one parcel has two different soil productivity ranks, we divide this parcel into two parcels with unique rank.

The mean elevation of each parcel in the baseline map is calculated via GIS. The elevation data is obtained from the NASA Shuttle Radar Topographic Mission (SRTM) 90m Digital Elevation Dataset. The SRTM provides digital elevation data (DEMs) for over 80% of the globe and the resolution of the dataset is 3 arc-seconds (approximately 90m resolution).

Agricultural Data

We construct our parcel-level agricultural data using the agricultural land use percentage within each parcel. First, we calculate the agricultural rate using cropland data from CropScape-Cropland Data Layer (CDL)[10] in the year 2015. The CDL is a raster, geo-referenced, crop-specific land cover data layer produced using satellite imagery. Classification accuracy is generally 85% to 95% for the major, crop-specific land cover categories. The CDL database covers the entire Uintah reservation. We obtain 9,304 parcels of 40 acres

[10]CropScape dataset is hosted by National Agricultural Statistics Service, United State Department of Agriculture. Agricultural land layer is in year 2015. Website:https://nassgeodata.gmu.edu/CropScape/

from CDL cropland classification. The average agricultural rate was approximately 30% in the Uintah region. The second database we use is the Water Related Land Use (WRL) data set published annually by the Utah Division of Water Resources.[11] This database provides more accurate agricultural and non-agricultural land cover on portions of the Uintah reservation, but it does not cover the entire study region. The total number of observations is 8,178 parcels, with a 60% agricultural rate. We test the agricultural rate across the boundary using both the CDL and WRL datasets and compare the results.

Irrigation Data

Irrigation rate and sprinkler irrigation rate data come from the WRL data for the year 2012. There are two primary irrigation methods used in the region, sprinkler and flood. Because drip-irrigated acreage is small, its effect on our empirical results is inconsequential and is thus dropped from the analysis of irrigation. Parcel level irrigation and sprinkler irrigation rates are captured by overlaying the irrigation map and sprinkler map on our baseline map. We obtain the sprinkler irrigation rate by dividing the sprinkler irrigated land by total irrigated land. The formulas to calculate the irrigation rate and sprinkler irrigation rate can be found in Table 2.1.

Figure 2.2 shows irrigation by type in the Uintah study region. The left panel shows the correspondence between WRL parcels and the 1905 allotment boundary, and right panel shows the correspondence with the 2017 land ownership. The solid black line indicates the 1905 allotment boundary on the left, and the 2017 tribal land boundary on the right.

High-value Crops Rate Data

We obtain crop data used in this study from the CDL and WRL data. We divide the crops grown in the Uintah reservation into two groups: (i) high-value crops, such as corn and beans, and (ii) low-value crops, such as alfalfa (See Table A.1 for crop value classification).

[11] Final water related land use data describing agricultural related land use in the Uintah region are between 2011 and 2016. The survey year of Uintah region is Year 2012. The last update of this dataset is August 3, 2017.

Table 2.1 shows that more high-value crops are grown on average on private land than tribal land in both data sets.

Figure 2.3 demonstrates the crop value distribution on tribal and private land in the Uintah reservation using WRL data set. The left panel provides the distribution using 1905 allotment boundary, while the right panel is for the 2017 tribal land boundary. In both panels, it appears that more low value crops are inside the tribal boundary.

Climate Data

Temperature and precipitation raster datasets were collected from WorldClim1.4: Current condition (1960-1990). The raster dataset provides the average value of climate statistics between year 1960 and 1990. The resolution of the raster datasets is 30 arc-seconds (1km). We obtain three temperature indicators, including annual mean temperature, maximum temperature of the warmest month, and minimum temperature of the coldest month. In addition, we include precipitation indicators, such as annual precipitation, to control for differences in agricultural productivity across the reservation boundary.

2.5.2 Regression Discontinuity Design

We adopt a spatial regression discontinuity (RD) design to study the cross-border variation in agriculture in the Uintah region. The spatial RD approach has been broadly implemented in different contexts in recent years to study intervention or treatment effects (Bayer, Ferreira, and McMillan 2007; Dell 2010; Grout, Jaeger, and Plantinga 2011; Dachis, Duranton, and Turner 2011; Dell 2015; Card and Giuliano 2016; Pan, Smith, and Sulaiman 2018). Our first empirical strategy exploits the exogenous allocation boundary of 1905 land allotment to explore the impacts of historical tribal land allotment on recent agricultural activities in the context of a sharp RD design.

The sharp RD approach used in this paper hinges on two identifying assumptions. First, the local randomization assumption requires that within a bandwidth of pre-specified size around the 1905 allotment boundary, whether or not an observation receives the treatment is essentially randomly determined. This assumption implies that all the relevant variables

should vary smoothly at the 1905 allotment boundary, and observations located just outside of the 1905 allotment boundary should be an appropriate counterfactual for those located just inside the boundary. To assess the validity of this requirement, we examine the climate statistics, land, and soil variables inside and outside of the 1905 allotment boundary.

Table 2.2 presents the balance test of climate, land, and soil variables for five bandwidth choices (0.5, 0.75, 1, 1.25, 1.5 miles) around the 1905 allotment boundary. In particular, the Welch t-test with log transformation and nonparametric Wilcoxon test are used to test for the difference in means between tribal and private land. The Welch t-test statistics are reported in parentheses, while the Wilcoxon test statistics are in brackets. In the first three columns, the sample includes only parcels located within less than 0.5 miles from the 1905 allotment boundary, and this threshold is gradually increased to 0.75, 1, 1.25, and 1.5 miles in the succeeding columns. It is apparent that the annual mean temperature, annual precipitation, and precipitation of driest month are statistically identical within 1 mile (0.5, 0.75, 1 mile bandwidths) distance across the boundary. As the distance from the boundary increases (1.25 and 1.5 mile bandwidths), however, the values of the balance test variables become statistically different across the boundary. This is consistent with the identification of the treatment effect under RD design. The eighth row shows small statistically significant differences in elevation. The elevation differences are due in part to the location of the Uintah reservation, which is surrounded by a mountain range. The soil productivity is identical within small bandwidths (0.75, 1, 1.25 mile bandwidths).

The second identifying assumption of sharp RD is a continuity assumption, which requires that the only change that occurs at the 1905 allotment boundary is the shift in treatment status. McCrary (2008) proposed an estimator designed to test the continuity of the density function of the forcing variable. He argued that if observations are able to sort themselves across a given bandwidth, the observations just to the left of the cut-off are likely to be substantially different from those to the right. In contrast, De la Cuesta and Imai (2016) argued that the local randomization assumption is stronger than the continuity assumption, and nothing in the continuity assumption requires the expected potential out-

comes on both sides of the threshold to be identical. That means imbalance in pretreatment covariates just below and above the cut-off does not necessarily imply the violation of the identification assumption for a valid RD design.

Under the spatial RD setting the selective sorting assumption would, however, be violated if a direct 1905 allotment effect triggered significant out-migration of relatively highly irrigated land parcels, leading to a larger indirect effect. However, because American Indian Reservations were initially enacted for the express purpose of allowing the tribal members to utilize the land for agricultural production, the continuity assumption is unlikely to hold. For this reason, we recognize the possibility of land switching around the discontinuity and build our models to identify treatment effects under these conditions. Because tribal land boundaries have changed since 1905, we first apply our sharp RD approach only on the lands that do not change ownership to examine the impacts of current agricultural choices across the 1905 allotment boundary. These lands are not affected by the land transactions since 1905 and for this reason retain random assignment. Table 2.3 presents the balance test for the 1905 tribal land boundary with the lands that do not change ownership. The results for lands that never change hands (Table 2.3) are similar to those from Table 2.2. Specifically, the parcels adjacent to the 1905 allotment boundary tend to be similar in reasonable characteristics within smaller distance for the boundary. However, they are different with further distances. Some of the observed differences between Table 2.2 and 2.3 can be explained by the fact that fewer tribal parcels are selected in the dataset that never change hands.

Our second empirical strategy utilizes a fuzzy RD design, which allows us to explore the impact of recent tribal land ownership on agricultural investment today. In the fuzzy RD design, instead of using the lands that do not change ownership, we use all the parcels located within the designated bandwidth of the 2017 boundary. The right panel of Figure 2.1 illustrates the land ownership changes between 1905 and 2017. Green areas represent the land held by the tribe in both 1905 and 2017; red areas represent the land that was allocated to the tribe and became fee-simple between 1905 and 2017; grey areas represent

the land not allotted to the tribe and opened for settlement in 1905; and blue areas represent land that was not originally allocated and was transferred back to the tribe after 1905. It is evident that most of the land returned to the tribe is located on the periphery of tribal land, while most of the land sold to private owners is intermingled with the tribal land. This checkerboard pattern of tribal and private land causes considerable fragmentation of the tribal boundary today.

Table 2.4 presents the balance test across the 2017 tribal land boundary. It is clear that all the climate and land variables are statistically different across the 2017 boundary. This is the result of tribal landowners selling land to non-tribal members (recall that more than 30,000 acres of Uintah agricultural land were sold or leased to non-Indian neighbors (Cuch 2000, p.207)), which considerably altered the original 1905 allotment boundary. Climate and land quality have likely affected whether a parcel has changed ownership since 1905. Consequently, these transactions cause fuzziness in our sample along the 2017 boundary, and we address this by applying a fuzzy RD design, using 1905 allotment boundary as an instrument for current land ownership.

Empirical framework for the 1905 allotment boundary

The 1905 allotment boundary treatment is a straight-line discontinuous function. Thus, we implement a sharp RD design to examine the impact of tribal trust ownership on the agricultural rate, irrigation rate, sprinkler-irrigation rate, and high-value crops rate across the 1905 allotment boundary. For simplicity, we name the treatment in the sharp RD model Allotment1905, which is an indicator, equal to 1 if parcel i is within x miles inside of boundary and equal to 0 if parcel i is within x miles outside of boundary. $dist1905_i$ is the running variable, representing shortest distance of parcel i from the 1905 allotment boundary ($\overline{dist1905}$). $\overline{dist1905}$ is the threshold value (boundary position), equal to 0 in this model. Since the assignment to treatment is sharply determined by the 1905 allotment boundary, the relationship between the treatment indicator and the running variable $dist$ is established by

$$Allotment1905_i = \begin{cases} 1 \quad if \quad dist1905_i \geq \overline{dist1905} \\ 0 \quad if \quad dist1905_i < \overline{dist1905} \end{cases}$$

The parametric linear RD model with a control for distance from the cutoff is:

$$\begin{aligned} (2.17) \quad R1905_i \quad = \quad & \alpha + \beta_1 Allotment1905_i + \beta_2 f(dist1905_i - \overline{dist1905}) \\ & + \beta_3 f(dist1905_i - \overline{dist1905}) \times Allotment1905_i + X^{'} \varphi + \epsilon_i \end{aligned}$$

where $R1905_i$ is the outcome variable of interest of parcel i within x-miles distance from either side of the boundary. X is a vector of controls that includes soil productivity, township and elevation. In our model, we test four different outcome variables: agricultural rate, irrigation rate, sprinkler-irrigation rate, and high-value crop rate. $f(\cdot)$ is a polynomial distance function and ϵ_i is an error term with standard properties. The parameter of interest is β_1, which captures the treatment effect.

As long as a parcel is near the cutoff, $\overline{dist1905}$, the treatment effect of $Allotment1905$ is valid. Hence, an estimate of average treatment effect can be obtained by comparing average $R1905_i$ of those just above and those just below $\overline{dist1905}$. However, the bandwidth has to be large enough to encompass sufficient observations to get a reasonable amount of precision in the estimated average value of $R1905_i$. A larger bandwidth yields more precision but potentially introduces bias.

Empirical framework for the 2017 tribal land ownership

To understand the difference in irrigation rates across the current land ownership boundary, we utilize a fuzzy regression discontinuity. The relationship between land ownership today ($Uintah2017_i$) and the running variable $dist2017_i$ is established by:

$$Uintah2017_i = \begin{cases} 1 \quad if \quad dist2017_i \geq \overline{dist2017} = 0 \\ 0 \quad if \quad dist2017_i < \overline{dist2017} = 0 \end{cases}$$

We cannot compare the average treatment effect immediately above and below the

2017 boundary because the average treatment effect around the cut-off will understate the causal effect. Instead, we can adopt $Allotment1905$ from the sharp RD specification above as an instrumental variable.

There are two basic assumptions about the instrumental variable. First, the relevance condition: $Allotment1905_i$ should have the potential to affect the probability that $Uintah2017_i = 1$. From Figure 2.2, it is clear that the 2017 tribal boundary is strongly related to the 1905 allotment boundary. Second, the exclusion condition: $Allotment1905_i$ has to be unrelated to $R2017_i$, conditional on $Uintah2017_i$ and other controls such as climate and land quality. While not directly testable, we believe this is a plausible assumption for several reasons. First, the 1905 allotment utilized several straight-line boundaries, which were unlikely to have been selected in a way that is correlated with future irrigation scheme. Second, the allotment borders were assigned before the irrigation infrastructure was built on the Uintah reservation.

Because the irrigation project delivered water to both tribal and non-tribal lands, it is also not the case that these boundaries were subsequently used to determine irrigation access. Moreover, the balance tests across the 1905 allotment boundary do not indicate substantial differences in land and climate characteristics that might have been observable at the time of assignment (see Table 2.2).

The fuzzy RD design is a two-stage estimation process. The first stage involves regressing the 2017 treatment indicator on the 1905 boundary and additional controls (soil productivity, township and elevation):

$$Uintah2017_i = \lambda + \gamma Allotment1905_i + g(dist1905_i - \overline{dist1905}) + X^{'}\varphi + \nu_i \tag{2.18}$$

To estimate the first stage we fit a generalized linear model with a probit function. Once we obtain the fitted value of $Uintah2017_i$ from stage 1, we use $\widehat{Uintah2017}_l$ to evaluate the average treatment effect in stage 2:

$$R2017_i = \delta + \beta_1 \widehat{Uintah2017}_l + h(dist1905_i - \overline{dist1905}) + X^{'}\varphi + \epsilon_i \tag{2.19}$$

The treatment effect is captured by β_1.

Bandwidth and functional form selection

Identification of the local spatial RD treatment effect requires data points in the immediate neighborhood around the border, whether it is a sharp or fuzzy design. As the neighborhood expands, the estimate of the average treatment effect becomes less noisy, while the risk of bias of the estimate increases as the trends and variations in other variables across the discontinuity may affect the estimates. While some of these effects can be controlled using additional regressors and polynomial order trends in distance, the selection of the bandwidth around the discontinuity remains an important consideration. In the present study, we use a bandwidth selection procedure based on Calonico, Cattaneo, and Titiunik (2014, 2015), who suggest using a simple kernel and then verifying the robustness of the results to different choices of bandwidth. Accordingly, we analyze the data with 0.5-mile, 0.78-mile,1-mile, 1.25-mile and 1.5-mile bandwidths around the 1905 allotment boundary, using both sharp and fuzzy RD designs, in addition to the optimal bandwidth.

Furthermore, we implement the non-parametric, bias-corrected robust inference procedure proposed by Calonico, Cattaneo, and Titiunik (2014) to select the functional form for the running variable ($f(\cdot), g(\cdot)$ *and* $h(\cdot)$) and to study the discontinuities at the boundary more closely. This approach can be used in contexts with a large number of observations very close to the treatment threshold (Imbens and Lemieux 2008). The nonparametric technique has the advantage of not relying on functional form assumptions and is commonly used in spatial RD design (Dell 2010). To build the nonparametric function of the running variables, we fit the 1st, 2nd, 3rd, 4th order local polynomial regression. It is common in regression discontinuity analysis to control for 3rd, 4th, or higher-degree polynomials of the forcing variable. However, Gelman and Imbens (2018) argue that high-order polynomials are ill-suited to regression discontinuity analysis because they lead to noisy estimates, sensitivity to the degree of the polynomial, and poor coverage of confidence intervals. Instead, they recommended using estimators based on local linear or quadratic polynomials. We present results using 2nd order polynomial and include the 1st, 3rd, and 4th order

polynomial results in the Appendix as robustness checks.

2.6 Results

2.6.1 Sharp RD Regression Results

We begin by testing the 1905 allotment boundary impact on agriculture and crop choice variables using the sharp-RD design on the lands where ownership does not change. This is similar to an intent-to-treat specification. Figure 2.4 plots soil quality and agricultural activities by distance to boundary by land ownership changes. Land moving from tribal to private ownership has higher irrigation rates, sprinkler-irrigation rates and high-value crop rates. This implies that better land was transferred from tribal to private ownership between 1905 and 2017.

Table 2.5 shows the empirical result of the sharp RD design using different bandwidth and a second-order polynomial of the running variable. First, we estimate the effect on soil productivity, using the soil productivity index as the dependent variable. Column 1 of Table 2.4 limits the sample to parcels within 1.5 miles of the 1905 allotment boundary, and columns 2 – 5 restrict it to fall within 1.25, 1, 0.75, and 0.5 miles, respectively. Column 6 reports the allotment effect with the optimal bandwidth obtained from nonparametric specification and column 7 indicates the optimal bandwidth. Rows 2-7 present the results for agricultural rate, irrigation rate, sprinkler-irrigation rate, and high-value crops rate as the dependent variable. Controlling for township, soil productivity index and elevation ensures that we are comparing parcels in close geographic proximity with similar soil quality and elevation. Appendix Table A.2 to Table A.8 examine robustness of the main specification to multiple control variables of 1st, 3rd, 4th order polynomials.

The results for soil productivity indicate that the treatment coefficients are positive but not statistically significant at the 10% level. While there are no apparent differences in rates of agriculture and irrigation on allotted lands, allotted lands see around a 12-percentage point lower rate of sprinkler irrigation (-0.146 to -0.102). These negative effects decrease as the bandwidth is reduced, but remain statistically significant at the 1% level. Hence,

the results consistently indicate that there is a negative effect of being inside the 1905 allotment border on investment in sprinkler irrigation. Moreover, the allotment coefficients are economically similar across the four specifications of the RD polynomial,[12] and we are unable to reject that they are statistically identical (shown in Appendix Table A.6). The coefficients for high-value crops also show a decrease across the 1905 allotment boundary using both the CDL and WRL data. The negative allotment coefficients range from -0.046 to -0.037 in CDL dataset and -0.014 to -0.009 in WRL dataset.[13] This implies that even excluding a large portion of land with high-value crops that has been transferred out of tribal control, tribal lands have lower levels of high-value crops.[14]

2.6.2 Fuzzy RD Regression Results

Table 2.6 reports the estimates of average treatment effect from the two-stage, fuzzy RD approach under different bandwidth choices. In all but the last set, we consider a fixed bandwidth from the 2017 tribal boundary to each parcel boundary, while the last set evaluates the average treatment effects with the optimal bandwidth choice. Each cell in this table reports an estimate of the average treatment effect for different bandwidths for a second-order polynomial of the running variable. Appendix Table A.9 to Table A.15 examine the robustness of the main specification, which demonstrate that the 2017 tribal boundary effects on each dependent variable are generally similar across different polynomial orders. As the bandwidth is increased, there is a modest decrease in the size of the treatment effect for 2nd, 3rd and 4th polynomial order.

Table 2.6 provides a rigorous analysis of average treatment effect by considering the effect of key covariates that might affect irrigation rate inside and outside the tribal land

[12]Table A.6 shows the sharp RD results of sprinkler irrigation rate using 1st, 2nd, 3rd, 4th order polynomial. The negative allotment effect (ranging from -0.146 to -0.093) is statistically significant at the 1% level.

[13]This negative effect still exists when we choose the different order polynomials, see Table A.7 and Table A.8 .

[14] The sharp RD results are illustrated in Fig A.1 to A.5. Each subfigure shows one choice of polynomial order plot. Based on the inspection of these plots, it is evident that the 2nd order polynomial regression models fit the data better than 1st, 3rd and 4th order polynomial models. Higher order polynomial regression models are more easily affected by outliers but generally provide a better fit for the data. The subfigures uniformly confirm the presence of a significant discontinuity at 1905 allotment boundary, thus corroborating the main findings from Table 2.5.

boundary. While soil productivity is consistently higher inside the boundary, the rate of irrigation, sprinkler irrigation, and high-value crops are lower. The tribal boundary effect lowers the irrigation rate by around eighteen percentage points (-0.195 to -0.177) within the reservation. The treatment coefficients are economically similar when we apply multiple controls, and we are unable to reject that they are statistically identical. The treatment effects remain similar as the bandwidth decreases and polynomial order increases (Appendix Table A.12). Similarly, tribal land sees an approximately 31 percentage point lower (-0.323 to -0.305) sprinkler-irrigated rate compared to non-tribal land. After controlling for covariates, the treatment effect coefficients are still statistically significant at the 1% level. The average treatment estimates are consistent across different bandwidth choices and multiple covariates.

In row 1 of Table 2.6, we describe estimates of treatment effect, using soil productivity index as the dependent variable. The tribal boundary effect increases the soil productivity index by one full rank in subjected parcels (1.011 to 1.143). After controlling for township and elevation, the treatment effect coefficients are still statistically significant at the 1% level. This indicates that land quality is statistically higher on tribal land. This result moves in the opposite direction of the selection effect we might suspect, where better land outside the reservation has more investment and higher-value crops. Instead, we see more irrigation and high-value crops on the (relatively) poorer land outside the reservation. Rows 6 and 7 report the high-value crop difference across the tribal boundary today. Each cell reports the coefficient on $\widehat{Uintah2017}$ for different bandwidth choices. As an example, the average tribal boundary effect of high-value crop rate is 14.7 percentage points lower on tribal land than on private land in the CDL data set, and 4.3 percentage points lower on tribal land in the WRL data set.[15]

[15]The negative effect ranges from -0.153 to -0.143 in CDL dataset and -0.045 to -0.042 in WRL dataset. The non-parametric fuzzy RD results are illustrated in Fig A.6 to A.10. Each subfigure shows one choice of polynomial order plot. Based on the inspection of these plots, it is evident that the 2nd order polynomial regression models fit the data better than 1st, 3rd and 4th order polynomial models.

2.7 Conclusion

This paper explores the effect of tribal land ownership on agricultural development caused, at least in part, by insecure property rights on American Indian Reservations. Our economic framework suggests fee-simple landowners with secure property rights are more likely to obtain access to commercial credit and borrow money to invest capital intensive improvements. The effect is that Uintah reservation lands see less intensive cultivation and lower value crops. Our findings illustrate that when controlling for land quality and geographic location, fee-simple land has an irrigation rate approximately 18 percentage points higher than tribal land. Conditional on being irrigated, tribal land is 31 percentage points less likely to be sprinkler-irrigated today. Moreover, fee-simple farms have higher amounts of high-value crops within 1.5 miles of the boundary with reservation land.

Evaluating allotment effects in 1905 and tribal boundary effects together, we conclude that agricultural irrigation development on the Uintah reservation is suppressed relative to non-reservation land. This lack of investment is consistent with our expectation of the effect of insecure property rights on tribal trust land. However, there are several alternative explanations for the observed results, including issues with BIA irrigation projects, a lack of access to USDA loan programs, and other government support and subsidies to which fee-simple farmers may receive preferential access. The results do suggest that trust ownership creates significant barriers to the acquisition of capital for agricultural investment. While access to investment capital may have multiple causes, it appears clear that improving access to capital, so tribal farmers can invest in irrigation systems at the level of their fee-simple neighbors, is key to improving lagging agricultural development on reservations.

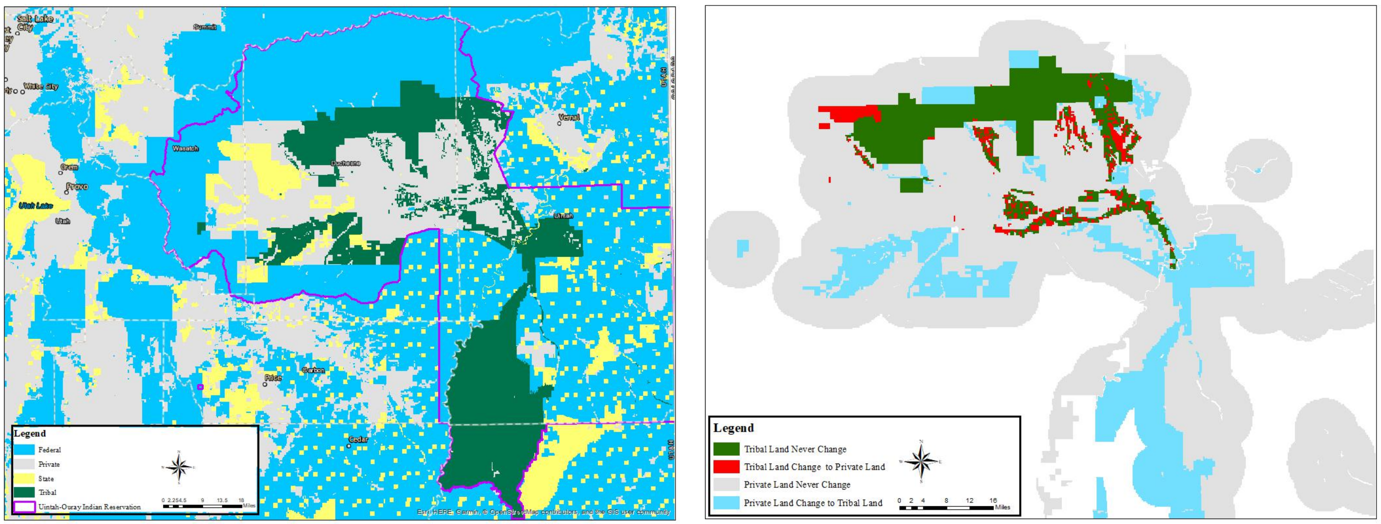


Figure 2.1. Land ownership map of Unitah Ouray Indian Reservation
Note: Left: Land ownership map of Uintah and Ouray Indian Reservation in 2017. Federal (Blue) represents land owned by federal government, private (Grey) represents privately owned land (the owner can be tribal member or non-tribal member), tribe (Green) represents tribal land, and state (Yellow) represents state owned land. Right: Land ownership changes map of Uintah and Ouray Indian Reservation in 2017. Green represents tribe owned land in both 1905 and 2017, red represents the land changes from tribe to private owner, grey represents private owned land in both 1905 and 2017, and blue is the land changes from private owner to tribe.

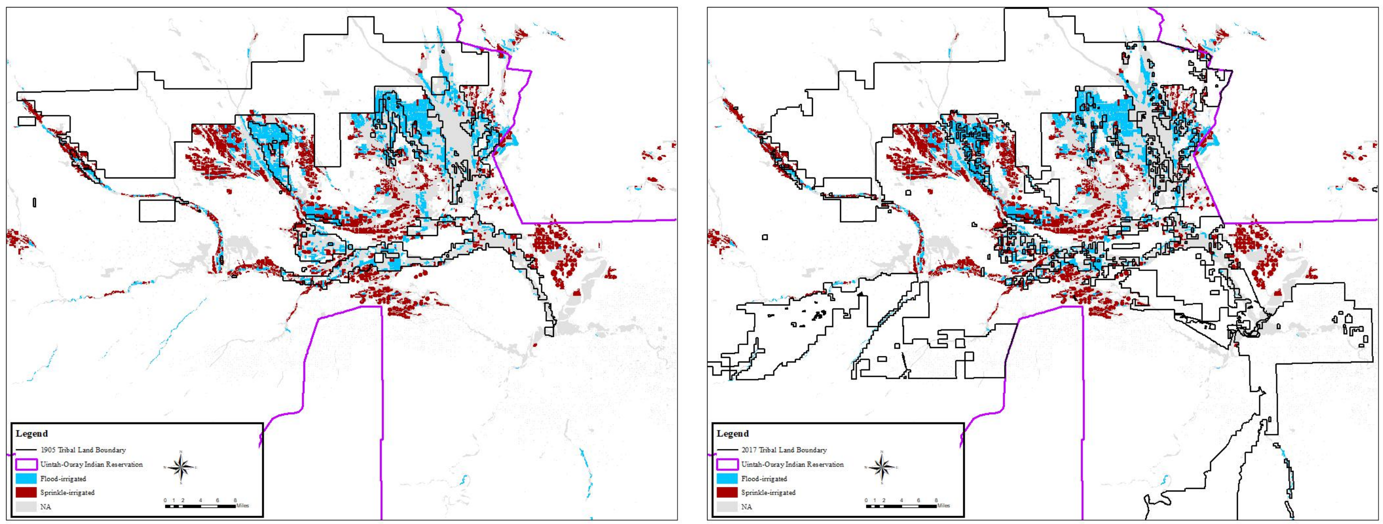


Figure 2.2. Water-related land use in Uintah and Ouray Indian Reservation in 1905 and 2017
Note: Left: Water-related-land-use parcels in Uintah and Ouray Indian Reservation in 1905. Right: Water-related-land-use parcels in Uintah and Ouray Indian Reservation in 2017. Blue represents flood-irrigated land, brown represents sprinkler-irrigated land, light grey represents non-irrigated land.

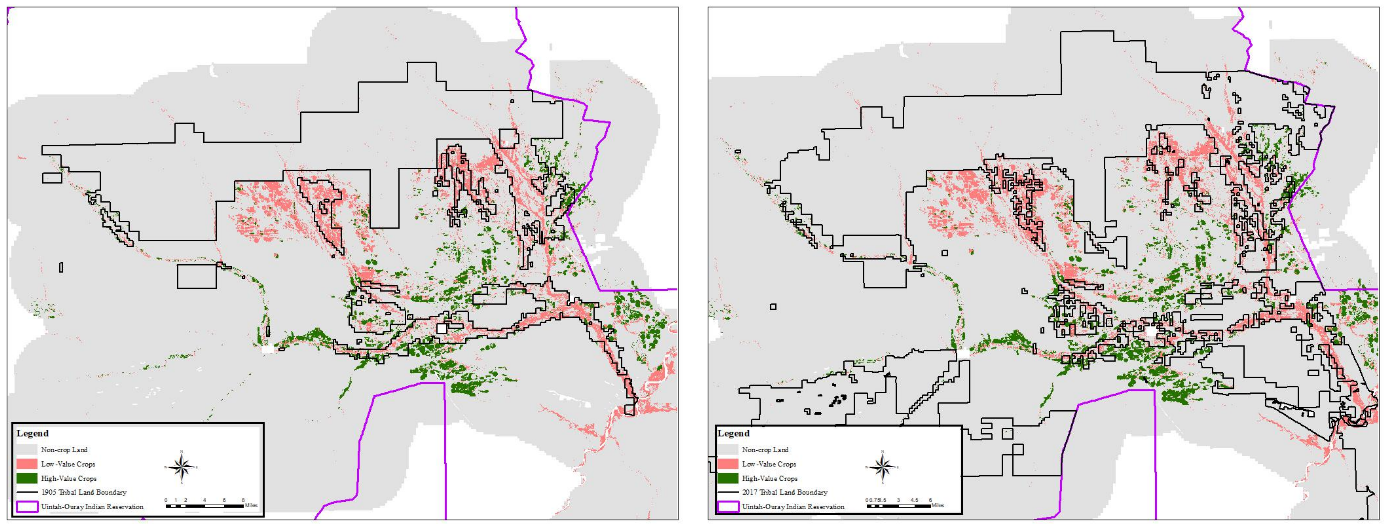


Figure 2.3. Crop value distribution in 1905 and 2017
Note: Left: Crop value distribution map for 1905 allotment boundary. Right: Crops value distribution map for 2017 tribal boundary. Pink represents low value crops, green represents high value crops, and light grey represents non-crop land.

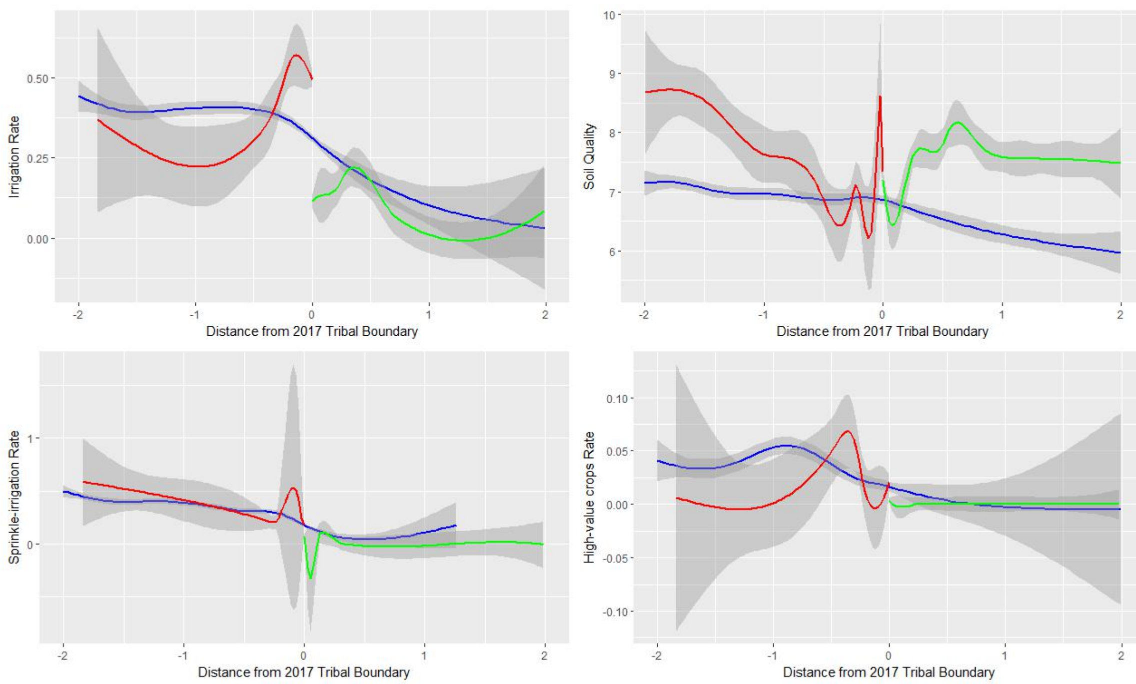


Figure 2.4. Outcome variable changes between 1905 and 2017
Note: Top left: Irrigation rate difference due to land ownership changes. Top right: Soil quality difference due to land ownership changes. Bottom left: Sprinkler-irrigation rate difference due to land ownership changes. Bottom right: High-value crops rate difference due to land ownership changes. Positive range of x-axis represents tribal land in 2017, and negative range of x-axis represents private land in 2017. Red line represents the trend of lands changing land ownership from tribal to private between 1905 and 2017. Green line represents the trend of lands changing landownership from private to tribal between 1905 and 2017. Blue line represents the trend of lands with no change in land ownership.

Table 2.1. Summary statistics

	Formula	< 2-mile 1905 Allotment Boundary				< 2-mile 2017 Tribal Boundary			
		Observation		Mean		Observation		Mean	
		Tribal	Private	Tribal	Private	Tribal	Private	Tribal	Private
Agricultural Rate (Cropscape)	$AgRate = \frac{AgriculturalLand}{TotalLand}$	3,343	5,961	0.342	0.300	3,411	8,560	0.301	0.301
				(0.305)	(0.303)			(0.309)	(0.305)
Agricultural Rate (Utah GIS)		2,753	5,425	0.659	0.588	2,190	7,202	0.574	0.590
				(0.361)	(0.352)			(0.397)	(0.351)
Irrigation Rate	$IrrRate = \frac{IrrigatedLand}{TotalLand}$	3,168	6,101	0.347	0.384	2,659	8,032	0.228	0.397
				(0.402)	(0.380)			(0.357)	(0.381)
Sprinkle-irrigated Rate	$SprkRate = \frac{SprinkleLand}{IrrigatedLand}$	2,060	4,498	0.129	0.326	1,431	6,065	0.073	0.312
				(0.286)	(0.361)			(0.219)	(0.361)
High Value Crops Rate (Cropscape)	$HcropRate = \frac{HcropLand}{TotalLand}$	3,343	5,961	0.060	0.116	3,411	8,560	0.024	0.119
				(0.177)	(0.233)			(0.109)	(0.237)
High Value Crops Rate (Utah GIS)		2,753	5,425	0.015	0.027	2,190	7,202	0.008	0.032
				(0.099)	(0.130)			(0.071)	(0.144)

Note: The sample contains 54,377 observations located less than 2 miles from either 1905 Allotment Boundary or 2017 Tribal Boundary. In Column 3 to 6, the sample includes only observations located less than 2 miles from the 1905 Allotment Boundary. Column 7 to 10 represent the sample located less than 2 miles from the 2017 Tribal Boundary. In the Uintah and Ouray Indian Reservation, only two types of irrigation system are used: flood irrigation system and sprinkler irrigation ($IrrigatedLand = FloodLand + SprinkleLand$). Standard errors are provided in the parenthesis.

Table 2.2. 1905 allotment border balance test

	< 0.5 Miles			< 0.75 Miles			< 1 Miles			< 1.25 Miles			< 1.5 Miles		
	inside	outside	statistics	inside	outside	statistics	inside	outside	statistics	inside	outside	statistics	inside	outside	statistics
							Climate Statistics								
Observations	1,001	1,229		1,199	1,602		1,368	1,954		1,503	2,290		1,624	2,578	
Annual Temp(^oF)	43.8	43.7	(1.26)	43.6	43.7	(-0.544)	43.4	43.6	(-1.3)	43.2	43.5	(-3.11) ***	43.1	43.5	(-4.17) ***
			[6.2E+5]			*[9.2E+5]* *			*[1.2E+6]* ***			*[1.5E+6]* ***			*[1.8E+6]* ***
Max Temp(^oF)	85.4	85.3	(0.74)	85.0	85.2	(-1.23)	84.6	85.0	(-2.41) **	84.2	85.0	(-4.33) ***	83.9	84.9	(-5.66) ***
			[6.2E+5]			*[9.2E+5]* **			*[1.2E+6]* ***			*[1.5E+6]* ***			*[1.8E+6]* ***
Min Temp(^oF)	5.5	5.4	(0.90)	5.6	5.4	(2.89) ***	5.8	5.4	(5.50) ***	5.8	5.4	(7.61) ***	5.9	5.4	(9.78) ***
			[6.4E+5] *			*[1.1E+6]* ***			*[1.5E+6]* ***			*[2.1E+6]* ***			*[2.6E+6]* ***
Annual Precipitation(mm)	249	257	(-1.89) *	256	258	(-0.29)	261	262	(0.58)	266	263	(2.33) **	269	265	(3.41) ***
			[5.9E+5]			*[9.6E+5]*			*[1.4E+6]* *			*[1.8E+6]* ***			*[2.3E+6]* ***
Min Precipitation(mm)	15.0	15.7	(-1.94) *	15.5	15.8	(-0.59)	16.0	16.1	(0.53)	16.4	16.2	(2.10) **	16.8	16.3	(3.30) ***
			[5.9E+5] **			*[9.4E+5]*			*[1.4E+6]*			*[1.8E+6]* **			*[2.2E+6]* ***
							Land Statistics								
Observations	4,476	5,511		5,330	7,290		5,992	8,833		6,596	10,369		7,088	11,810	
Elevation(m)	1,872	1,855	(2.51) **	1,901	1,865	(5.89) ***	1,927	1,876	(8.75) ***	1,949	1,886	(11.50) ***	1,966	1,894	(13.80) ***
			[1.3E+7] **			*[2.1E+7]* ***			*[2.9E+7]* ***			*[3.9E+7]* ***			*[4.8E+7]* ***
Soil Productivity	7.1	7.1	(1.65) *	7.0	7.0	(1.01)	7.0	7.0	(-0.352)	6.9	6.9	(-1.53)	6.8	6.9	(-2.61) ***
			[1.3E+7] **			*[2.0E+7]* *			*[2.7E+7]*			*[3.4E+7]*			*[4.2E+7]*

Note: Annual Temp represents annual mean temperature. Max Temp represents Max Temperature of Warmest Month and Min Temp represents Min Temperature of Coldest Month. Min Precipitation represents precipitation of Driest Month. The unit of observation is PLSS QuarterQuarter Section. Summary statistics table show two datasets: climate statistics and land statistics dataset. The land statistics dataset contains all the private and tribal parcels in the Uintah and Ouray Indian reservation, while the climate statistics dataset contains only most representative PLSS parcels. Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%.

Table 2.3. 1905 allotment border balance best (never switched land)

	< 0.5 Miles			< 0.75 Miles			< 1 Miles			< 1.25 Miles			< 1.5 Miles		
	inside	outside	statistics	inside	outside	statistics	inside	outside	statistics	inside	outside	statistics	inside	outside	statistics
							Climate Statistics								
Observations	748	1,030		892	1,345		1,053	1,637		1,164	1,939		1,282	2,195	
Annual Temp (^oF)	43.6	43.7	(-0.715)	43.4	43.7	(-1.99) *	43.2	43.5	(-3.19) ***	43.1	43.5	(-4.05) ***	42.9	43.4	(-5.2) ***
			[3.6E+5] *			*[5.5E+5]* ***			*[7.5E+5]* ***			*[9.5E+5]* ***			*[1.2E+6]* ***
Max Temp(^oF)	84.9	85.2	(-1.16)	84.5	85.1	(-2.71) ***	84.1	84.9	(-4.31) ***	83.8	84.9	(-5.4) ***	83.6	84.8	(-6.83) ***
			[3.6E+5] **			*[5.4E+5]* **			*[7.4E+5]* ***			*[9.4E+5]* ***			*[1.1E+6]* ***
Min Temp(^oF)	5.68	5.41	(2.27) **	5.81	5.41	(4.32) ***	5.97	5.43	(7.16) ***	6.06	5.44	(9.04) ***	6.15	5.45	(11.40) ***
			[4.4E+5] ***			*[7.1E+5]* ***			*[1.1E+6]* ***			*[1.5E+6]* ***			*[1.9E+6]* ***
Annual Precipitation(mm)	254	258	(-0.0469)	259	260	(0.97)	265	263	(2.21) **	268	265	(2.97) ***	271	266	(4.16) ***
			[3.9E+5]			*[6.3E+5]* *			*[9.4E+5]* *			*[1.2E+6]* ***			*[1.6E+6]* ***
Min Precipitation(mm)	15.6	15.8	(0.12)	16	15.9	(1.01)	16.5	16.2	(2.49) **	16.7	16.3	(3.18) ***	17.1	16.4	(4.45) ***
			[3.8E+5]			*[6.1E+5]*			*[9.1E+5]* **			*[1.2E+6]* ***			*[1.5E+6]* ***
							Land Statistics								
Observations	3,202	4,574		3,904	6,067		4,487	7,382		5,030	8,697		5,479	9,936	
Elevation(m)	1,898	1,853	(5.96) ***	1,928	1,867	(9.07) ***	1,953	1,880	(11.50) ***	1,973	1,892	(13.90) ***	1,990	1,900	(16.10) ***
			[7.9E+6] ***			*[1.3E+7]* ***			*[1.9E+7]* ***			*[2.6E+7]* ***			*[3.3E+7]* ***
Soil Productivity	7.0	7.2	(-3.81) ***	6.9	7.1	(-5.2) ***	6.8	7.1	(-6.93) ***	6.7	7.1	(-8.53) ***	6.7	7.1	(-10) ***
			[7.0E+6] ***			*[1.1E+7]* ***			*[1.6E+7]* ***			*[2.1E+7]* ***			*[2.6E+7]* ***

Note: Annual Temp represents annual mean temperature. Max Temp represents Max Temperature of Warmest Month and Min Temp represents Min Temperature of Coldest Month. Min Precipitation represents precipitation of Driest Month. The unit of observation is PLSS QuarterQuarter Section. Summary statistics table show two datasets: climate statistics and land statistics dataset. The land statistics dataset contains all the private and tribal parcels in the Uintah and Ouray Indian reservation, while the climate statistics dataset contains only most representative PLSS parcels. Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%.

Table 2.4. 2017 tribal boundary balance test

	< 0.5 Miles				< 0.75 Miles				< 1 Miles				< 1.25 Miles				< 1.5 Miles			
	inside	outside	statistics		inside	outside	statistics		inside	outside	statistics		inside	outside	statistics		inside	outside	statistics	
Climate Statistics																				
Observations	2,332	2,830			2,824	3,653			3,215	4,395			3,536	5,232			3,810	5,919		
Annual Temp(^oF)	44.0	44.4	(-4.73)	***	43.9	44.3	(-5.82)	***	43.8	44.3	(-6.93)	***	43.7	44.2	(-6.27)	***	43.6	44.1	(-6.95)	***
			[3.0E+6]	***			*[4.7E+6]*	***			*[6.4E+6]*	***			*[8.4E+6]*	***			*[1.0E+7]*	***
Max Temp(^oF)	85.5	86.1	(-4.61)	***	85.3	86.1	(-5.73)	***	85.0	85.9	(-7.27)	***	84.9	85.8	(-6.91)	***	84.7	85.6	(-7.91)	***
			[3.1E+6]	***			*[4.7E+6]*	***			*[6.3E+6]*	***			*[8.4E+6]*	***			*[1.0E+7]*	***
Min Temp(^oF)	5.9	5.8	(0.73)		6.0	5.9	(1.18)		6.1	5.9	(2.93)	***	6.2	5.9	(3.66)	***	6.3	6.0	(4.88)	***
			[3.4E+6]	***			*[5.4E+6]*	***			*[7.6E+6]*	***			*[1.0E+7]*	***			*[1.2E+7]*	***
Annual Precipitation(mm)	270	263	(3.30)	***	274	265	(4.13)	***	278	268	(5.15)	***	280	272	(4.52)	***	283	275	(5.02)	***
			[3.5E+6]	***			*[5.4E+6]*	***			*[7.5E+6]*	***			*[9.8E+6]*	***			*[1.2E+7]*	***
Min Precipitation(mm)	16.3	15.7	(3.42)	***	16.6	15.9	(4.28)	***	17.0	16.1	(5.37)	***	17.1	16.4	(4.98)	***	17.3	16.6	(5.52)	***
			[3.5E+6]	***			*[5.4E+6]*	***			*[7.5E+6]*	***			*[9.8E+6]*	***			*[1.2E+7]*	***
Land Statistics																				
Observations	10,480	12,750			12,627	16,421			14,233	19,811			15,630	23,232			16,804	26,516		
Elevation(m)	1,887	1,852	(7.39)	***	1,910	1,864	(10.60)	***	1,927	1,876	(12.50)	***	1,940	1,887	(13.90)	***	1,951	1,896	(15.00)	***
			[6.9E+7]	***			*[1.1E+8]*	***			*[1.5E+8]*	***			*[2.0E+8]*	***			*[2.4E+8]*	***
Soil Productivity	7.3	6.9	(7.35)	***	7.3	6.9	(8.92)	***	7.3	6.9	(9.03)	***	7.3	6.9	(8.60)	***	7.3	6.9	(8.09)	***
			[7.3E+7]	***			*[1.1E+8]*	***			*[1.5E+8]*	***			*[2.0E+8]*	***			*[2.4E+8]*	***

Note: Annual Temp represents annual mean temperature. Max Temp represents Max Temperature of Warmest Month and Min Temp represents Min Temperature of Coldest Month. Min Precipitation represents precipitation of Driest Month. The unit of observation is PLSS QuarterQuarter Section. Summary statistics table show two datasets: climate statistics and land statistics dataset. The land statistics dataset contains all the private and tribal parcels in the Uintah and Ouray Indian reservation, while the climate statistics dataset contains only most representative PLSS parcels. Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%.

Table 2.5. 1905 allotment border non-parametric RD results (second order polynomial)

Sample Within	Estimated Average Treatment Effects							Control Variables		
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth		Elevation	Townships	Soil Productivity
Soil Productivity Index							Optimal Miles			
Allotment1905	0.139	0.117	0.103	0.084	0.067	0.139	1.289	Ã	Ã	
	(0.115)	(0.118)	(0.120)	(0.122)	(0.127)	(0.115)				
Agricultural Rate (Cropscape)										
Allotment1905	0.004	0.004	0.007	0.008	0.015	0.003	1.328	Ã	Ã	Ã
	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)	(0.018)				
Agricultural Rate (Utah GIS)										
Allotment1905	0.040	0.044	0.041 *	0.040	0.037	0.037	1.699	Ã	Ã	Ã
	(0.024)	(0.025)	(0.026)	(0.026)	(0.027)	(0.023)				
Irrigation Rate										
Allotment1905	0.005	0.015	0.017	0.018	0.016	-0.002	1.595	Ã	Ã	Ã
	(0.024)	(0.025)	(0.025)	(0.026)	(0.027)	(0.023)				
Sprinkle-irrigated Rate										
Allotment1905	-0.146 ***	-0.133 ***	-0.124 ***	-0.119 ***	-0.102 ***	-0.120 ***	0.954	Ã	Ã	Ã
	(0.026)	(0.027)	(0.028)	(0.028)	(0.029)	(0.028)				
High-value Crops Rate (Cropscape)										
Allotment1905	-0.046 ***	-0.046 ***	-0.042 ***	-0.038 ***	-0.037 **	-0.045 ***	0.972	Ã	Ã	Ã
	(0.013)	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)				
High-value Crops Rate (Utah GIS)										
Allotment1905	-0.012	-0.013 *	-0.014 *	-0.014 *	-0.013	-0.009	2.162	Ã	Ã	Ã
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.006)				

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis. The second order polynomial sharp RD results of five different dependent variables are shown in the table. Row 2 and 3 shows the agricultural rate RD effects of two different dataset, CDL and WRL, while the row 6 and 7 shows the RD results of high-value crop rate in these two datasets. Five different bandwidths choices results are listed in column 1 to 5. Elevation, townships and soil productivity are controlled.

Table 2.6. 2017 allotment border non-parametric RD results (second order polynomial)

Sample Within	Estimated Average Treatment Effects							Control Variables		
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	Optimal Miles	Elevation	Townships	Soil Productivity
Soil Productivity Index										
Tribal2017	1.011 ***	1.060 ***	1.111 ***	1.123 ***	1.143 ***	1.025 ***	1.521	Ã	Ã	
	(0.385)	(0.400)	(0.414)	(0.422)	(0.435)	(0.359)				
Agricultural Rate (Cropscape)										
Tribal2017	-0.048	-0.043	-0.030	-0.019	-0.003	-0.048	1.323	Ã	Ã	Ã
	(0.033)	(0.034)	(0.036)	(0.036)	(0.037)	(0.034)				
Agricultural Rate (Utah GIS)										
Tribal2017	-0.020	-0.026	-0.027	-0.025	-0.024	-0.024	1.355	Ã	Ã	Ã
	(0.039)	(0.041)	(0.042)	(0.043)	(0.044)	(0.040)				
Irrigation Rate										
Tribal2017	-0.195 ***	-0.193 ***	-0.187 ***	-0.182 ***	-0.177 ***	-0.192 ***	1.396	Ã	Ã	Ã
	(0.042)	(0.043)	(0.044)	(0.045)	(0.046)	(0.043)				
Sprinkle-irrigated Rate										
Tribal2017	-0.323 ***	-0.319 ***	-0.313 ***	-0.309 ***	-0.305 ***	-0.320 ***	1.319	Ã	Ã	Ã
	(0.044)	(0.046)	(0.047)	(0.047)	(0.048)	(0.045)				
High-value Crops Rate (Cropscape)										
Tribal2017	-0.147 ***	-0.147 ***	-0.147 ***	-0.143 ***	-0.144 ***	-0.153 ***	1.316	Ã	Ã	Ã
	(0.024)	(0.025)	(0.026)	(0.026)	(0.027)	(0.026)				
High-value Crops Rate (Utah GIS)										
Tribal2017	-0.043 ***	-0.044 ***	-0.045 ***	-0.044 ***	-0.042 ***	-0.044 ***	1.168	Ã	Ã	Ã
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.013)				

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis. The second order polynomial fuzzy RD results of five different dependent variables are shown in the table. Row 2 and 3 shows the agricultural rate RD effects of two different dataset, CDL and WRL, while the row 6 and 7 shows the RD results of high-value crop rate in these two datasets. Five different bandwidths choices results are listed in column 1 to 5. Elevation, townships and soil productivity are controlled.

CHAPTER 3

REGIONAL WATER TRADE IN GENERAL EQUILIBRIUM: THEORY AND EVIDENCE FROM THE UNITED STATES' LARGEST EVER AG-TO-URBAN WATER TRANSFER

3.1 Abstract

In arid regions, water sale of water from rural to urban areas may offer a solution to reduce costly water shortfalls among high-value urban users by transferring water originally used for relatively low-value agriculture. However, such sales have been criticized for reducing economic activity and the availability of water for ecosystem services in the rural region. In this study, we examine the impact of an agreement in California that transferred significant amounts of water from Imperial County to San Diego County between 2004 and 2016. We develop a general-equilibrium representation of a hydrologic-ecological-economic system to explore the theoretical effect of this trade between regional economies. The model predicts increases in the value of water in Imperial County and losses in employment income. We test the effect empirically using a synthetic counterfactual analysis. Consistent with the model, our estimates show a decline in the number of low- and high-skill jobs in Imperial County corresponding to the water transfer, as well as a reduction in overall crop production. In addition, increased crop yields indicate higher water values in the post-trade period. The increased volume of water trading appears to have reduced water availability in Imperial County, as predicted by the model, leading to corresponding declines in the Salton Sea.

3.2 Introduction

Water stress in the world's arid regions is increasing due to increasing urban water demand and is likely to be aggravated by decreasing water availability due to climate change

and land use alteration (Schewe et al. 2014; Vörösmarty et al. 2000). The reallocation of water between different regions and sectors potentially addresses the need to meet human water demand. However, such sales have been criticized for reducing economic activity and the availability of water for ecosystem services in the rural region. Few studies have explored the effects of cross-sectoral water allocation but effects are varied across the regions(e.g., see papers on Australia (Brooks and Harris 2008), China (Cai 2008), South Africa (Juana, Strzepek, and Kirsten 2011), and Europe (Wimmer et al. 2015)). Wimmer et al. (2015) and Juana, Strzepek, and Kirsten (2011) found that cross-sectoral water allocation between agricultural and non-agricultural sectors in Europe and South Africa would lead to physical water shortage in most of water exporting region and consequently cause socioeconomic disadvantage and aquatic ecosystem degradation. However, Cai (2008) and Brooks and Harris (2008) studied the water allocation in Northern China and Australia and oppositely concluded that the water allocation between sectors will generate an efficient water market and therefore increase the sectoral output in the future.

In this study, we examine the impact of Quantification Settlement Agreement (QSA) in California that transferred significant amounts of water from Imperial County to San Diego County between 2004 and 2016. We develop a general-equilibrium representation of a hydrologic-ecological-economic system to explore the theoretical effect of this trade between regional economies. This is the first study we are aware of that links general equilibrium analysis with empirical results in examining the efficiency of a regional water transfer agreement between agriculture and urban areas. A water transfer between agriculture and urban areas benefits the buyer and seller who are directly involved, but the impact on other parties is hotly debated. Howe, Lazo, and Weber (1990) summarized potential negative externalities including reductions in water quality, water availability and instream flows. In addition, so-called pecuniary externalities have been cited as negative trade outcomes, including the loss of jobs in the region from which water is transferred (Mann and Wüstemann 2008; Holcombe and Sobel 2001). In contrast, a few studies find positive impacts of the water transfer policy. For example, Grafton et al. (2011) argued

that water allocation will create direct benefits for water buyers and sellers, and the benefits increase as water availability decreases. Hadjigeorgalis and Lillywhite (2004) calculated the gains from water allocation in Chile and found that the positive gains from water allocation account for approximately 8 to 32 percent of the total regional GDP.

In 2003, a water transfer agreement called QSA was signed between the Imperial Irrigation District (IID), Metropolitan Water District of Southern California, Coachella Valley Water District, the State of California, and the United States Department of the Interior. The intent of the QSA is to require water conservation in the IID so that water can be transferred to other regional uses, including the San Diego County Water Authority (SDCWA). This agreement resulted in a water transfer between the Imperial Valley and San Diego for a duration of 35 years, the largest agricultural to urban water transfer in the United States.[1] The QSA provides a case study to test the performance of a regional water transfer agreement between agricultural and urban sectors.

We develop a general-equilibrium representation of a hydrologic-ecological-economic system to model the post-trade scenario of the water exporting region. This model predicts that in the post-trade period, a decline in water availability would result in the reduction of both skilled and unskilled labor in the water exporting region. In contrast to the reduction of employment, an increase in water values would be observed simultaneously in the post-trade period.

In the empirical analysis, we apply a synthetic control methodology to test the predictions of the water trading model previously developed. Two sets of outcome variables are tested in the empirical analysis. We explore the direct effect of the water transfer between the pre-trade and post-trade periods by comparing the crop production statistics, including harvested acreage, crop values and crop yield per acre. Not surprisingly, we notice a decrease in both harvest acreage and crop values in the post-trade period, which implies that less water availability in the agricultural sector will hamper the crop production in the water exporting region. In contrast to the cut of crop values and harvest acreage, we find

[1] Quantification Settlement Agreement, Water Education Foundation, http://www.watereducation.org/aquapedia/quantification-settlement-agreement

an increase of crop yield per acre of the water exporting region in the post-trade period. The increasing crop yield per acre is an indicator of the increasing water values in the post-trade period. We then test a set of labor statistics including skilled and unskilled labor employment and earnings in the crop-production sector. Supporting our theoretical model, a significant decline of both skilled and unskilled labor employment in the crop production sector has been observed in our empirical analysis.

The paper proceeds as follows: section two provides background on the Quantification Settlement Agreement and the fallowing program under the supervision of the QSA. Section three describes a general equilibrium model related to the water trading and provides predictions. Section four provides details on the econometric approach and results. Section five concludes the paper.

3.3 Background

The QSA was signed October 10, 2003, to resolve the basic problem of California's excess use of Colorado River water. The Colorado River has been called the most legislated and managed river in the world because of its multiple overlapping jurisdictions and strong contrast in legal and administrative styles of the two neighboring countries, the United States and Mexico (Pulwarty, Jacobs, and Dole 2005). It is the principal source of water for irrigation and domestic use in "lower basin" states: Arizona, southern California, and southern Nevada. Among the lower basin states, California was the only state that had irrigation districts that could use the water immediately (Reisner 1993). Taking this into consideration, California had been allocated 4.4 out of 7.5 million acre-feet of water among the lower basin states under the Boulder Canyon Project Act in 1928. The California Seven Party Agreement[2] was then signed to distribute the water between irrigation districts and urban water uses. This agreement was signed before rapid urban development happened in lower basin states. With the increasing population and urban expansion in southern

[2]The California Seven Party Agreement awarded the agricultural water consumers (Imperial Irrigation District, Palo Verde Irrigation, the federal Yuma Project, and the Coachella Valley County Water District) senior water rights of 3.85 million acre-feet water. Out of 3.85 million acre-feet, IID diverted 2.6 million acre-feet. The Metropolitan Water District of Southern California, the City of Los Angeles, and San Diego only have been secured 0.662 million acre-feet for any surplus.

California, the water allocation under the California Seven Party Agreement was unable to meet the needs of this growing development. The QSA was signed to reduce California's over-dependence on Colorado River and balance the inequality between agriculture and urban water uses. It provides for large-scale water transfers from IID to SDWA, Metropolitan Water District (MWD) and Coachella Valley County Water District (CVWD).[3] Moreover, as part of the QSA, the state of California is obligated to undertake the restoration of the Salton Sea ecosystem with the mitigation water saved from a voluntary fallowing program in IID (San Diego County Water Authority 2019). Figure 3.1 shows the long-term water transfer in California since 1970. It is clear from Figure 3.1 that Imperial County exported the largest amount of water to the surrounding counties between 1970 and 2016. No county is comparable to Imperial County in the amount of water exported in our sample years.

Imperial County, where IID is located, is one of the top agricultural producing counties in the United States. Agriculture-related industry is the leading employer and a fundamental component of the county's economy. The agricultural production and processing industry were the top-ranked sector in terms of direct economic output in 2016 (Ortiz and Dessert 2017). It contributed a total of $4.5 billion to the local economy and created 24,429 jobs in 2016. Imperial County is located in an arid region near the Mexico and Arizona borders. It borders San Diego County to the west and Riverside County to the north. The Colorado River is the main source of irrigation water in the county, and roughly 2.8 million acre-feet fresh water has historically been diverted to the Imperial Irrigation District for agricultural irrigation.

To meet the water conservation requirement of the QSA, Imperial County ultimately has agreed to promote a 15-year voluntary fallowing program. The goal of this program is to eliminate potential side effects to the Salton Sea caused by the large amount of water transfer out of Imperial County. The Salton Sea is fed solely by agricultural run-off and provides important habitat for fish and migratory birds. However, the shoreline has receded and the lake bed has been exposed due to diminished water supply. IID sends water to the

[3] IID-SDCWA transfer ramps up to 200,000 acre-feet per year in 2021 for up to 75 years. IID-MWD transfer is 105,000 acre-feet per year. IID-CVWD transfer ramps up to 103,000 acre-feet per year.

Salton Sea to mitigate for environmental impacts associated with reduced crop irrigation runoff. The water fallowing program requires reducing water use without decreasing crop production. The qualified farmers are supported by payments for water conservation using the money generated by water sold to the San Diego County Water Authority. This program was posited to have no direct changes to the local economy due to the unchanging crop production overall but may create costs through indirect mechanisms.

Figures 3.2 and 3.3 show the monthly and annual fallowing volume in Imperial County since the beginning of the water transfer. In Imperial County, the water fallowing program officially started at the end of 2003 and only 3,445 AF water had been conserved due to the voluntary fallowing program by December 2003. Consequently, the post-trade period for this paper begins in 2004. The fallowing program can be separated into two periods. In the first period (2004–2011), there is no evident seasonality of monthly volume or big spikes in the annual volume in the Imperial County. However, in the second period (2012-2016), we notice a clear seasonality trend in monthly volume. In addition, a spike is observed in year 2014 in Figure 3.3. This overall observation of fallowing is important for understanding our theoretical and empirical results in Section 4 and 5.

3.4 Water Trading Model

In this section we develop a general-equilibrium representation of a simple coupled hydrologic-ecological-economic system. These types of simple models are commonly used in economics to provide analytical insight into complex behavior. We assume a single regional economy (Imperial County in this case) with three sectors, the agricultural sector (A), an ecosystem service-based sector (S), and manufacturing sector (M). Among these three sectors, the agricultural sector is the domain sector with the biggest share of labor. Assume regional water availability follows the equation of motion:

$$\frac{dW}{dt} \equiv \dot{W} = \bar{\sigma} - f(W) - W_A - W_M \tag{3.1}$$

W: A measure of the amount of water

$\bar{\sigma}$: Water inflow

$f(W)$: Water outflow

W_A: Amount of water used in the agricultural sector

W_M: Amount of water used in the manufacturing sector

The inflow and outflow terms represent areas of coupling between the economic system and the natural system. $f(W)$ is a key coupling feature, linking our understanding of the hydrologic system to the coupled model. For the purposes of analytic understanding, we represent this term as a function of the amount of water in the system.

The regional economy consists of three primary production sectors, the agricultural sector (Q_A), manufacturing sector (Q_M) and ecosystem sector (Q_S). Each sector requires unskilled and/or skilled labor, U and L, respectively. We assume different sectors have preference to hire labor from different levels of skill, such that skilled labor can only be hired in the agricultural and manufacturing sectors while unskilled labor can be hired in the agricultural and ecosystem service sectors (e.g. services related to the natural system). Labor is assumed to be freely mobile across sectors, meaning $\bar{L} = L_A + L_M$ and $\bar{U} = U_A + U_S$. The agricultural sector produces output using labor and water, which are the same factors used by the manufacturing sector. Different from the agricultural and manufacturing sectors, the service sector uses unskilled labor, and water remains in the system. Hence, the amount of water used in production is $\bar{W} = W_A + W_M$. Production technologies are represented by the following production functions:

$$Q_A = Q_A(U_A, L_A, W_A) \tag{3.2}$$

$$Q_M = Q_M(L_M, W_M) \tag{3.3}$$

$$Q_S = Q_S(U_S, W) \tag{3.4}$$

3.4.1 Regional Trade in Water

We can add a second regional economy that works in a similar way, albeit with a potentially different equation of motion and different ecosystem, agricultural and manufacturing production functions. Since we only focus on the water selling region, we only define

the production function of the water selling region here. Given the perfect competition assumption, zero-profit condition implies that:

$$\begin{aligned}
(3.5) \quad \bar{P_A} &= \alpha_{L_A}\gamma_L + \alpha_{W_A}\gamma_W + \alpha_{U_A}\gamma_U \\
(3.6) \quad \bar{P_S} &= \alpha_{U_S}\gamma_U \\
(3.7) \quad \bar{P_M} &= \alpha_{L_M}\gamma_L + \alpha_{W_M}\gamma_W
\end{aligned}$$

$\bar{P_A}$, $\bar{P_S}$ and $\bar{P_M}$ denote the price of agricultural, service and manufacturing output respectively, γ_L and γ_U standing for wage rate for skilled and unskilled labor, γ_W for the water price, and α_{L_i}, α_{U_i} and α_{W_i} for the respective per-unit amount of skilled and unskilled labor and water wage in sector i, $i = c(A, S, M)$.

Moreover, full employment condition implies:

$$\begin{aligned}
(3.8) \quad \bar{U} &= \alpha_{U_A}Q_A + \alpha_{U_S}Q_S \\
(3.9) \quad \bar{L} &= \alpha_{L_A}Q_A + \alpha_{L_M}Q_M \\
(3.10) \quad \bar{W} &= \alpha_{W_A}Q_A + \alpha_{W_M}Q_M
\end{aligned}$$

By total differentiating equation 3.5 , using $\hat{x} = dx/x$, we get:

$$(3.11) \quad \theta_{L_A}\hat{\gamma}_L + \theta_{U_A}\hat{\gamma}_U + \theta_{W_A}\hat{\gamma}_W = \hat{\bar{P}}_A - (\theta_{L_A}\hat{\alpha}_{L_A} + \theta_{U_A}\hat{\alpha}_{U_A} + \theta_{W_A}\hat{\alpha}_{W_A})$$

where θ_{ji} is the factor j's share in sector i, e.g., $\theta_{U_i} = \frac{\gamma_U \alpha_{U_i}}{\bar{P_i}}$. At the equilibrium, $\theta_{L_A}\hat{\alpha}_{L_A} + \theta_{U_A}\hat{\alpha}_{U_A} + \theta_{W_A}\hat{\alpha}_{W_A} = 0$. Thus, we get:

$$(3.12) \quad \theta_{L_A}\hat{\gamma}_L + \theta_{U_A}\hat{\gamma}_U + \theta_{W_A}\hat{\gamma}_W = \hat{\bar{P}}_A$$

According to the definition of the α_{U_S}, we have relationship, $\alpha_{U_S} = \frac{U_S}{Q_S}$. Solving together with the equation 3.4 and 3.6, we get:

$$(3.13) \quad \gamma_U = W \times \bar{P_S}$$

By total differentiating equation 3.13, we get:

$$\hat{\gamma}_U = \hat{\hat{P}}_S + \hat{W} \tag{3.14}$$

A similar process is followed to yield:

$$\theta_{L_M}\hat{\gamma}_L + \theta_{W_M}\hat{\gamma}_W = \hat{\hat{P}}_M \tag{3.15}$$

Therefore, equation 3.12, 3.14 and 3.15 can be reduced to the simple matrix form

$$\begin{bmatrix} \theta_{U_A} & \theta_{W_A} & \theta_{L_A} \\ 0 & \theta_{W_M} & \theta_{L_M} \\ 1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \hat{\gamma}_U \\ \hat{\gamma}_W \\ \hat{\gamma}_L \end{bmatrix} = \begin{bmatrix} \hat{\hat{P}}_A \\ \hat{\hat{P}}_M \\ \hat{\hat{P}}_S - \hat{W} \end{bmatrix} \tag{3.16}$$

Since the regional economy of Imperial County is a small open economy, the price is exogenous and we have $\hat{\hat{P}}_A = \hat{\hat{P}}_M = \hat{\hat{P}}_S = 0$.

$$\begin{bmatrix} \theta_{U_A} & \theta_{W_A} & \theta_{L_A} \\ 0 & \theta_{W_M} & \theta_{L_M} \\ 1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \hat{\gamma}_U \\ \hat{\gamma}_W \\ \hat{\gamma}_L \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \hat{W} \end{bmatrix} \tag{3.17}$$

Solving the equation 3.17, we get:

$$\hat{\gamma}_U = \hat{W} \tag{3.18}$$

$$\hat{\gamma}_W = -\frac{\theta_{U_A}\theta_{L_M}}{\theta_{W_A}\theta_{L_M} - \theta_{L_A}\theta_{W_M}}\hat{W} \tag{3.19}$$

$$\hat{\gamma}_L = \frac{\theta_{U_A}\theta_{W_M}}{\theta_{W_A}\theta_{L_M} - \theta_{L_A}\theta_{W_M}}\hat{W} \tag{3.20}$$

The sign of γ_W and γ_L is determined by the sign of equation $\det\theta = \theta_{W_A}\theta_{L_M} - \theta_{L_A}\theta_{W_M}$. Rewriting this equation with the factor shares, we have:

$$
\begin{aligned}
(3.21)\quad \det\theta &= \theta_{W_A}\theta_{L_M} - \theta_{L_A}\theta_{W_M} \\
&= \frac{\gamma_W\alpha_{W_A}}{\bar{P_A}} \times \frac{\gamma_L\alpha_{L_M}}{\bar{P_L}} - \frac{\gamma_L\alpha_{L_A}}{\bar{P_A}} \times \frac{\gamma_W\alpha_{W_M}}{\bar{P_M}} \\
&= \frac{\gamma_L\gamma_W}{\bar{P_A}\bar{P_M}} \times (\alpha_{W_A}\alpha_{L_M} - \alpha_{L_A}\alpha_{W_M}) \\
&= \frac{\gamma_L\gamma_W}{\bar{P_A}\bar{P_M}} \times (\frac{W_A}{Q_A}\frac{L_M}{Q_M} - \frac{L_A}{Q_A}\frac{W_M}{Q_M}) \\
&= \frac{\gamma_L\gamma_W L_M L_A}{\bar{P_A}\bar{P_M}Q_A Q_M} \times (\frac{W_A}{L_A} - \frac{W_M}{L_M})
\end{aligned}
$$

Given that agricultural sector is more water intensive than the manufacturing sector (i.e. $\frac{W_A}{L_A} > \frac{W_M}{L_M}$), and $\det\theta = \theta_{W_A}\theta_{L_M} - \theta_{L_A}\theta_{W_M} > 0$. Therefore, for a water exporting region, $\hat{W} < 0$, the relationship of labor-water equation system in Imperial County is:

$$
\begin{aligned}
(3.22)\quad \hat{\gamma}_U &= \hat{W} < 0 \\
(3.23)\quad \hat{\gamma}_W &= -\frac{\theta_{U_A}\theta_{L_M}}{\theta_{W_A}\theta_{L_M} - \theta_{L_A}\theta_{W_M}}\hat{W} > 0 \\
(3.24)\quad \hat{\gamma}_L &= \frac{\theta_{U_A}\theta_{W_M}}{\theta_{W_A}\theta_{L_M} - \theta_{L_A}\theta_{W_M}}\hat{W} < 0
\end{aligned}
$$

We conclude the following predictions from our theoretical model.

i) Water trading raises the return to water in the water exporting region, referring to equation 3.23.

ii) Water trading reduces the return to skilled and unskilled labor in the water exporting region, referring to equations 3.22 and 3.24.

In the following empirical analysis, we apply the synthetic control methodology to test if the two predictions we proposed above hold in the water exporting region. If the two predictions hold, we would expect to observe a statistically significant increase in the water values and decrease in the employment of both skilled and unskilled labor in the agricultural sector.

3.5 Empirical Framework

We are interested in the impact of water trading on local economies and water conservation in California. Imperial County is the treatment unit in our model because it consistently traded a large amount of water to San Diego County, and no other county is comparable to it in terms of water trading volume. The selection of control groups is crucial but difficult. If control groups are not sufficiently similar to the treatment county, any difference in outcome variables between treated counties and control counties may be caused by the inconsistencies in their characteristics (Abadie, Diamond, and Hainmueller 2010). The synthetic control approach uses a panel dataset to construct a convex combination of control units that best represents characteristics of the treated groups prior to the intervention. The effect of the intervention is quantified by the difference of outcome variables in the post-intervention period between treated and synthetic groups.

Abadie, Diamond, and Hainmueller (2010) discuss the theoretical properties of the synthetic control method comprehensively and our model is derived from this formulation. Let y_{it} be the set of outcome variables, in county i with $i = 0$ for Imperial County and $i > 0$ for all the other counties in California, in year t for time periods $t = 1, \ldots, t_0, \ldots, T$, where t_0 is the water transfer starting year, 2004 in our case. Let $y_{it}^{noTrade}$ be the outcome that would be observed for county i at time t without the water trading, for counties $i = 1, \ldots, N$, and time periods $t = 1, \ldots, T$. Let y_{it}^{Trade} be the outcome that would be observed for county i at time t if county i is subjected to the water trading in periods $t_0 + 1$ to T. We assume that the water trading has no effect on the outcome before 2004, i.e., $t \in \{1, \ldots, t_0\}$ and all $i \in \{1, \ldots, N\}$. Let $\delta_{it} = y_{it}^{Trade} - y_{it}^{noTrade}$ be the effect of the water trading for county i at time t, and let D_{it} be an indicator that equals to one if county i is subjected to the water trading at time t and zero otherwise. In practice, the water trading may have an impact prior to implementation. We assume that outcomes of the untreated counties are not affected by the water trading implemented in the treated counties. The observed outcome for county i at time t is:

$$y_{it} = y_{it}^{noTrade} + \delta_{it} D_{it} \tag{3.25}$$

Since only the Imperial County ($i = 0$) is subjected to the QSA and only after period t_0 ($t_0 = 2004$), we have:

$$D_{it} = \begin{cases} 1 \quad if \quad i = 0 \quad and \quad t > t_0 \\ 0 \quad if \quad otherwise \end{cases}$$

We aim to estimate $\delta_0 = (\delta_{0,t_0+1}, \delta_{0,t_0+2}, \ldots, \delta_{0,T})'$. For $t > t_0$:

$$(3.26) \qquad \delta_{0t} = y_{0t}^{noTrade} = y_{0t} - y_{0t}^{noTrade}$$

Since y_{0t}^{Trade} is observed, to estimate δ_{0t}, we need to estimate $y_{0t}^{noTrade}$. Suppose that $y_{it}^{noTrade}$ is given by a factor model, that is:

$$(3.27) \qquad y_{it}^{noTrade} = \eta_t + \boldsymbol{\theta}_t \boldsymbol{Z}_i + \lambda_t \mu_i + \varepsilon_{it}$$

where η_t is an unknown common time effects, $\boldsymbol{Z}_i$ is a $r \times 1$ vector of observed covariates (not affected by the water trading policy), μ_i are permanent unobserved variables, $\boldsymbol{\theta}_t$ and λ_t are unknown parameters, and ε_{it} are unobserved error term with zero mean.

Consider a $N \times 1$ vector of weights, $\boldsymbol{W} = (w_1, \ldots, w_N)' \in [0,1]^N$, summing to 1, to minimize distance in pre-trade characteristics between treated and weighted average of controls. Estimated treatment effect is the simple difference between them, i.e.:

$$(3.28) \qquad \hat{\delta}_{0t} = y_{0t} - \sum_{i=1}^{N} w_i y_{i,t} \quad for \quad t = t_0 + 1, t_0 + 2, \ldots, T$$

We may choose the vector $\boldsymbol{W}^* = (w_1^*, \ldots, w_N^*)'$ satisfying,

$$(3.29) \qquad \sum_{i=1}^{N} w_i^* \boldsymbol{Z}_i = \boldsymbol{Z}_0, \sum_{i=1}^{N} w_i^* y_{i,1} = y_{0,1}, \ldots, \sum_{i=1}^{N} w_i^* y_{i,t_0} = y_{0,t_0}$$

The vector $\boldsymbol{W}^*$ is chosen to minimize the mean squared prediction error (MSPE) of

the set of outcome variables for the pre-trade periods.[4] Once we find the vector $\boldsymbol{W}^*$ which minimizes equation 3.29, we can estimate the effect of the water trading using equation 3.28.

3.6 Data

Our outcome of interest can be divided into two groups, labor market statistics in the crop production sector and crop production statistics. We choose the subsector crop production, because this sector is most affected by the QSA fallowing program officially started in 2004. Four labor market indicators, including skilled labor employment, unskilled labor employment, skilled labor earnings, and unskilled labor earnings, in the crop production sector are selected as the outcome variables to represent the impacts on both skilled and unskilled labor. Skilled labor here refers to the labor with high school degree or higher, while the unskilled labor represents the labor without a high school degree. In addition, annual acreage harvested,[5] annual crop values, and annual crop yield per acre are selected to represent the level of crop production. Table 3.1 shows the summary statistics of outcome variables and predictors. The dataset contains one treatment county, Imperial County, and 56 control counties. We exclude Alpine County in the analysis because it lacks data for all the outcome variables in the sampling period. The synthetic control requires a set of predictors to construct a synthetic Imperial County by calculating the weighted average of potential control counties to best reproduce the outcomes of interest in the pre-trade period. Figure 3.1 shows the annual aggregate level long-term water transfer among California counties since 1970. The vertical line represents the starting year of water transfer. Thirteen counties, including Imperial County, have long-term water transfer programs between 1992 and 2016, but none of the other counties are comparable in volume of water delivery to Imperial County in the post-trade period. Because the synthetic control is meant to

[4] The MSPE is the mean of the squared deviations between the outcome for the treated counties and the synthetic counties in all pre-QSA periods, $MSPE = \frac{1}{t_0-1}\sum_{t=1}^{t=t_0-1}(y_{1,t} - w_i^* y_{i,t})^2$.

[5] Due to the high skewness and wide range of the selected outcome variables, the labor statistics, including skilled labor employment, unskilled labor employment, skilled labor earnings, unskilled labor earnings, and annual harvested acreage all fails to converge for the Imperial County. We apply the log transformation to rescale them to make the relationship clearer.

reproduce the outcomes of interest that would have been observed for Imperial County in the absence of water transfer, and Imperial County had some transfers prior, we include all the other counties in our donor pool. We use annual county-level panel data for the period 1992 – 2016. The QSA was passed in October of 2003, IID was only able to fallow a small amount of land in that year, and therefore we use 2004 as the treatment start date, giving us 12 years of pre-intervention observations (1992-2003) and 13 years of post-intervention data (2004-2016).

Data on quarterly average employment and monthly earnings are collected from the Longitudinal Employer-Household Dynamics (LEHD) program of the United States Census Bureau.[6] The LEHD provides the quarterly county-level labor information in the United States from 1992. Quarterly employment is the estimate of the number of jobs that are held on both the first and last day of the quarter with the same employer, while quarterly earnings are the average monthly earnings of employees with stable jobs, measured in monthly earnings per capita within the one quarter. The quarterly average employment and monthly average earnings are calculated by average the stable quarterly employment and monthly earnings in each quarter within the same calendar year.[7]

The visualization of quarterly employment and monthly earnings in each quarter are shown in Figure B.1 in the Appendix B. We notice an unusual spike of both earnings and employment in 2002 but it goes back to usual in the following year. This indicates a data issue of the 2002 labor data. This spike occurs in the pre-intervention period, which means the imputation of problematic data will actually improve the reliability of the synthetic counterfactual method. Then we use the quarterly data of 2001 and 2003 to impute the labor data of 2002. Taking the first quarter of 2002 for an example, the imputation equations

[6] https://ledextract.ces.census.gov/static/data.html

[7] We are only able to collect quarterly data for employment and monthly average earnings for each quarter. We turn the quarterly employment to quarterly average employment using the equation below (taking the employment in 1992 as an example): $Employment_{1992} = 1/4(Employment_{Q1-1992} + Employment_{Q2-1992} + Employment_{Q3-1992} + Employment_{Q4-1992})$. I turn the monthly earnings for each quarter data into annual data using the equation below: $Earnings_{1992} = 1/4(Earnings_{Q1-1992} + Earnings_{Q2-1992} + Earnings_{Q3-1992} + Earnings_{Q4-1992})$

are:

$$Employment_{Q1-2002} = 1/2(Employment_{Q1-2001} + Employment_{Q1-2003})$$
$$Earnings_{Q1-2002} = 1/2(Earnings_{Q1-2001} + Earnings_{Q1-2003})$$

The quarterly employment and monthly earnings in each quarter after imputation is shown in Figure B.2 in the Appendix B. It is clear that there is no unusual spike but the seasonality still holds in the data.

The annual acreage harvested, annual yield per acre and annual crop values are collected from the annual report of USDA's National Agricultural Statistics Service California Field Office.[8] This is an annual crop report compiled by the California County Agricultural Commissioners (CCAC) with providing the most detailed annual agricultural production data in the county level.

The set of predictors to identify the synthetic Imperial County is listed in the Table 3.1. The predictors to control local economic development, including farm proprietor's income and employment,[9] wage and salaries and wage salary employment,[10] and proprietor's income and employment,[11] are obtained from Bureau of Economic Analysis (BEA).[12] The Ag-labor ratio characteristics including the ratio of white ag-labor, the ratio of male ag-labor, the ratio of Hispanic ag-labor and the ratio of ag-labor with high school or higher education

[8] https://www.nass.usda.gov/Statistics_by_State/California/Publications/AgComm/index.php

[9] According to BEA's definition, farm proprietors' income consists of the income that is received by the sole proprietorships and the partnerships that operate farms. It excludes the income that is received by corporate farms. Farm proprietor's employment consists of sole proprietors and non-corporate partners in the farm industry. The farm proprietors' income is in millions of dollars and farm proprietors' employment is in thousands of jobs.

[10] Wage and salary employment measures the average annual number of full-time and part-time jobs in each area by place of work. Wage and salaries is in billions of dollars and wage and salary employment is in thousands of jobs.

[11] Proprietors' income is the current-production income (including income in kind) of sole proprietorships, partnerships, and tax-exempt cooperatives. Corporate directors' fees are included in proprietors' income. Proprietors' income includes the interest income received by financial partnerships and the net rental real estate income of those partnerships primarily engaged in the real estate business. The proprietors' employment consists of farm proprietors' employment and nonfarm proprietors' employment. The proprietor's income is in billions of dollars and proprietor's employment is in number of jobs.

[12] https://apps.bea.gov/regional/histdata/releases/1117lapi/index.cfm

degree is obtained from the LEHD.[13] The county-level agricultural characteristics are from two different data sources. Total irrigation water withdrawals and total irrigation acres are from the United States Geological Survey (USGS),[14] while the total agricultural acres and total agricultural sales are from the United States Department of Agriculture (USDA) National Agricultural Statistics Services.[15] Lastly, the county-level annual agricultural production data, including cattle values, alfalfa hay values, lettuce values, melon values and other vegetable values, to control the agricultural output of synthetic Imperial County, is from the annual report of USDA's National Agricultural Statistics Service California Field Office. The summary statistics of all the predictors are listed in Table 3.1.

The synthetic control methodology described earlier is used to construct a synthetic Imperial County that best predicts the pre-QSA outcomes of interest. We estimate the effect of water trading as the difference in the seven outcome variables between Imperial County and synthetic Imperial County in the post-QSA period. We then perform the placebo tests to each outcome variable following Abadie, Diamond, and Hainmueller (2010) to test whether the estimated effects of Imperial County are unusually large relative to other counties by reapplying the same analysis to all the other control counties.

3.7 Results

3.7.1 Trends in Outcome Variables

Figure 3.4 and 3.5 illustrate the comparison of annual average value of seven outcome variables between Imperial County and the average of the other 56 counties of California

[13] LEHD provides the number of stable jobs in agricultural sector of different demographic characteristics and the ratio of ag-labor is based on author's calculation.

[14] `https://water.usgs.gov/watuse/data/index.html`. USGS maintains the county-level water use information of the nation from 1985. This data is collected every five year. The data year in our sample is 1995, 2000, 2005, 2010, 2015. Total irrigation acres is in thousands of acres and the total irrigation water withdrawals is in Mgal/d.

[15] `https://quickstats.nass.usda.gov/#351328CE-C207-3092-ABF4-C45A89232234`. The county-level agriculture characteristics of California is collected by USDA every 5 year starting from 1997. The data year in our sample is 1997, 2002, 2007, 2012. The total agricultural acres is in thousands of acres and the total crop sales is in millions of dollars.

(excluding Alpine County due to missing data).[16] The blue line represents the average value of one outcome variable of the other 56 counties in California, while the black line represents the value of that outcome variable in Imperial County. The vertical line represents the beginning of trade year. The two plots on the top of Figure 3.4 show that Imperial County has significantly higher harvested acreage (logged) and crop values compared to the other 56 counties in California. The plot on the bottom of Figure 3.4 shows an unusual spike in crop yield per acre in 2004 in our sample of the other 56 counties. This spike could be caused by the extreme data points in one or two counties. However, synthetic control method uses the predictors to find the optimal combination of the other 56 counties to best-replicated synthetic Imperial County, the extreme point will be automatically weighted as zero in our synthetic Imperial County. This argument can be supported when we take a look at the plot on the bottom of Figure 3.6. After calculating the weighted average of the other 56 counties, there is no spike in 2004 of synthetic Imperial County (dashed line in Figure 3.6 bottom).

Two plots on the left of Figure 3.5 shows the comparison of monthly earnings between Imperial County and the other 56 counties in California. These two plots tell us that both skilled and unskilled labor earnings (logged) in the crop production sector is higher than the average of the other 56 counties. In addition, the earnings for both skilled and unskilled labor increase gradually from 1992 to 2016. The gaps of both skilled and unskilled labor between the Imperial County and the average of the other 56 counties are reduced within the sampling period. Two plots on the right of Figure 3.5 shows the comparison of quarterly employment between Imperial County and the other 56 counties in California. The skilled and unskilled labor employment are stable in the sampling period. Not surprisingly, Imperial County has more labor employment of both skilled and unskilled labor in the crop production sector compared to the average of the other 56 counties. This supports that average value of the other 56 counties in California is a suitable control group for Imperial County to study the seven selected outcome variables. The big gaps between Imperial

[16] All the other counties in California refers to the other 56 counties except Imperial County (treatment county) and Alpine County (excluded due to missing values).

County and the other 56 counties motivate us to apply the synthetic control methodology to search for a comparable synthetic Imperial County, especially in the pre-trade period.

The synthetic Imperial County is constructed by a convex combination of the other 56 counties that most closely predicts Imperial County in terms of outcome variables in the pre-trade period. Table 3.2 and 3.3 illustrate predict means of observed covariates which is not affected by water transfer. The first column represents the mean of observed covariates of Imperial County, while the last column represents the average of the other 56 control counties. There is an evident difference of all the observed covariates between Imperial County and average of the other 56 control counties. However, after the convex combination of the other 56 control counties, the predict means of synthetic Imperial County is much similar to the predict means of Imperial County in most of observed covariates. For example, farm proprietors' income in the pre-trade period is much lower in the average of the 56 controls ($71,573) than in Imperial County ($240,711), but it is closer in synthetic Imperial County ($244,673) (see Table 3.2).

However, some observed covariates, such as total irrigation water withdrawals, cannot reproduce the pre-trade scenario accurately with the mean of 2088.288 Mgal/d in Imperial County and 411.734 Mgal/d in average of 56 controls and 1336.1 Mgal/d in the synthetic Imperial County for the test of harvested acreage (logged) (see Table 3.3 row 11). To minimize the MSPE of outcome variables during the pre-trade period, different weights of each observed covariates are signed. For example, since total irrigation water withdrawals fails to mirror the pre-trade scenario accurately, it does not have considerable weights (weights varying from 0 to 0.076) to predict the outcome variables of synthetic Imperial County in the pre-trade period (see Table 3.4 row 11).

Table 3.5 shows the weights of each county in the synthetic Imperial County. The column 1 to 7 illustrate the best combination of counties to reproduce the synthetic Imperial County for different outcome variables. We hide the counties with zero weight in this table to diminish the table size. For example, the column 1 tells us that the combination of Inyo, Kings, Mono, Monterey and Tulare will minimize the MSPE in the pre-trade period when

reproducing the synthetic Imperial County.

3.7.2 Impacts on Crop Production

Figure 3.6 demonstrates the time paths of outcome variables related to crop production output between Imperial County and the synthetic Imperial County in the sample period. The annual harvested acreage, annual crop values and annual yield per acre are reported here. These three variables are closely affected by the fallowing program and we gather these three variables to indicate the direct impact of fallowing program but using our synthetic counterfactual.

The synthetic Imperial County is able to successfully track the trajectory of Imperial County in the pre-trade period. Compared to the significant gaps between the outcome variables in the pre-trade period in Figure 3.4, this implies that the synthetic Imperial County provides an approximately accurate estimate of selected outcome variables in the scenario in the absence of the water transfer. The time path of harvested acreage (logged) starts to diverge around 2007, three years after starting transfer water (See subfigure on the top). What's more, the gap between the Imperial County and the synthetic Imperial County becomes larger the longer time has passed post water transfer. It shows up more clearly when we plot the gap between Imperial County and synthetic Imperial County. The gap plots can be found in Figure 3.7. Since a good match of Imperial County and synthetic Imperial County in the pre-trade period will show a small gap between them, Figure 3.7 states that two of the outcome variables, the annual harvested acreage (logged) and annual crop values, both have a relatively small gap in the pre-trade period and an obviously negative gap in the post-trade period. This suggests that the overall estimated effect of water transfer on the annual harvested acres and annual crop values is negative in Imperial County.

When we compare the gap plots of harvested acreage (logged) and annual crop values to the amount of land fallowed in Figure 3.3, we notice striking similarities. This suggests that the harvested acreage and crop value decreases could be the result of the fallowed land being taken out of production. In particular, we observe a sharp decrease in all three plots

around the year of 2014, which was the year with the greatest amount of fallowing. These results give us some confidence in the ability of our synthetic counterfactual to pick up changes as a result of the water transfer.

3.7.3 Labor in the Crop Production Sector

Figures 3.8 and 3.9 show the time path and gap plots of labor statistics between Imperial County and synthetic Imperial County. Figure 3.8 left panel top to bottom shows the time paths of skilled labor earnings and unskilled labor earnings separately. There are no evident differences in skilled labor earnings between Imperial County and synthetic Imperial County, but a small negative effect can be found in the unskilled labor earnings in the post-QSA period. This finding can be supported when we switch to the gap plots in Figure 3.9 (left panel top to bottom). In contrast to the invisible gaps in the earnings, we observe evidently negative gaps in both skilled and unskilled labor employment in both 3.8 and 3.9 (right panel top to bottom).

Again, to compare the gap plots of labor statistics with the fallowing land in Figure 3.3, we find an even more interesting relationship existing between the skilled and unskilled labor employment and the land fallowed each year. The unskilled labor employment strictly follows the very similar trajectory of Figure 2B, while the skilled labor employment loosely follows this trajectory at the same time. To understand this, we need to record our water trading model in the sector three, especially equation 3.22 and 3.24. What the model tells us is that the amount of the skilled and unskilled labor loss in the water exporting region ideally should be equal to the amount of water trading. These interesting results suggest that the fallowing program might have a negative impact to both the skilled and unskilled labor employment in the water selling region, in our case, Imperial County. However, since the farmers will get the direct benefits, such as water payment, due to the water trading, the QSA program will not affect the labor earnings at the visible level. However, these figures tell us nothing about the significance of our results, that is, we are not able to tell if these gaps plot in the Figure 3.7 and Figure 3.9 are statistically significant from zero or not in the post-trade period. In the next section, we run placebo tests that can be used for

hypothesis testing.

3.8 Placebo Tests

In order to get the significance level of our estimation, we perform a placebo test by reapplying the synthetic control method to each county that is not affected by QSA or any comparable large-scale water transfer program during the sampling period. The placebo test is suggested by Abadie, Diamond, and Hainmueller (2010) in their paper evaluating the effect of California's tobacco control program. The reason for the placebo test is to see whether effects similar in or of larger magnitude exist for the same treatment date in the control counties. If the control counties have comparable or larger magnitude of estimated effects when treating them as treatment counties, then the estimated effects of our selected outcome variable in the post-trade period are not statistically significant. If the differences of the control counties are smaller than Imperial County, it provides evidence of significance. The p-value from the test can be interpreted as "the probability of obtaining an estimate at least as large as the one obtained for the unit representing the case of interest when the intervention is reassigned at random in the data set" (Abadie, Diamond, and Hainmueller 2015, p. 500).

To conduct the placebo test we run the synthetic control method separately on each control county treating them as a treatment county when using all the other remaining counties, including Imperial County, as before. The gap between the placebo county and its synthetic county should be randomly signed in the results because none of the control counties is actually exposed to the fallowing program. We can evaluate the probability that the treatment effect exists by comparing the gap between Imperial County and synthetic Imperial County (black line) to the gaps of the placebo counties and their synthetic controls (grey lines). The placebo counties with the large pre-trade MSPE, we define as the poorly fitted units. A large pre-trade MSPE value indicates that the pre-trade time path of this placebo county is not accurately reproduced by the convex combination of the remaining 56 controls. Hence, these poorly fitted values should be deleted when comparing significance level of treatment effect. We exclude the counties that have pre-trade MSPE of more than

20 times, 10 times, 5 times, and 2 times the MSPE of Imperial County. The results with counties of more than two times pre-trade MSPE of Imperial County are reported in the Figure 3.10 and 3.11, while the other results can be found in Figure B.3 to B.10 in the Appendix B. The 3.10 and 3.11 suggest that among the counties remaining in the figure, the Imperial County gap line is about the most unusual line in crop values (see Figure 3.10), skilled labor employment and unskilled labor employment.

Another way to evaluate the significance of the Imperial County gap relative to the gaps obtained from placebo runs is to rank the ratios of post/pre-trade MSPE or RMSPE. The advantage of this methodology is that it removes choosing a cut-off for the exclusion of poorly fitted placebo runs. Hence, the rank of the MSPE or RMSPE ratio can represents the significant level of the treatment effects.

Figures 3.12 to 3.18 show the rank of MSPE and RMSPE ratio of different outcome variables. The plot of MSPE ratio rank is presented on the left of each figure, while the plot of RMSPE ratio rank is presented on the right. The ratio for Imperial County ranked on the top in the skilled labor employment (logged) (see Figure 3.14). Post-trade MSPE is about 25 times the MSPE for the pre-trade period. Only Tuolumne and Santa Barbara are ranked higher than it. If one county were to assign the water transfer program at random in the dataset, its probability of obtaining a post/pre-trade MSPE or RMSPE ratio as large as Imperial County is $3/55 = 5.5\%$. However, surprisingly, the ranks of the other two significant outcome variables, unskilled labor employment and crop values, are not high in the second placebo test, with a $12/54 = 22.2\%$ in unskilled labor employment and $13/53 = 24.5\%$ in crop value. The reason could be that even Imperial County has relative large post-trade MSPE in unskilled labor employment and crop values, but its pre-MSPEs of these two outcome variables are relative large due to limited sample size and short pre-trade sample years. We can improve the results of these two variables by adding more sampling year or extending our sample from California to the West of US.

3.9 Conclusion

This paper explores the causal effect of the largest agriculture-to-urban water transfer in

US history. A general-equilibrium representation of a simple coupled hydrologic-ecological-economic system is constructed in the theoretical section. The water trading model is between two small open economies of three separate sectors (Agricultural, Manufacturing, and Ecosystem sectors) in each economy. One economy is rural with agricultural sector as the domain sector which has the biggest share of labor. The rural economy is in a water abundant region exporting water to the urban economy. The water trading model demonstrates that after exporting a large amount of water, the return to water in the water exporting region will increase while the return to skilled and unskilled labor in the water exporting region will decrease.

The synthetic control methodology is apply to test the validity of the water trading model. The goal of the empirical framework is to explore empirically whether the existence of water transfer program will negatively impact the labor employment and earnings in the crop production sector. Our results show that there is a decline in the number of skilled and unskilled jobs in Imperial County after the water transfer while the average earnings of both skilled and unskilled labor do not change after trade. These results imply that due to benefits of water payment from water trading, the loss of water will not affect the labor earnings at the visible level in the water exporting region. However, voluntary fallowing program decreases the demand of both skilled and unskilled on-farm labors in the water exporting region. As a results of water trading, labor will move in the same direction as water. At the same time, the overall crop production shrinks in response to limited water availability. We observe a statistically siginificant decrease of crop value in the Imperial County after water transfer. The higher return to water is indicated by the increased crop yields, but this result is not statistically significant. This result can be improved by adding more pre-trade sample years or extending the research region from California to the West of US.

The increased volume of water trading appears to have reduced water availability in Imperial County, as predicted by the model, leading to corresponding declines in the Salton Sea. One of the important goals of this large-scale water transfer program is to restore the

Salton Sea ecosystem with the mitigation water saved from efficient irrigation practice in IID. However, this year is the 15th anniversary of the signing of the QSA, the problems Salton Sea is facing still remain unsolved. Less water availability caused by the large water transfer program has potentially reduced the water inflow of Salton Sea. This might worsen the salinity problem of Salton Sea because there is no outlet for Salton Sea and water is lost only through evaporation.

In conclusion, reductions in employment and environmental damage may be attributable to water transfers. In order to mitigate the rapid decline of both skilled and unskilled labor, more job opportunites should be created in the water exporting region. Instead of letting seasonally unemployed agricultural labor move out of the region, unemployed labor because of fallowing program should be placed into other related sectors. Another alternative policy could focus on targeting low agricultural and environmental productivity water for transfers to minimize these impacts while still creating gains from trade.

Table 3.1. Summary statistics

Statistic	N	Mean	St. Dev.
Outcome Variables			
Skilled Labor Employment	1,425	673.74	932.65
Unskilled Labor Employment	1,425	621.82	911.87
Skilled Labor Monthly Earnings	1,425	6,406.74	2,572.03
Unskilled Labor Monthly Earnings	1,425	1,608.76	603.28
Log (Harvested Acreage)	1,425	12.39	2.07
Annual Yield (per acre)	1,425	195.09	449.74
Crop Value (Million Dollars)	1,425	503.1	848.54
Predictors			
Farm proprietors' income	1,425	110.48	216.49
Farm proprietors' employment	1,425	1.32	1.41
wage and salary employment	1,425	268.66	619.83
wage and salary	1,425	12.33	30.5
proprietors' employment	1,425	73.5	174.48
proprietors' income	1,425	2.43	6.03
White Ag-labor ratio	1,425	0.8	0.2
Male Ag-labor ratio	1,425	0.66	0.2
Hispanic Ag-labor ratio	1,425	0.49	0.19
High school or higher Ag-labor ratio	1,425	0.45	0.18
Total irrigation water withdrawals	285	441.15	651.43
Total irrigation acres	285	169.57	241.5
Total agriculture acres	222	304.44	321.35
Total crop sales	216	413.62	612.62
Annual Cattle Values	1,425	41.5	92.86
Annual Alfalfa Hay Values	1,425	17.52	35.33
Annual Lettuce Values	1,425	14.15	80.77
Annual Melons Values	1,425	11.55	38.41
Annual Other Vegetable Values	1,425	13.6	34.02

Note: Employment and Earnings are all in quarterly average. Annual crop, cattle, alfalfa hay, lettuce, melons, and other vegetable values are all in millions of dollars. Total irrigation water withdrawals and total irrigation acres are in year 1995, 2000, 2005, 2010, 2015. Total agriculture acres and total crop sales are in year1997, 2002, 2007, 2012. All the other variables are in 1992-2016. Farm proprietorÕs income and employment, wage and salary employment and proprietorÕs employment are in thousands, while wage and salary and proprietorÕs income are in millions. Total agriculture acres is in thousands, and total crop sales, annual cattle values, annual alfalfa hay values, annual lettuce values, annual melon values and annual other vegetable values are in millions of dollars.

Table 3.2. Outcome variables predict means (labor Statistics)

	Imperial	Synthetic				Average of 56 Controls
		Skilled Labor Employment (logged)	Unskilled Labor Employment (logged)	Skilled Labor Monthly Earnings (logged)	Unskilled Labor Monthly Earnings (logged)	Sample Mean
Farm proprietors' income	240.711	244.673	224.419	309.873	452.907	71.573
Farm proprietors' employment	0.565	3.777	3.054	2.659	2.73	1.492
wage and salary employment	51.277	196.427	126.322	326.543	133.059	259.653
wage and salary	1.208	5.212	3.336	9.196	3.539	9.428
proprietors' employment	8.21	50.877	33.279	89.52	30.262	63.101
proprietors' income	0.469	1.417	0.993	2.315	1.288	1.837
White Ag labor ratio	0.947	0.854	0.875	0.872	0.832	0.815
Male Ag labor ratio	0.852	0.738	0.759	0.713	0.696	0.692
Hispanic Ag labor ratio	0.851	0.646	0.675	0.729	0.664	0.506
High school or higher Ag labor ratio	0.314	0.357	0.396	0.371	0.334	0.421
Total irrigation water withdrawals	2088.288	1431.895	1311.353	1180.76	955.761	411.734
Total irrigation acres	553.96	557.852	512.943	455.078	457.291	162.706
Total agriculture acres	480.661	741.701	742.996	804.88	622.196	293.337
Total crop sales	814.099	1105.093	945.804	1554.47	1606.799	384.98
Annual Cattle Values	193.202	82.4	76.923	39.937	116.951	24.383
Annual Alfalfa Hay Values	125.702	52.054	56.534	49.754	27.786	11.648
Annual Lettuce Values	56.418	12.823	21.666	75.228	160.096	7.624
Annual Melons Values	102.825	45.569	28.16	49.276	56.848	7.371
Annual Other Vegetable Values	23.753	16.981	19.507	63.608	61.793	10.863

Table 3.3. Outcome variable predict means (crop land statistics)

	Imperial	Synthetic			Average of 56 Controls
		Harvested Acreage (logged)	Crop Values	Yield Per Acre	Sample Mean
Farm proprietors' income	240.711	301.732	156.101	229.01	71.573
Farm proprietors' employment	0.565	4.098	2.025	3.567	1.492
wage and salary employment	51.277	119.571	83.283	87.535	259.653
wage and salary	1.208	2.947	2.21	2.112	9.428
proprietors' employment	8.21	29.804	17.541	19.682	63.101
proprietors' income	0.469	1.021	0.581	0.694	1.837
White Ag labor ratio	0.947	0.827	0.861	0.87	0.815
Male Ag labor ratio	0.852	0.703	0.755	0.766	0.692
Hispanic Ag labor ratio	0.851	0.647	0.626	0.664	0.506
High school or higher Ag labor ratio	0.314	0.344	0.374	0.406	0.421
Total irrigation water withdrawals	2088.288	1336.1	1333.258	1569.4	411.734
Total irrigation acres	553.96	539.367	532.087	624.43	162.706
Total agriculture acres	480.661	650.144	766.214	794.56	293.337
Total crop sales	814.099	1119.211	924.142	981.04	384.98
Annual Cattle Values	193.202	191.871	67.271	147.09	24.383
Annual Alfalfa Hay Values	125.702	50.652	50.49	60.871	11.648
Annual Lettuce Values	56.418	38.979	0.237	3.578	7.624

Table 3.4. Weights of observed covariates

	Skilled Labor Employment (logged)	Unskilled Labor Employment (logged)	Skilled Labor Monthly Earnings (logged)	Unskilled Labor Monthly Earnings (logged)	Harvested Acreage (logged)	Crop Values	Yield Per Acre
Farm proprietors' income	0.127	0.036	0.003	0.008	0.025	0.042	0.07
Farm proprietors' employment	0.068	0.1	0.206	0.194	0.004	0.151	0.061
wage and salary employment	0.022	0.027	0.045	0.047	0.064	0.075	0.051
wage and salary	0.006	0.009	0.009	0.046	0.012	0.028	0.072
proprietors' employment	0.022	0.024	0.015	0.047	0.3	0.072	0.055
proprietors' income	0.012	0.01	0.005	0.046	0.16	0.048	0.069
White Ag labor ratio	0	0.008	0.032	0.024	0.003	0.061	0.045
Male Ag labor ratio	0.001	0.039	0.009	0	0.002	0.051	0.06
Hispanic Ag labor ratio	0	0.003	0.111	0.022	0.037	0.001	0.059
High school or higher Ag labor ratio	0.077	0	0.053	0.112	0.065	0.087	0.064
Total irrigation water withdrawals	0.076	0.006	0.001	0.002	0.017	0.019	0.053
Total irrigation acres	0.022	0.017	0.022	0.053	0.015	0.076	0.058
Total agriculture acres	0.131	0.205	0.084	0.133	0.123	0.117	0.069
Total crop sales	0.213	0.169	0.022	0.045	0	0.086	0.091
Annual Cattle Values	0.033	0.023	0.003	0.139	0.11	0.005	0.035
Annual Alfalfa Hay Values	0.055	0.103	0.057	0.01	0.016	0.068	0.059
Annual Lettuce Values	0	0.026	0.203	0	0.046	0.013	0.029
Annual Melons Values	0.066	0.02	0.12	0.068	\	\	\
Annual Other Vegetable Values	0.069	0.174	0	0.004	\	\	\

Note: The annual melon values and annual other vegetable values are excluded from the crop production analysis due to the convergence issue.

Table 3.5. County weights in the synthetic Imperial County

	Skilled Labor Employment (logged)	Unskilled Labor Employment (logged)	Skilled Labor Monthly Earnings (logged)	Unskilled Labor Monthly Earnings (logged)	Harvested Acreage (logged)	Crop Values	Yield Per Acre
Riverside	0.209	0.123	0.46	0	0.015	0	0
Fresno	0.209	0	0	0	0	0	0
Inyo	0	0	0	0	0	0	0
Kern	0	0.035	0.328	0	0	0.184	0.04
Kings	0	0.05	0	0	0	0.357	0.264
Merced	0.467	0.722	0	0	0	0.348	0.385
Mono	0.103	0	0	0.146	0.128	0.111	0
Monterey	0.012	0.062	0.212	0.527	0.123	0	0.012
San Benito	0	0	0	0	0.021	0	0
Santa Cruz	0	0	0	0	0.109	0	0
Stanislaus	0	0	0	0	0	0	0
Tulare	0	0.007	0	0.327	0.605	0	0.299

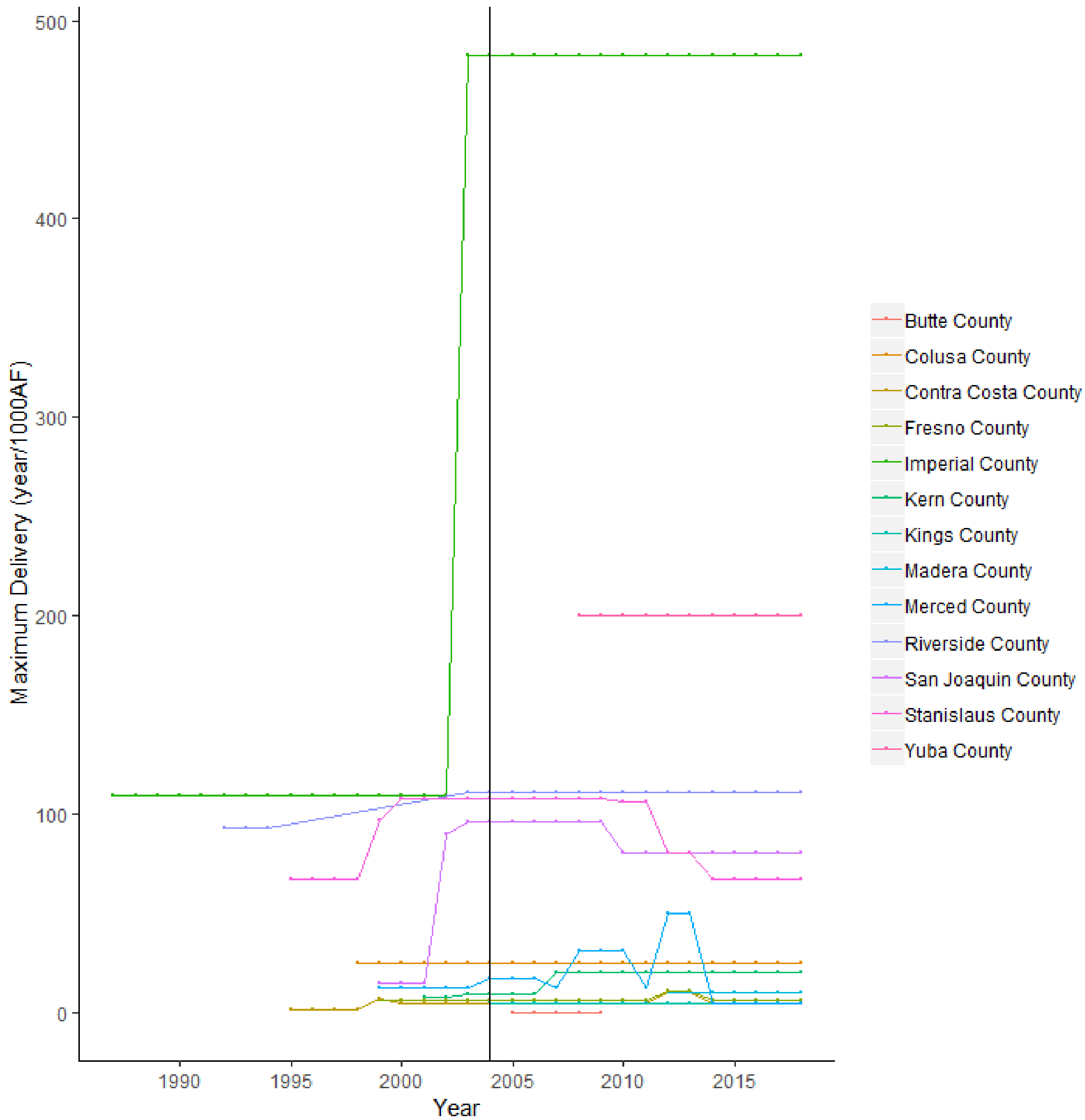


Figure 3.1. Long term water transfer in California since 1970
Note: The water transfer data is aggregated at county-level. In most of case, there are more than one transfer happening in the county within the same data year. County-level data is based on authors' calculation. The vertical line represents the QSA effective year, 2004. Data source: the data is summarized from summary report written by Hanak and Stryjewski (2012).

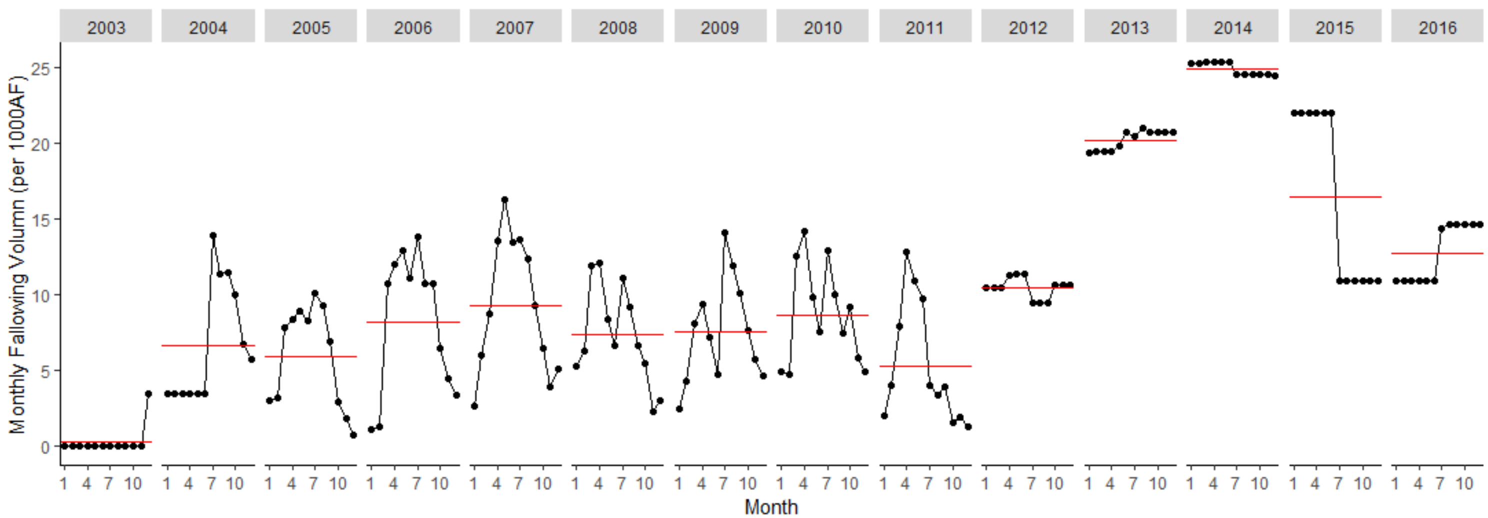


Figure 3.2. Monthly visualization of fallowing volume in Imperial County
Note: Red line represents the annual average fallowing volume. Data source: the data is summarized from summary report written by Hanak and Stryjewski (2012). https://www.ppic.org/publication/californias-water-market-by-the-numbers-update-2012/

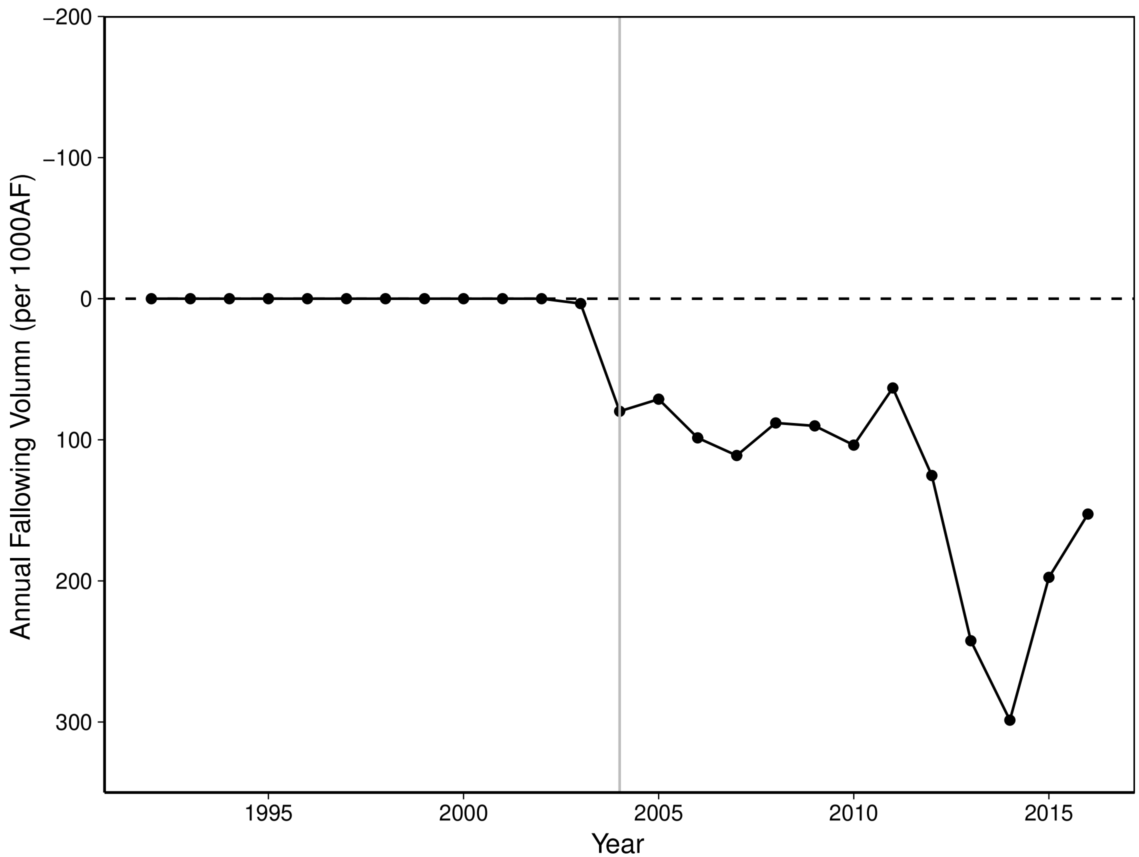


Figure 3.3. Annual visualization of fallowing volume(mirrored with x-axis)
Note: Imperial County only fallowed 3445AF water in 2003 and the QSA program officially went into effect in the year of 2004. Data source: the data is summarized from summary report written by Hanak and Stryjewski (2012). https://www.ppic.org/publication/californias-water-market-by-the-numbers-update-2012/

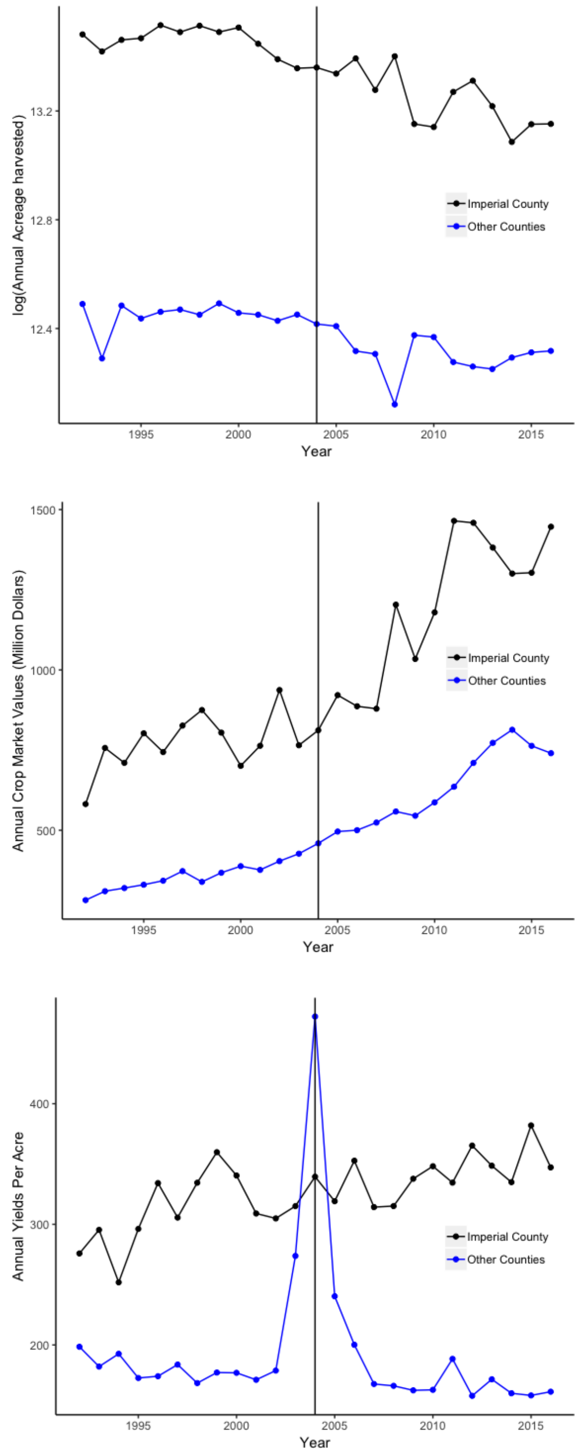


Figure 3.4. Comparison of crop production statistics
Note: All the other counties in the California here refers to all the other 56 counties except Imperial and Alpine County in California. The vertical line represents the QSA effective year, 2004.

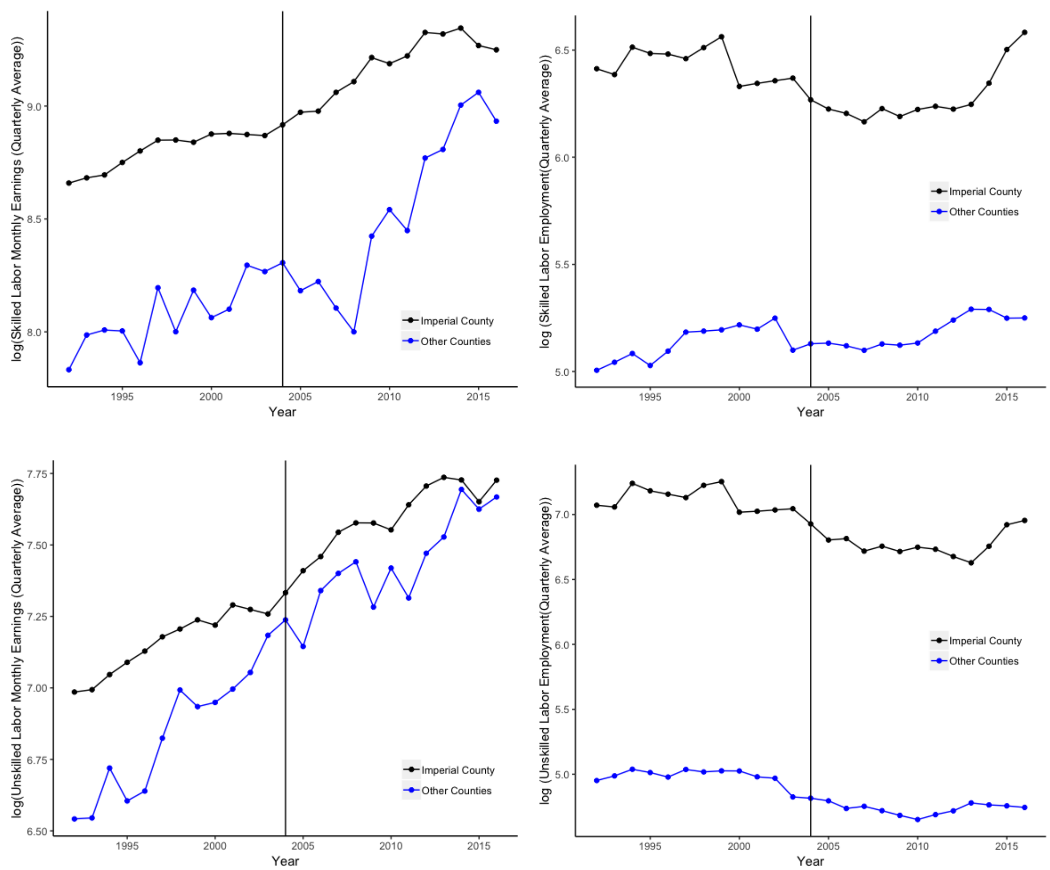


Figure 3.5. Comparison of labor statistics in crop production sector
Note: All the employment and earnings data are calculated by the quarterly average. Earnings are in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

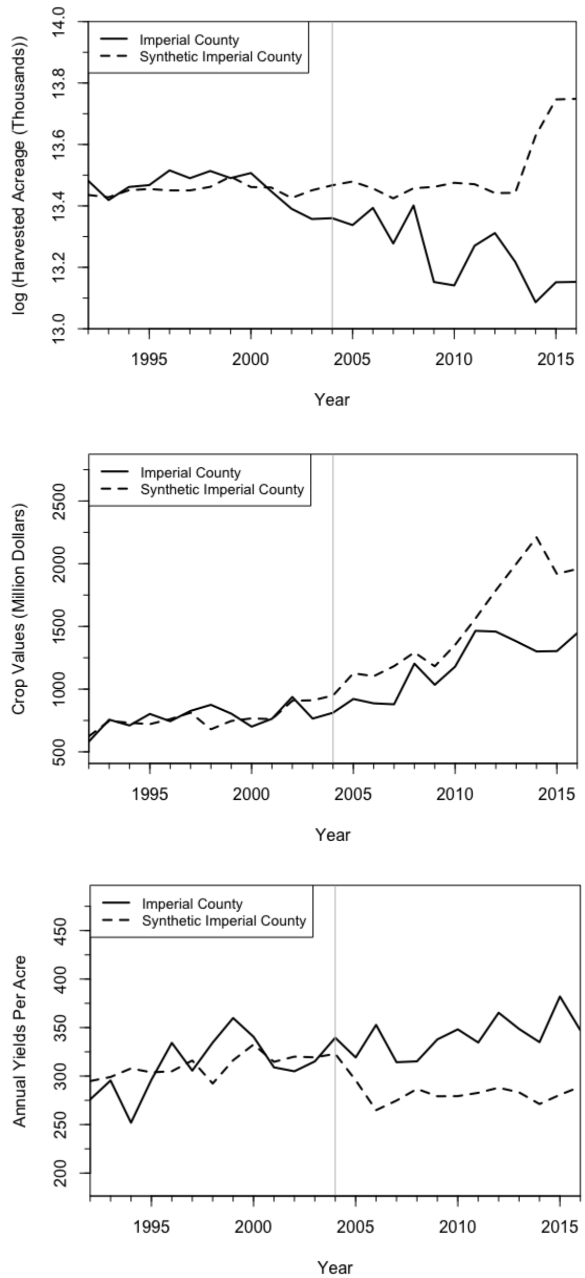


Figure 3.6. Time path in crop production statistics
Note: The crop value is in millions of dollars. The vertical line represents the QSA effective year, 2004.

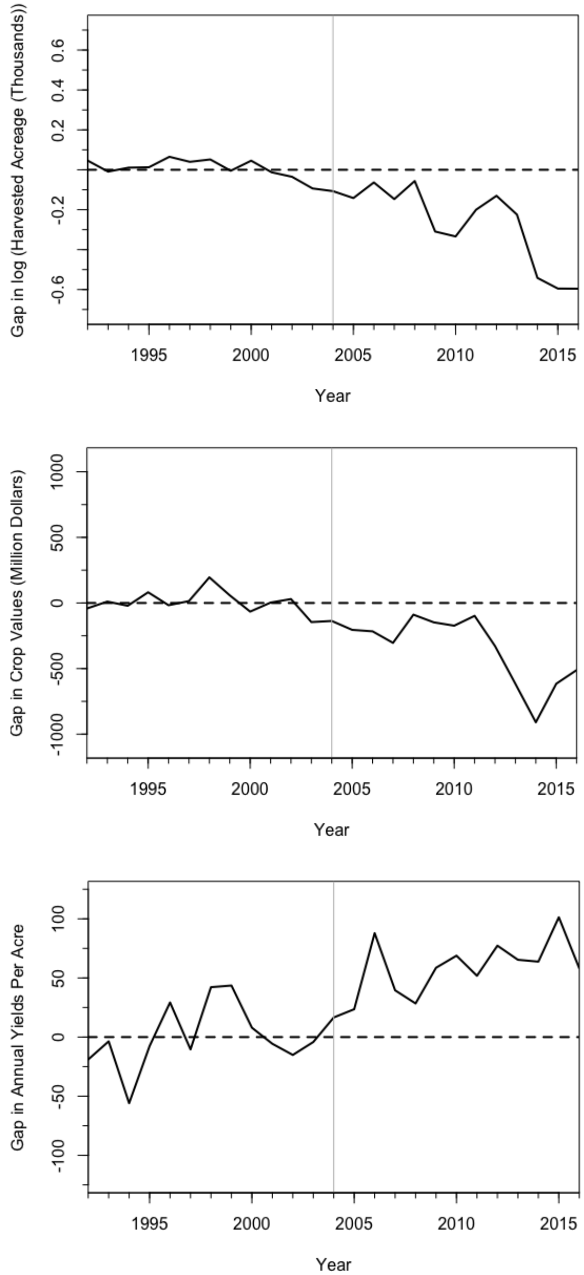


Figure 3.7. Gap plots in crop production statistics
Note: The crop value is in millions of dollars. The vertical line represents the QSA effective year, 2004.

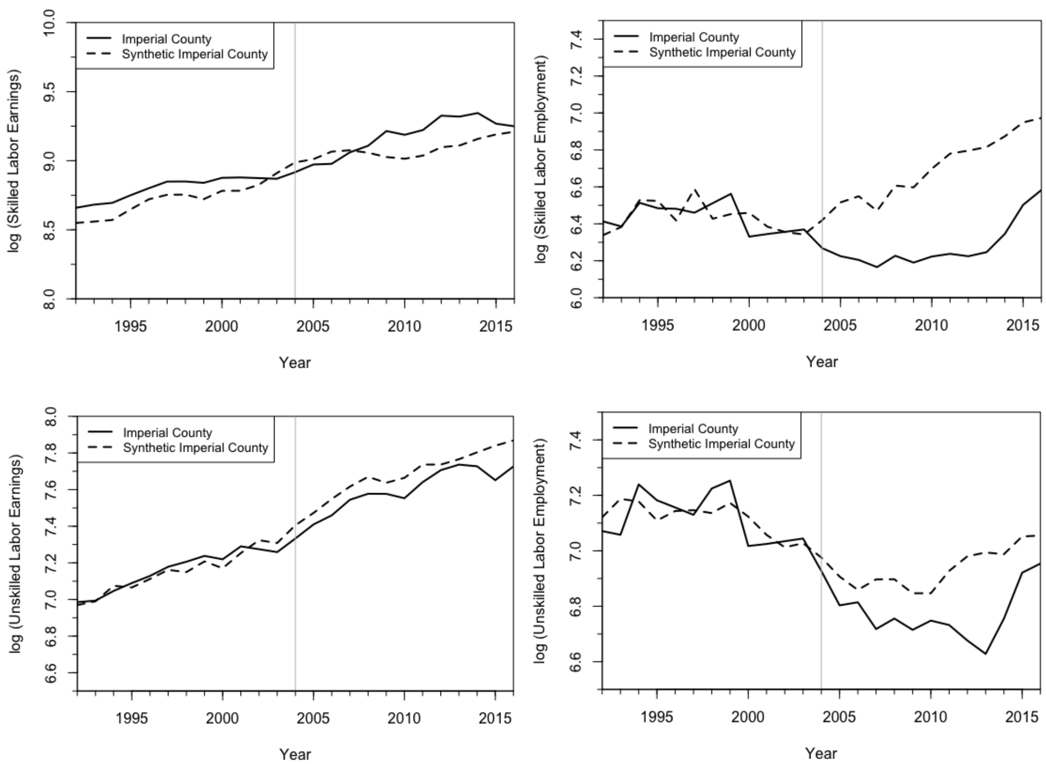


Figure 3.8. Time path in labor statistics in crop production sector
Note: All the employment and earnings data are calculated by the quarterly average. Earnings is in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

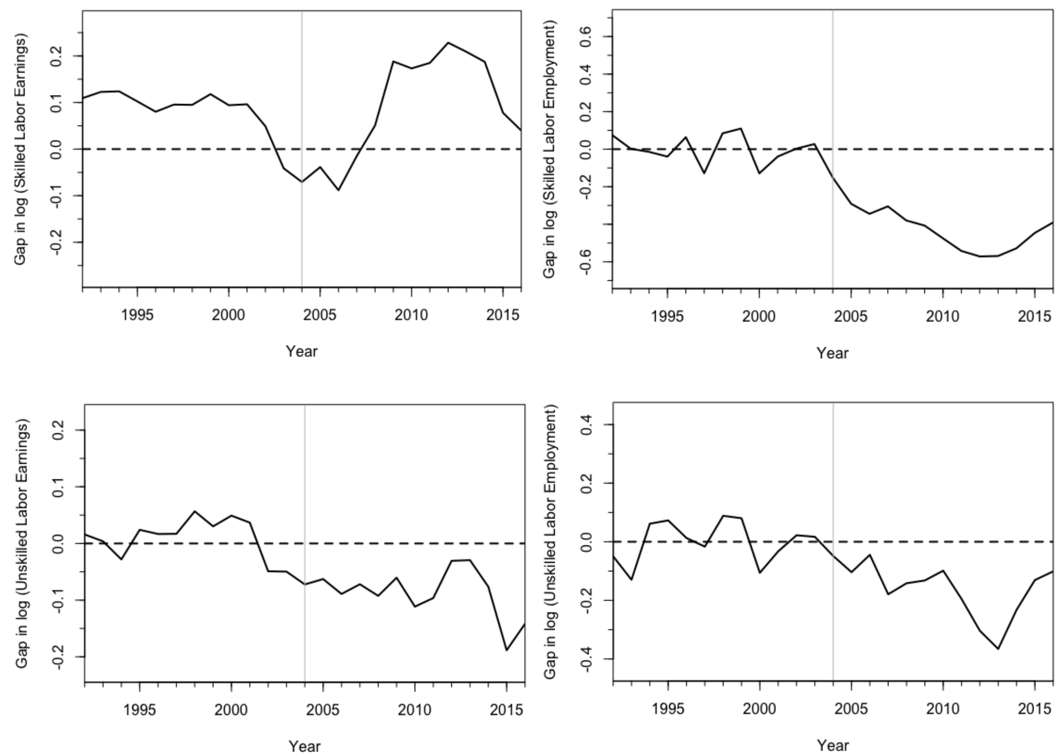


Figure 3.9. Gap plots in labor statistics in crop production sector
Note: All the employment and earnings data are calculated by the quarterly average. Earnings is in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

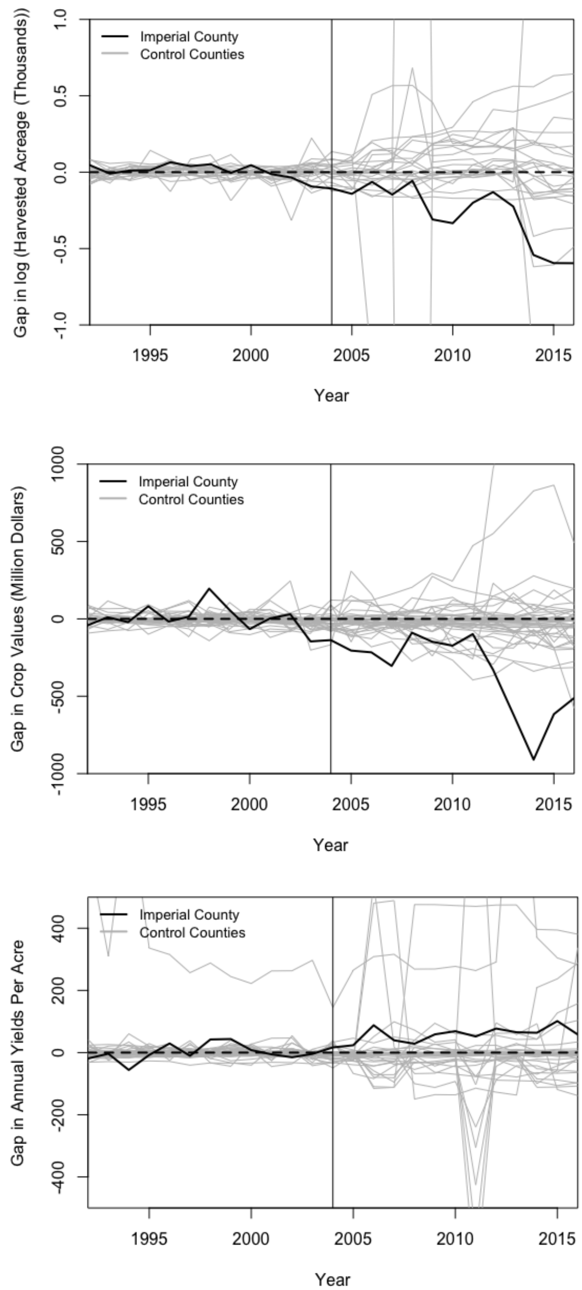


Figure 3.10. Restricted sample size placebo test plots in crop production
Note: The county with pre-QSA MSPE greater than 2 times Imperial County pre-QSA MSPE are excluded from the sample. The crop value is in millions of dollars. The vertical line represents the QSA effective year, 2004.

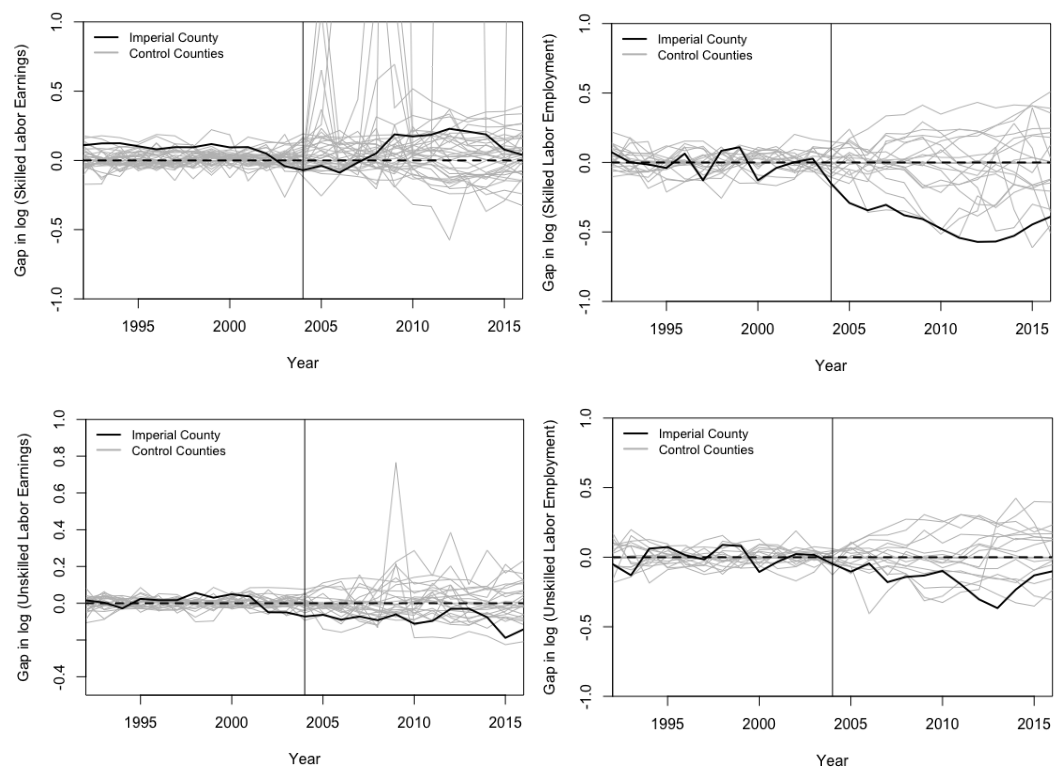


Figure 3.11. Restricted sample size placebo test plots in labor statistics in crop production sector

Note: The county with pre-QSA MSPE greater than 2 times Imperial County pre-QSA MSPE are excluded from the sample. All the employment and earnings data are calculated by the quarterly average. Earnings is in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

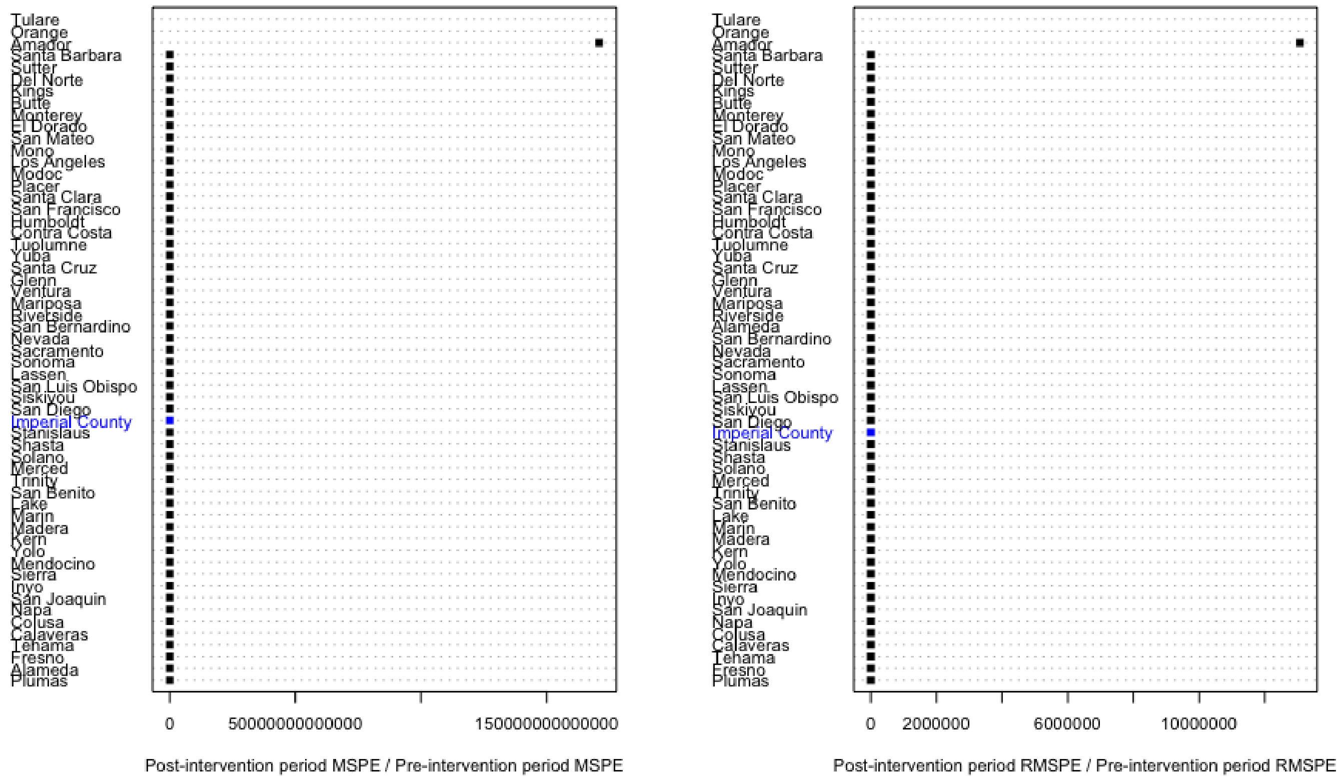


Figure 3.12. Ratio of MSPE and RMSPE for log(Skilled Labor Earnings)
Note: Ratio of post-QSA MSPE and pre-QSA MSPE (Left) and ratio of post-QSA RMSPE and pre-QSA RMSPE (Right) with log (Skilled Labor Earnings). The earning is calculated in quarterly average. Orange and Tulare County are skipped due to convergence issue.

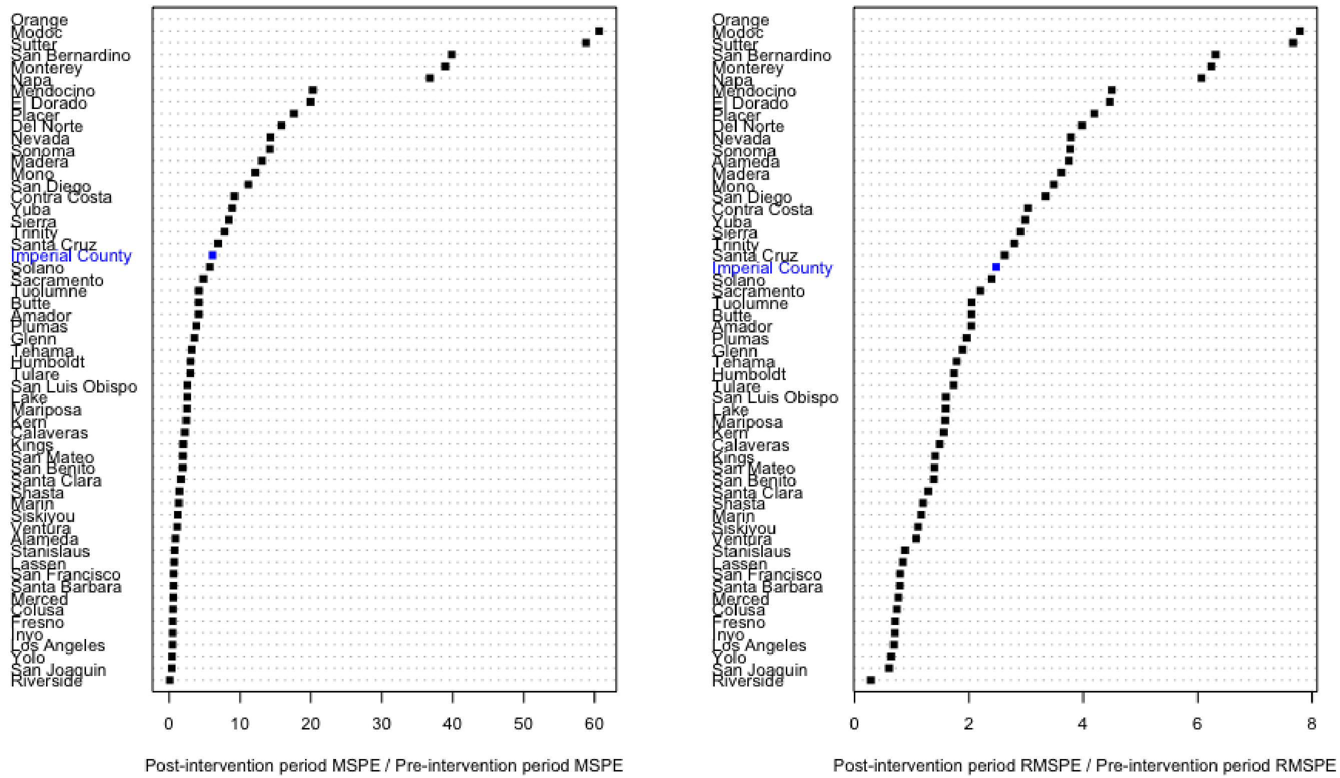


Figure 3.13. Ratio of MSPE and RMSPE for log(Unskilled Labor Earnings)
Note: Ratio of post-QSA MSPE and pre-QSA MSPE (Left) and ratio of post-QSA RMSPE and pre-QSA RMSPE (Right) with log (Unskilled Labor Earnings). The earning is calculated in quarterly average. Orange County is skipped due to convergence issue.

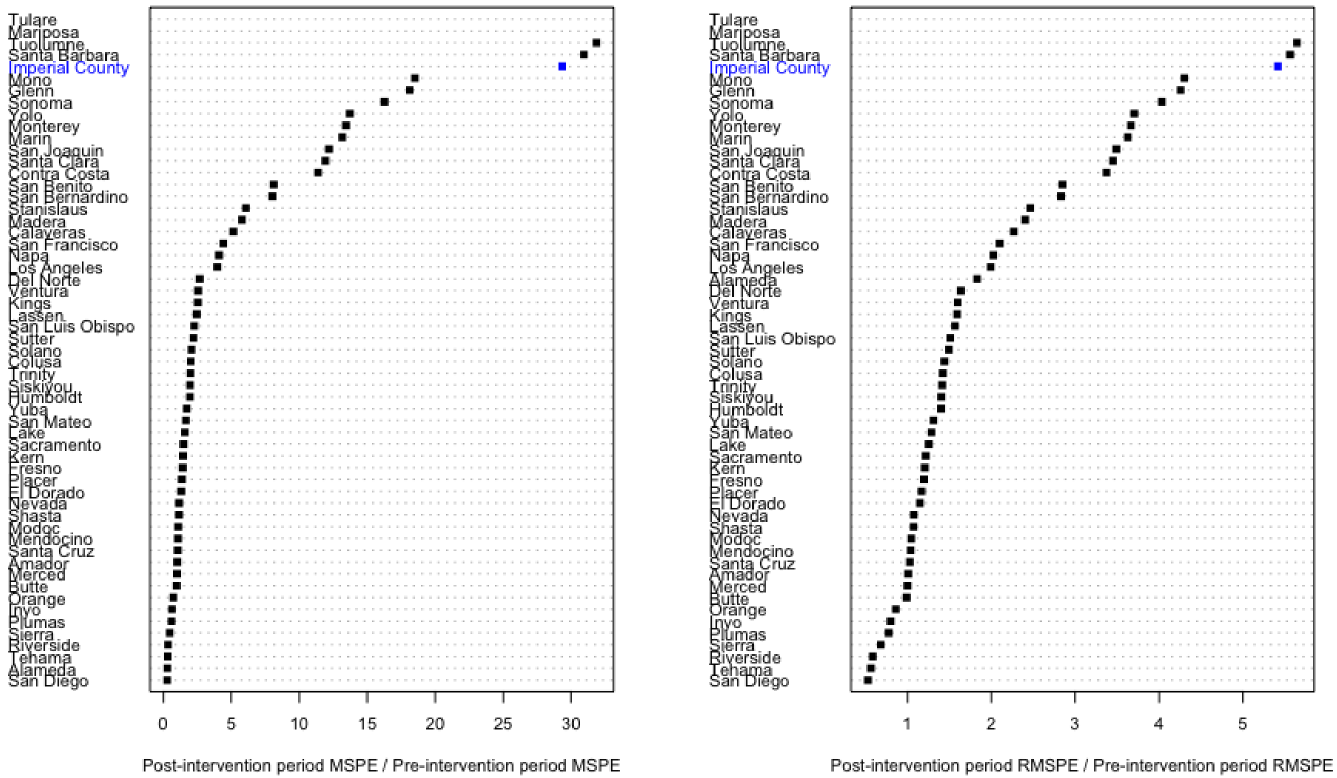


Figure 3.14. Ratio of MSPE and RMSPE for log(Skilled Labor Employment)
Note: Ratio of post-QSA MSPE and pre-QSA MSPE (Left) and ratio of post-QSA RMSPE and pre-QSA RMSPE (Right) with log (Skilled Labor Employment). The employment is calculated in quarterly average. Tulare and Mariposa County are skipped due to convergence issue.

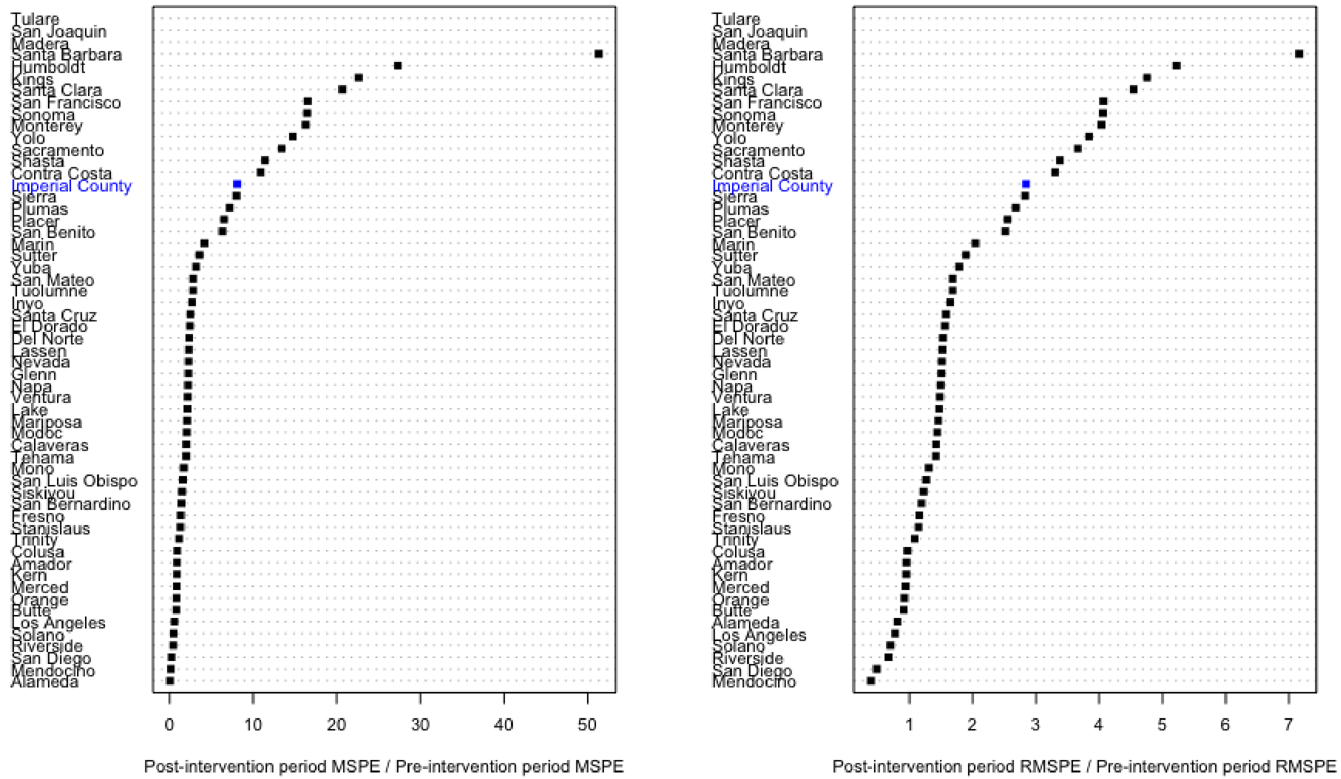


Figure 3.15. Ratio of MSPE and RMSPE for log(Unskilled Labor Employment)
Note: Ratio of post-QSA MSPE and pre-QSA MSPE (Left) and ratio of post-QSA RMSPE and pre-QSA RMSPE (Right) with log (Unskilled Labor Employment). The employment is calculated in quarterly average. Tulare, San Joaquin, Madera County are skipped due to convergence issue.

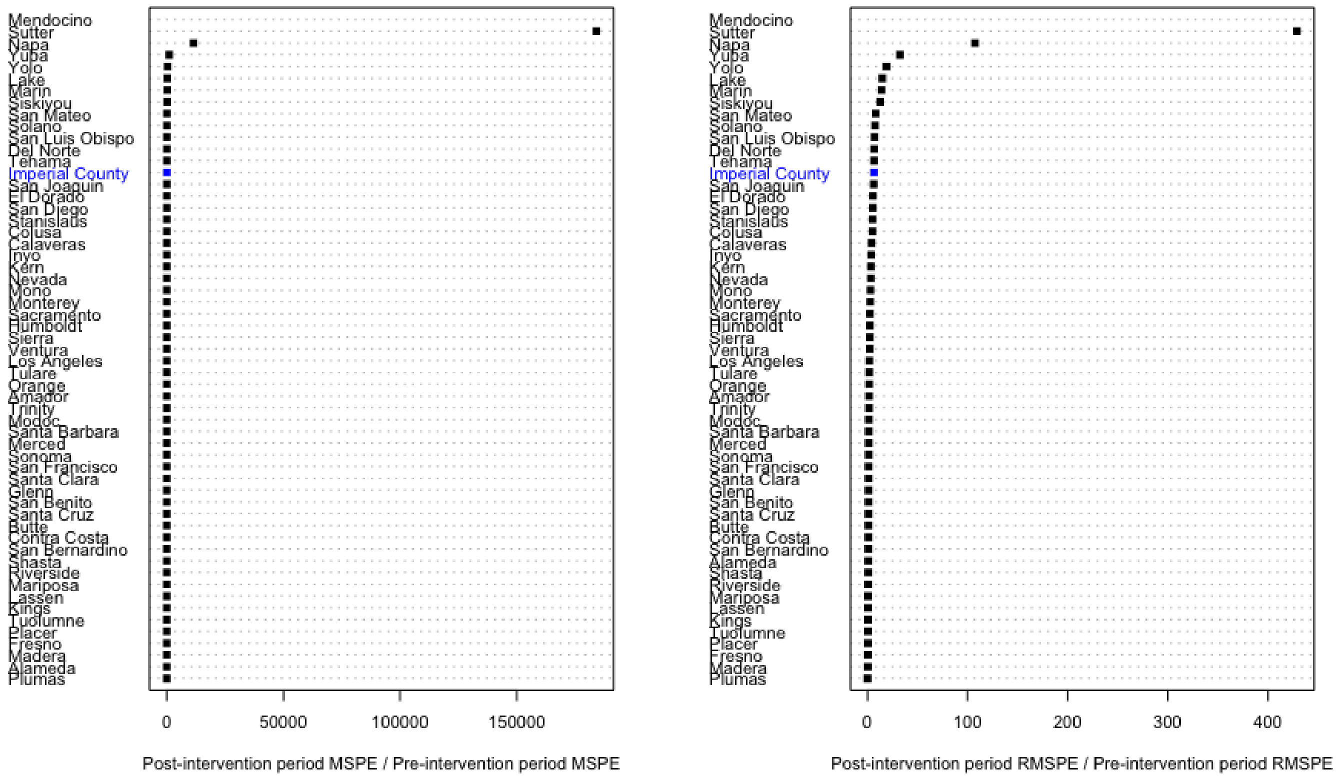


Figure 3.16. Ratio of MSPE and RMSPE for log(Harvested Acreage)
Note: Ratio of post-QSA MSPE and pre-QSA MSPE (Left) and ratio of post-QSA RMSPE and pre-QSA RMSPE (Right) with log (Harvested Acreage). Mendocino County is skipped due to convergence issue.

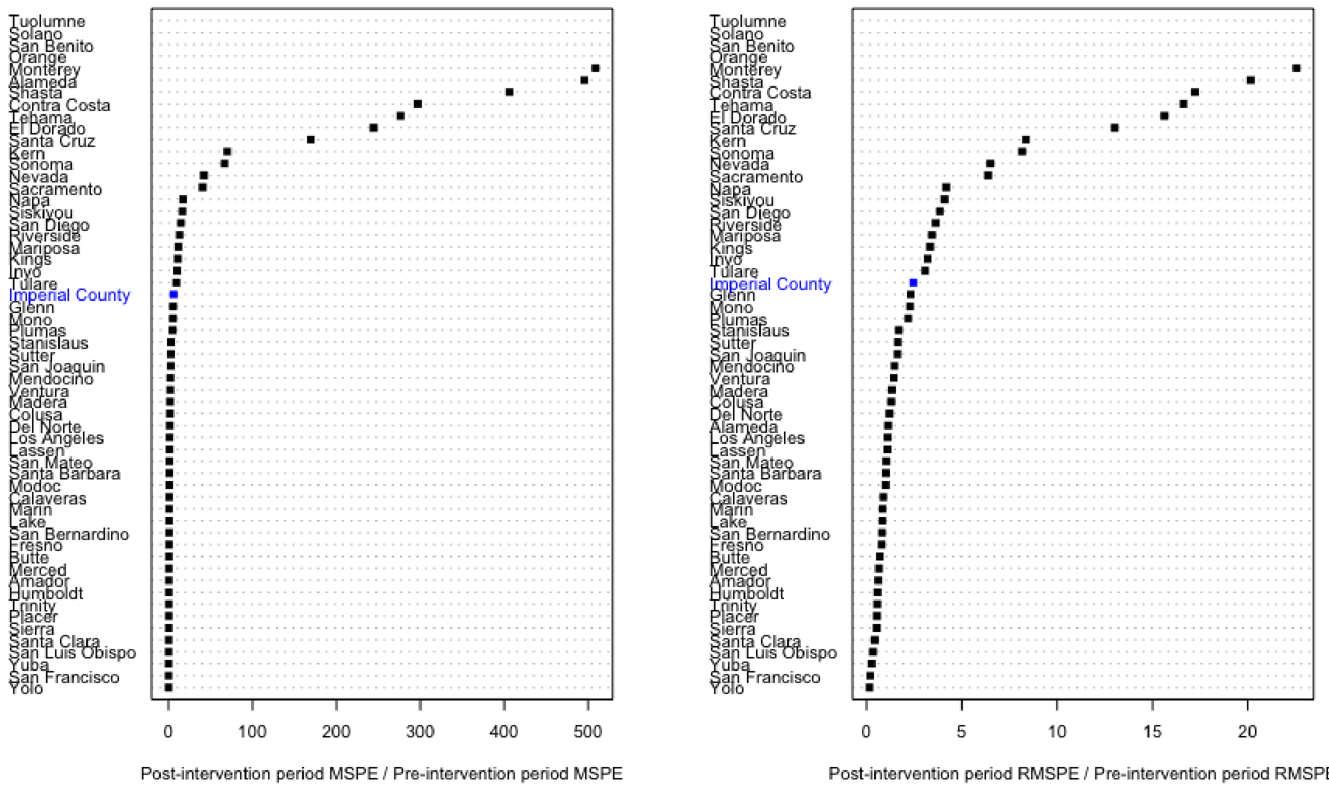


Figure 3.17. Ratio of MSPE and RMSPE for Yield (Per Acre)

Note: Ratio of post-QSA MSPE and pre-QSA MSPE (Left) and ratio of post-QSA RMSPE and pre-QSA RMSPE (Right) with Yield (Per Acre). Tuolumne, Solano, San Benito, and Orange County are skipped due to convergence issue.

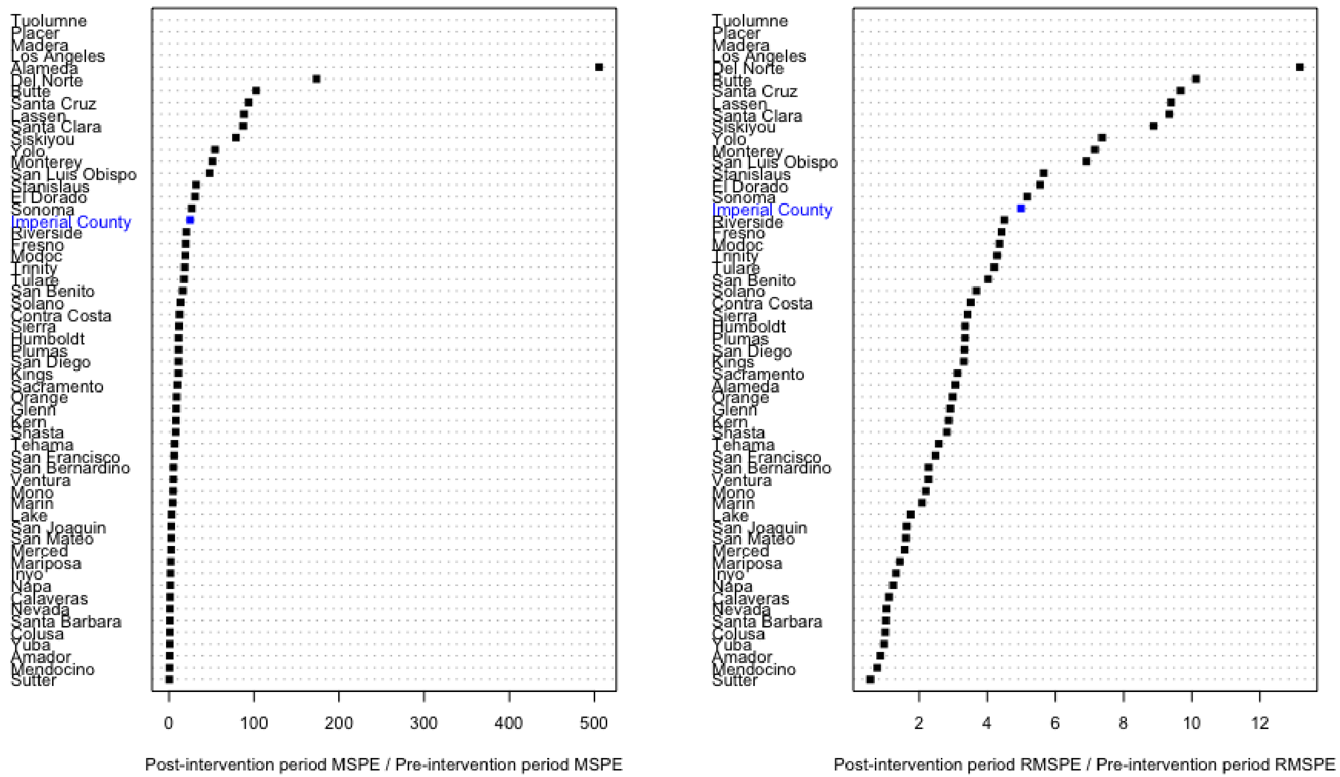


Figure 3.18. Ratio of MSPE and RMSPE for Annual Crop Values
Note: Ratio of post-QSA MSPE and pre-QSA MSPE (Left) and ratio of post-QSA RMSPE and pre-QSA RMSPE (Right) with Annual Crop Values. The crop value is in millions of dollars. Tulare, Placer, Madera, and Los Angeles County are skipped due to convergence issue.

CHAPTER 4

ORGANIC FARMING IN SHALE STATES: A COLORADO CASE STUDY

4.1 Abstract

Air, soil, and water pollution caused by every stage of developing and operating a fracking well could create substantial difficulties in obtaining and maintaining organic certification. More importantly, consumers' awareness of potential soil and water contaminations from fracking will make the organic products less marketable and further harm the organic farms around drilling pads. This paper explores the causal impacts of shale development on organic farming. A novel geospatial dataset of organic farm location and certification duration was created and used to find organic farms present variant distribution patterns across different shale states. Colorado, with the highest exposure possibility, suffers the most from the risk from fracking, and there is a small but significant negative impact on maintaining organic certification.

4.2 Introduction

Hydraulic fracturing, more commonly known as fracking, has caused a dramatic boom in the United States gas production since 2005, when the process was made exempt from important pro- visions, such as the Clean Water Act and other environmental and public health protections. Together with the production boom, rising concern about the negative societal impact of fracking has been documented. For example, the potential health risk and environmental damage caused by fracking has been widely studied in the Environmental Science field since 2010 (Allred et al. 2015; Vidic et al. 2013; McKenzie et al. 2012). More recently, hedonic rent frameworks have been used to explore the overall impact of gas development on local housing value declines (Rakitan 2018; Muehlenbachs, Spiller, and Timmins 2015). However, the overall effect of shale development on agriculture is uncertain

and under-studied.

The shale states, which are rich in gas development, are also agriculturally rich (Hitaj et al. 2014). Eight states, including Texas, North Dakota, Colorado, Wyoming, Louisiana, Pennsylvania, West Virginia and Ohio, have been defined as shale states.[1] The cash brought in by the oil and gas industry increases farmers' wealth through lease and royalty payments, but fracking hampers agricultural production by competing for water and labor with the local agricultural sector and potentially polluting the surrounding environment (Weber, Brown, and Pender 2013). In shale states, agriculture dominates surface land use, while energy development competes for water and labor (Hitaj, Boslett, and Weber 2017) and weakens the profitability of farms, particularly in rural and dry areas. For example, Hitaj, Boslett, and Weber (2017) documented that hydraulic fracturing used more water than farming, and the greatest displacement occurred in states that allow farmers to forgo irrigating, thereby allowing them to sell water to energy firms. Farah (2017) in Alberta, Canada, found a negative impact of hydraulic fracturing on irrigated crop production during the agriculturally active months, using wells and agricultural productivity data. However, the relationship between hydraulic fracturing on organic farming has not yet been studied.

Organic food sales have been expanding rapidly since 2000 in the United States (Greene et al. 2017). Generally, certified organic food is more expensive than the conventional agricultural products on the market. The high price premium encourages many conventional farmers to convert to organic farm (Sustainable Agriculture Research & Education 2012). The ideal experiment suggests that organic production has similar output, but lower cost compared to conventional production. Using Agricultural Resource Management Survey (ARMS) data in 2006 and 2009, McBride and Greene (2008) and McBride et al. (2012) found that organic soybeans and wheat both have higher returns compared to conventional soybeans and wheat. Even though production costs are relatively higher and annual yields are much lower than the conventional products, the returns of organic production remain

[1] According to the United States Energy Information Administration, seven regions account for 92% of the country's shale oil growth in recent year. Table 4.3 is obtained from website: https://www.fool.com/investing/2017/03/25/which-us-states-produce-the-most-shale-oil.aspx, and the original data is from United States Energy Information Administration.

high because of the high price premiums. Organic farmers have incentives to maintain organic certification as a result of high price premiums.

However, compared to conventional productions, organic productions seem to have higher risk. Hanson et al. (2004) mentioned two main risks that vary between organic and conventional production in their six focus group discussions with organic farmers. The first risk unique to organic farmers is that the opportunity cost of organic premium is too high. Organic farmers enjoy the price premiums only after they acquired the organic certification. The transition period is usually three years, and the certification itself could cost them hundreds or even thousands of dollars. Additionally, the organic certification is not guaranteed after the transition period. The second risk is that in order to follow the federal organic standards, the organic production system land and products are more vulnerable to contamination.

If a negative relationship between agricultural production and shale development exists, organic farms should be more sensitive than conventional farms since they will not only be affected by air, soil, and water pollution, but by every stage of developing and operating a fracking well (Royte 2012; Debatin 2014). These farms will also be heavily harmed by consumers' awareness of potential soil and water contamination from fracking. Ong (2014) states that even a single drill pad on the horizon of an organic farm substantially increases the possibility of the erosion of the marketability of certified organic products. Potential organic adopters may take the appearance of wells as a signal of the difficulty in obtaining certification, and current organic operators may decide against maintaining certification in the face of perceived contamination.

In this study, the survival length of maintaining organic certification in Colorado is examined using two novel survival models: Instrumental Variable Estimation in a Survival Analysis and a Joint Model with an endogenous time-dependent variable. Organic farm locations in shale states are geocoded using data from the United States Department of Agriculture (USDA) Organic Integrity Database (OID).[2] In Colorado, organic farms are clustered near gas wells. In other states, such as Pennsylvania, organic farms and gas

[2] https://organic.ams.usda.gov/Integrity/Reports/DataHistory.aspx

wells are separated. There is a lack of correspondence between oil and gas deposits and good agricultural land. Within the same shale state, fracking well placement decisions and regulations may be impacted by the presence of traditional and/or organic agriculture or organic agriculture. The United States Tight Oil and Gas Plays Map is used to show the geologic formations with available shale oil and gas deposits as an instrumental variable on well placement to eliminate the potential bias of county regulations.

The results show that Colorado has the most active fracking wells around organic farms of all the shale states, and that these farms face a higher risk of organic certification loss. There was a small but significant negative impact of fracking on maintaining organic certification in both short- and long-distance-from-well measures. Prior studies have found limited evidence that shale development has negative impacts to agricultural development (see for example, Rakitan 2018; Farah 2017; Hitaj, Boslett, and Weber 2017). However, none of these studies explore the impact of gas development to organic farming using farm-level dataset.

The paper proceeds as follows: Section 2 provides background information on organic farm and energy regulations. Section 3 describes the data set construction while Section 4 provides details on the empirical design and econometric approach. The econometric results are provided in Section 5, and Section 6 concludes this work.

4.3 Background

4.3.1 Impacts of Fracking on Organic Farming

Organic farms aren't allowed to use synthetic chemicals such as fertilizers and pesticides, and the production processes rely heavily on the use of cover crops, crop rotation, manure application, etc. In order to maintain the quality of organic products, organic farmers face additional challenges due to the restrictions and regulations from the National Organic Program (NOP), which defines national standards for the organic production system. These restrictions include stricter requirements on land quality. Qualified organic farms are required to have no prohibited substances applied for a period of three years im-

mediately preceding harvest of the crops. They must also have distinct, defined boundaries and buffer zones, such as runoff diversions to prevent the unintended application of a prohibited substance to the crop or contact with a prohibited substance applied to adjoining land that is not under organic management (U.S. Department of Agriculture 2009). The land requirements regulated by the NOP make location selection of organic farms more sensitive than non-organic agriculture to environmental containments generated by energy development, which can affect the quality of organic production through air and water emitted from the adjacent fracking wells.

By law, the Agricultural Marketing Service (AMS) National Organic Program accredits and oversees organizations as certifiers to enforce the USDA organic regulations (U.S. Department of Agriculture 2019). The certifiers inspect every certified organic farm and collect samples to check for prohibited substances at least annually. The organic certification will be suspended or revoked if the organic farms fail to comply with the Organic Production and Handling Requirements. A suspended or revoked operation cannot sell, label, or represent its products as having been organically produced or handled, and this information will be recorded in the OID.

Pollution of the surrounding environment can occur at almost every stage of developing and operating a fracking well. Radioactive material, volatile organic compounds, and petrochemicals from the adjacent wells could containment air, soil, and both ground and surface water used by surrounding organic farms (Royte 2012). Debatin (2014) depicted that fracking wastewater disposal consequently results in a nearby organic farmer selling their operation, due to the fear of losing organic certification. Essentially, organic inspectors or certification agencies have the right to take soil, product, or tissue samples at any time to verify compliance with the NOP. If any prohibited substances, including fracking chemicals, are detected on a certified organic farm, the producer may have to wait at least three years before becoming eligible for recertification (Kuhlman 2014). The concern about fracking has been widely raised among organic farmers in both in New York and Ohio (Gsell 2012; Kuhlman 2014).

4.3.2 Fluid Disclosure Regulation and Trade Secret Protection

During the fracking process, a large number of chemical additives will be injected into the fracking wells to open the fissures in the shale formation by increasing the well pressure. The chemical additives mixed with fracking fluid could potentially soak into nearby aquifers and contaminate both surface and ground water. The harmful additives, if overused, could pollute the surrounding environment and even threaten human health (Adgate, Goldstein, and McKenzie 2014; Finkel and Law 2011). Despite the environmental issues caused by the fracking fluids, the Energy Policy Act passed by Congress in 2005 forbade the EPA from regulating fracking fluids. Since then, the right to regulate fracking fluid has been turned over to state governments. The Department of Energy has advocated for a nationwide voluntary fracking chemical disclosure project, consisting of a listing of the chemical constituents in fracking fluid on FracFocus website. However, the FracFocus only identifies 59 chemical additives most commonly used in fracking processes and there are still a large number of undisclosed additives (Marie 2017). Twenty-three state governments, including Colorado, have required the fracking companies to disclose chemicals used. Due to the existence of trade secret laws, most of the fracking companies do not have to fully report all the chemicals, especially commercial additive products, used in the fracking process. Even if the drilling companies disclose the chemical additives, the information they are providing is in general terms, such as "friction reducer" or "clay stabilizer" (Craven 2013). More importantly, the vague descriptive information causes more difficulties in preventing potential risks resulting from chemical additives. This unknown nature of the process increases the risk to organic farmers that contamination will disrupt their organic certification.

In the state of Colorado, chemicals used in fracking are considered a trade secret; therefore, the drilling companies may indicate that on the chemical disclosure registry form and provide only the chemical family (Murrill and Vann 2012). In order to encourage the innovation of new products and environmentally friendly fluids in the drilling industry, only state regulators and health professionals have the right to ask for chemical disclosure in emergency situations. The existence of trade secret protection provides an obstacle to

implement state-level fluid disclosure regulation and encourages the development of fracking in shale states. This implies that if an organic farmer wants to avoid the potential of losing certification around the drilling location, it would cost more to conduct additional monitoring, as the additive's information is not fully transparent to the public. If the farm fails to pass the annual accrediting process from USDA, organic certification will be suspended or revoked. This is another potential cost in maintaining an organic farm. Additionally, if there is any evidence showing that the organic farms have violated the regulations, or have represented products as organic without certification, penalties can be assessed up to $17,952 per violation.[3]

4.4 Methodology

The methodology used in this study can be divided into three parts, starting with a simple Kaplan-Meier estimate of the survival curve of the time an organic farm commences operation until it fails to be certified. In the second section, an instrumental variable (IV) approach in a survival analysis context is used to evaluate the causal effect of fracking on organic certification using time-to-event data. In the survival model with and without IV, the exposure to fracking wells is treated as a time-fixed variable, which is not allowed to change during the observation time. In fact, the real data indicates that the active fracking well number changes over time. Lastly, in order to estimate the impact, the exposure to fracking wells is treated as a time-dependent variable using the Joint Model to more accurately explore the impact of fracking on maintaining organic farm certification.

4.4.1 Kaplan-Meier Estimator

Organic farms in Colorado were divided into two groups in the simple Kaplan-Meier estimator. One group is comprised of the organic farms located on the shale plays; the other group is those organic farms not located on the shale plays. The goal of Kaplan-Meier estimator is to estimate the survival curve of the two groups from the organic farm samples. The survival probability is calculated as the number of organic farms surviving

[3] https://www.ams.usda.gov/services/enforcement/organic

divided by number of organic farms at risk. Kaplan-Meier is a simple way to explore the difference between survival curves and determine whether there is a statistically significant difference between the groups.

The Kaplan-Meier estimator is defined as

$$S\hat{(t)} = \prod_{T_i \leq t} (1 - \frac{1}{Y(T(i))}) \tag{4.1}$$

where T_i is the event time for organic farm i, and $Y(T_i)$ is the number of organic farms at risk before time T_i.

4.4.2 Instrumental Variable Approach in Survival Analysis

An instrumental variable approach is introduced here to solve the endogeneity problem caused by a lack of correspondence between oil and gas deposits and good agricultural land. The presence of agriculture or organic agriculture may affect the well placement decision and regulations. Organic farmers within a county specialized in drilling (fracking county) would value the organic certification differently than farmers in a county specialized in agricultural production (agriculture-based county), and farmers' behavior to the certification risk from surrounding drill pads will reflect that difference. For example, organic farmers in a fracking county might have higher tolerance to fracking than organic farmers in an agriculture-based county. In this case, length of maintaining an organic certification clearly depends on the density of fracking wells around each organic farm, but the density of fracking wells is not exogenously given since they are determined in part by the farmers' attitude of organic certification in different counties. A suitable instrument for the density of fracking wells is a variable that is correlated with density of fracking wells but does not directly affect length of maintaining an organic certification. An obvious candidate is a variable that affects oil and gas deposits, since this also affects density of fracking wells, but is not a direct determinant of the length of maintaining an organic certification. The United States Tight Oil and Gas Plays Map is treated as an instrumental variable to eliminate the bias coming from heterogeneity of counties.

A two-stage regression approach is introduced here. The two-stage approach is analogous to the two stages least squares approach commonly used for IV estimation in a linear model. The fitted value from a first stage regression of the exposure on the IV is entered in place of the exposure in the second stage hazard model to recover the causal effect (see for instance Lergenmuller (2017), who applies this approach using the Norwegian Mother and Child Cohort Study data to explore the effect of mothers' body mass index on the pregnancy duration).

The exposure to fracking affects the response of certificated status, and the instrumental variable shale plays affects the response of certified status only through the exposure to fracking. In addition, unobservable and observable factors affect both exposure and response, but not the instrument. In order to meet the requirements of IV, "shale plays" need to be:

1) Unconditionally independent of the unobservable factors. This condition is not testable.

2) Dependent with the endogenous explanatory variable of interest "exposure". We can test this assumption by performing the likelihood ratio test, and by looking at the adjusted R-squared of a linear regression between the exposure and the "shale plays". For this paper, the assumption intuitively makes sense as the density of fracking wells partly depends on the shale plays.

3) Conditional on observable, unobservable and "exposure" variables, "shale plays" has to be independent of the response. This assumption is not testable. However, "shale plays" was assigned before the energy development boom in early 2000s. One farm will not lose its organic certification just because it is located on the shale plays. It will have possibility to lose its organic certification if, and only if, it is polluted, or the owner voluntarily discontinues organic agriculture.

1. First stage

d_i^j is predicted by using linear regression model:

$$E[d_i^j|plays, aquifer] = \beta_0 + \beta_1 plays + \beta_x^{'} aquifer \tag{4.2}$$

Where d_i^j is an outcome variable of interest, equal to the number of active wells located within $i-mile$ buffer zone in state j. $plays$ is an instrumental variable, equal to one if the organic farm is located on the shale plays. $aquifer$ is to control for the heterogeneity of aquifers. In order to capture the impact of active fracking wells properly, $1-mile$, $3-mile$, $5-mile$, $10-mile$ were selected as bandwidths in Colorado. The predicted value of d_i^j obtained from this model is $\hat{d_i^j}$.

2. Second stage

Before performing the second stage, Aalen's Additive Hazard Model is fitted with all the observed variables. The model without IV is displayed as:

$$h[t|d_i^j, plays, aquifer] = h_0(t) + \gamma_d(t)d_i^j + \gamma_x^{'}(t)aquifer \tag{4.3}$$

where $\gamma_x^{'}(t)$ is selected corresponding to the selection of the covariate, $aquifer$. The two variables of interest are the estimated cumulative baseline hazard $\hat{H_0}(t)$ and the estimated cumulative parameter for d_i^j, $\hat{\gamma_d}(t)$. A non-parametric strategy in the survival analysis is used here. The non-parametric approach focuses on estimation of the regression coefficients γ leaving the baseline hazard $h_0(t)$ completely unrestricted (Cox 1975).

The second stage model with fitted values obtained from first stage is:

$$h[t|plays, aquifer] = h_0^s(t) + \gamma_d(t)\hat{d_i^j} + \gamma_x^{'}(t)aquifer \tag{4.4}$$

where $\gamma_x^{'}(t)$ and aquifer are the same as illustrated above, and where $\hat{d_i^j}$ is the predicted value of d_i^j using first stage model. The variables of interest remain the same as the Aalen Additive Model without IV as above. They are the estimated cumulative baseline hazard $\hat{H_0}(t)$ and the estimated cumulative parameter for d_i^j, $\hat{\gamma_d}(t)$. $h_0^s(t)$ is a baseline hazard function with two-stage approach.

3. Validity of IV

In order to check the validity of assumption 2), the correlation between d_i^j and $plays$ was determined. Table 4.1 shows the correlation between the instrument and the exposure,

d_i^j. The instrument was found to be relatively weak within narrow bandwidths selection such as $1-mile$ and $3-mile$, with correlations ranging from 0.195 to 0.217. The correlations of $10-mile$ bandwidth selection are comparatively high for all four bandwidths.

Table 4.2 shows that the instrument, *plays*, is statistically significant within all the selected bandwidth choices. However, the $adjusted-R^2$ of the smaller bandwidth choice is relatively small. For example, the reported $adjusted-R^2$ of $5-mile$ bandwidth choice is about 0.053, which means that the model only explains about 5% of the variation in the response, and so the instrument will be rather weak. Table C.1 inthe Appendix show the likelihood ratio test results. The instrument, *plays*, shows statistical significance at 99% level with all the bandwidth choices in Colorado.

4.4.3 Joint Model with Time-dependent Covariates

In the survival model with and without IV, active well impact has been treated as a time-fixed variable, meaning it is not allowed to change over time. However, the real number of active fracking wells does change over time. Figure 4.6 depicts how the selected characteristics of active fracking wells change within a $20-mile$ buffer of sixteen sampled organic farms. Four different variables were selected as indicators to evaluate the impact of active fracking wells. The *well number* is calculated by summing all active fracking wells within the defined time frame (0, 2, 4, 6, 8, 10 years from the certification of the nearest organic farm). The *production days* variable is the sum of total production days of the active wells around the nearest organic farm. Our observation here is organic farms in Colorado. If there are two farms near one well, that well can only be counted to the nearest farm. The *weighted distance* is defined as the reciprocal of the distance of each active well to its nearest organic farm. It is easy to understand that there is a negative correlation between the distance of each well to an organic farm; that is, the larger distance from an organic farm, the less impact this active well will have on that farm. Hence, the sum of the reciprocal of distance ($\sum \frac{1}{distance}$) is used to evaluate the impact of each active fracking well on its nearest organic farm. The last variable to evaluate the impact of active fracking well is called *indicator*, which is the sum of production days over distance ($\sum \frac{production\ day}{distance}$). The

length of production days is positively correlated with the impact of the active fracking well, while the distance is negatively correlated with the active fracking well. This implies that wells that are farther away from the organic farm will have smaller impact than the wells that are close to the farm. If the active fracking wells do have an impact on maintaining organic farm certification, the aggregated indicator may have more of a chance to capture it.

Figure 4.6 shows that all sixteen selected organic farms have different patterns of surrounding wells. Among these organic farms, farm No.23, No.260, No.519, No.531 were originally certified, but lost organic certification at the end of the sampling period. For example, farm No.92 maintains its organic certification in our sampling period. Well number, production days, weighted distance and indicator remain low level in its sample years of data. However, farm No.260, which lost its organic certification in its 10th year, has a significant increasing trend of the surrounding well characteristics within its sample years of data.

Hence, instead of assuming the impact of fracking is unchangeable over time, it is worthwhile to treat the exposure to fracking as a time-dependent variable. Ignoring the time dependency of exposure to fracking when fitting the hazard model might lead to an incorrect estimate of the hazard function (Munoz-Price et al. 2016). There are two main types of models which deal with time-dependent variables; one is the Extended Cox Model, and the other is the Joint Model. The Extended Cox Model is widely used to handle exogenous time-dependent covariates by treating the occurrence of events as the realization of a very slow Poisson Process (Fleming and Harrington 2011; Andersen and Gill 1982). The Joint Model focuses on the survival outcome accounting for the effect of endogenous time-dependent covariates measured with error. Here, a time-dependent covariate is endogenous if its value at any time point can be affected by an event occurring at an earlier time and is exogenous if its value at any point in time is not affected by an earlier event. In this case, fracking wells in Colorado cluster together around the shale plays. The existence of a fracking well at earlier time indicates large amounts of oil and gas

deposits. Therefore, the Joint Model is introduced to deal with the time-dependent impact of exposure to fracking. Well-level production data was collected to represent the time-dependent well impact. The Joint Model is adopted from Rizopoulos (2012). The Joint Model is used to measure the association between the longitudinal impact level and the risk for an event, while accounting for the endogenous property of the impact indicator. There are two components of a standard Joint Model, Relative Risk Model and Mixed Effects Model.

Relative Risk Model:

$$h_i[t|M_i(t), plays_i] = h_0(t)\exp(\gamma^T plays_i + \alpha m_i(t)) \tag{4.5}$$

where $m_i(t)$ is the true and unobserved value of active well characteristics; in this paper, *well number*, *production days*, *weighted distance* and *indicator*. $M_i(t), = m_i(s), 0 < s < t, t > 0$ is the longitudinal history denoting the history of the true unobserved longitudinal process up to time point t. α is the parameter of interest, representing the strength of the association between the active well characteristics and the risk for an event. $plays_i$ is a treatment indicator, equal to 1 if the organic farm is located on the shale plays, 0 otherwise. $h_0(t)$ is the baseline risk function as defined in the instrumental variable model.

The Linear Mixed Effects Model:

$$\begin{aligned} y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= x_i^T\gamma + z_i^T(t)b_i + \varepsilon_i(t) \\ &= \beta_0 + \beta_1 t + \beta_2 t \times plays_i + b_{i0} + b_{i1}t + \varepsilon_i(t) \end{aligned} \tag{4.6}$$

where $x_i(t)$ and β are fixed effects, while $z_i(t)$ and b_i are random effects. In this model, our parameter of interest is α, measuring the association between $m_i(t)$ (i.e. the four different well characteristics, *well number*, *production days*, *weighted distance* and *indicators*) and the risk for the organic certification loss.

4.5 Data

The dataset is drawn from several sources. The records of organic farms with the organic certification effective and suspended or surrendered dates come from USDA Organic Integrity Database (OID).[4] OID maintains the physical addresses of organic farms all around the world, which provides the opportunity to geocode the location of each organic farm with longitude and latitude information using google geocoding API service. One of the important variables in the standard survival analysis is the survival duration. Farm-level certification duration is calculated starting from the effective date of the certification and ending with the reported date of suspension or surrender recorded in the OID system.

Table 4.3 lists the shale regions and states. According to the United States Energy Information Administration, seven regions account for 92% of the country's shale oil growth in recent years. Hence, the eight states that house those regions are defined as the shale states. Location relationship maps between fracking wells and organic farms of selected states are presented in Figures 4.1-4.3. States with over 100 organic farms at any point in time up until the end of 2008, namely Pennsylvania, Ohio and Colorado, are depicted. The location maps of other shale states are presented in Figure C.1 of the Appendix C. The "active wells" here refers to the wells which are currently producing when surrounding organic farms are certified. In Colorado, uncertified organic farms and active wells are clustered together around the shale plays. Compared to a clustered pattern in Colorado, uncertified organic farms in Ohio are more randomly located within the state boundary. In Pennsylvania, most of the shale plays are located on the northwest corner, while organic farms are clustered in the southeast. The clustered pattern of fracking wells and organic farms leads to the selection of Colorado as the state to explore the effect of fracking on maintaining organic certification.

Geospatial well data is obtained from various sources. The data sources of each state are listed on Table 4.3. Active wells are defined as wells located within target buffer zones that have been drilled within the nearest organic farm survival time period. Four distinct buffer

[4]The United States Department of Agriculture, Agricultural Marketing Service, Organic Integrity Database, website: https://organic.ams.usda.gov/Integrity/Reports/DataHistory.aspx, Accessed: 04/13/2018

zones are adapted from Auch (2015). They are: core ($< 1 mile$), intermediate ($< 3 miles$), periphery ($< 5 miles$), and sub-watershed ($< 10 miles$). Figure 4.4 shows the number of active wells around organic farms in the three selected states. The number of wells with different bandwidths supports the Figures 4.1 - 4.3. Visually Colorado has the most fracking wells clustered around the organic farms; therefore, it has the highest number of active wells around organic farms. Comparing all three panels together, it is evident that only a small number of active wells fall within the sub-watershed zone in Ohio and Pennsylvania. In order to capture the impact in the proper size, Colorado is selected as the target state to study the impact of active fracking well on holding an organic farm certification.

The summary statistics of total active wells for Colorado are listed in Tables 4.4 and 4.5. The average certification duration in years of uncertified organic farms in Colorado is lower than the state average. This implies that surrounding an organic farm by more active fracking wells might cause a negative impact to overall organic farming in Colorado.

The well production status data for the Joint Model analysis is obtained from the Colorado Oil and Gas Conservation Commission,[5] who records the well-level production data between 1999 and 2018. We combined the production data with well geospatial location to evaluate the production days and distance of each active well to the nearest organic farm in Colorado.

The well location is not exogenous because well drilling selection is related to local economic and regulatory conditions. In order to omit the well location bias, "shale plays" is treated as an instrumental variable in the empirical framework. The instrumental variable is obtained from the United States Energy Information Administration (EIA).[6] "Shale plays" tracks the potential of oil and gas production on a shale basin, which is unrelated to any agricultural decision and is decided before energy development begins. However, it is highly related to the drilling location of a fracking well.

The main covariate added to the survival model is whether or not an area lies over an aquifer. The United States aquifer digital map is obtained from the Principal Aquifers of

[5] https://cogcc.state.co.us/data2.html#/downloads
[6] https://www.eia.gov/maps/maps.htm

United States dataset maintained by the United States Geological Survey (USGS).[7]. The aquifer layer was modified from the Ground Water Atlas of the United States maps with the resolution $1 : 2,500,000$. The aquifer type of each organic farm is collected by overlapping each organic farm address and the aquifer digital map. There are only two types of aquifers in Colorado, sandstone aquifers and unconsolidated sand and gravel aquifers.

The number of organic farms in each county for three separate census years, 2002, 2007 and 2012, was collected from the USDA Census of Agriculture. Figure 4.5 illustrates the organic farm trends from 2002 to 2012. Compared to Ohio and Pennsylvania, Colorado has the largest decrease in organic farm numbers since 2007.

4.6 Results

4.6.1 Kaplan-Meier Plots

Figure 4.7 depicts the probability of survival for a group of organic farms located on a shale play versus the group of organic farms not located on the shale plays. The group of organic farms not located on the shale plays has a slightly higher survival rate than the group of organic farms located on the "shale plays" after approximately 12 years time. The stable survival rate for organic farms located on the "shale plays" is approximately 50% while the survival rate for organic farms not located on the "shale plays" is approximately 80%. This result implies that the organic farms located on the shale plays will have a higher risk of organic certification loss than the farms located farther away from the shale plays in Colorado. The sudden decline of the red dashed line in the graph in year 22 is driven by one single organic farm. However, the simple Kaplan-Meier plot is not highly accurate because it does not count for the endogeneity problem or time-dependent effects.

4.6.2 Survival Curves Without IV Estimation

The survival plots without IV estimation baseline hazard for organic farms in Colorado

[7] The Principal Aquifers of United States dataset is published in 2003 and contains the shallowest principal aquifers of the conterminous United States. https://water.usgs.gov/ogw/aquifer/map.html

are shown in Figure 4.8, while the estimated cumulative parameters are shown in Figure 4.9. The baseline hazard function is akin to the intercept in a regression model and it must be positive over time. It describes the risk of losing the organic certification at time t in the control group. In this case, the control group is organic farms without active fracking wells in close proximity. There is a significant baseline effect on maintaining an organic certification in Colorado. Figure 4.8 indicates that even without the fracking wells near an organic farm, the risk of organic certification loss increases as time goes on. The baseline hazard rate stays the same when we change the bandwidth from $1 - mile$ to $10 - miles$. This result is valid because without exposure to fracking wells, organic farms in the control group should have identical baseline hazard rate with different bandwidth choices. The next parameter of interest is the estimated cumulative parameter, which is often known as the Nelson-Aalen estimator of cumulative hazard function. The baseline hazard is estimated relatively precisely for organic certification, lasting for a shorter time (the 95% confidence interval is tight around the estimated cumulative hazard) but as the certification duration increases, the confidence interval gradually widens for the cumulative hazard estimate for a certification lasting over 15 years. Due to the existence of the baseline hazard for all organic farms, more organic farms maintain their organic certification for a short time period and fewer organic farms are able to maintain their certification for more than 15 years in Colorado.

The pattern of cumulative parameters in Figure 4.9 is complex. The cumulative parameter indicates the risk of treatment farms. For the first 5 years, the slope is negative, and the upper confidence band lies slightly below zero, except for the plot with the $1 - mile$ bandwidth choice. This indicates a statistically significant negative early impact of fracking on maintaining organic certification with widened bandwidth choice. The time-varying coefficient in the Aalen model is zero and constant from 5 to 15 years, indicating no middle effect. It is nonzero after 15 years within the $1 - mile$ bandwidth choice, indicating a late effect with short distance, such as 1 mile. Overall, there is no significant effect of fracking on maintaining organic certification. Except for the time period after approximately 15

years, an increase in active fracking wells increases the hazard difference between control and treatment farms very slightly. For example, the hazard ratio is 0.03 at a $1-mile$ bandwidth choice. This indicates that one organic farm with exposure to active fracking wells within the $1-mile$ bandwidth will have a 3% greater risk compared to the organic farms without any fracking exposure. When comparing across the four plots from narrow to wide bandwidth, it is interesting to see that cumulative coefficients become smaller as bandwidths increase, indicating that the long-distance exposure might have less risk of maintaining the organic certification than short-distance exposure. The estimated cumulative coefficients have large confidence interval after 15 years in four plots because few organic farms still maintain organic certification after 15 years.

4.6.3 Survival Curves With IV Estimation

Figure 4.10 provides the estimated baseline hazard of active fracking wells within the $10-mile$ bandwidth choice in Colorado using IV estimates. It is evident that the overall trend of the two-stage approach is very close to the survival results without IV. After omitting the heterogeneity of counties by using "shale plays" as an instrumental variable, the baseline hazard remains similar with different bandwidth choices. The magnitude of different bandwidth choices slightly decreases from approximately 0.6 to 0.4 after introducing the IV estimate. However, the estimated cumulative hazard plots below are quite different from the results without IV.

Figure 4.11 depicts the estimated cumulative parameter with IV for Colorado. The cumulative coefficient curves hold the same trend as the results without IV, however, the cumulative hazard with IV shows a larger magnitude of coefficient than the one without IV. For example, within the $1-mile$ bandwidth choice, the cumulative parameter is 0.2 with IV compared to 0.03 without IV. Instead of having the late effect, the active fracking wells show no impact on maintaining organic certification within all four bandwidths with IV. However, the risk of organic certification loss with IV is not statistically significant from zero from all bandwidth choices within the sampling period. This might be due to the impact of active fracking wells being a time-dependent variable but the survival analyses with and

without IV fail to account for this. The results of the Joint Model Analysis accounting for the time-dependent variable are discussed below.

4.6.4 Joint Model Results

Table 4.6 shows the parameter of interest results for four different bandwidth choices of four active well characteristics. Again, α represents the strength of the association between the active well characteristics and the risk for an event. The Joint Model finds a statistically significant association between four selected active well characteristics and the risk for organic certification loss. All four selected characteristics, namely *well number*, *production days*, *weighted distance* and aggregated *indicator*, are positively correlated with the risk of losing organic certification. For example, the cell in column 3 row 1 of the table shows that one unit of decrease in the well number corresponds to a $1 - exp(-\alpha) = 1 - exp(-0.077) = 1 - 92.64\% = 7.56\%$ increase in the risk for losing certification. When comparing "indicator" impact across different bandwidths, the larger the bandwidth, the smaller the magnitude of impact. However, the significance of the parameter increases as the bandwidth increases from $1 - mile$ to $10 - mile$. This implies that exposure to fracking has both a short-distance impact and a long-distance impact on maintaining organic certification in Colorado. The short-distance impact has larger magnitude but less significance while the long-distance impact has smaller magnitude but more significance.

4.7 Conclusion

This paper explores the effect of energy development on organic farming by evaluating the length of maintaining organic certification near fracking wells in Colorado. The different distribution relationships between active fracking wells and organic farms in eight different shale states are illustrated. Relatively more active fracking wells are located near organic farms in Colorado than other shale states. A survival model with IV estimates is used to explore the short-distance and long-distance exposure impacts of fracking on organic certification. The results suggest that exposure to fracking has a small but negative impact on maintaining organic certification in both short-distance and long-distance band-

width choices. The impact of exposure to fracking is treated as a time-dependent variable. Findings suggest that both short-distance and long-distance impact of fracking exist, but the overall impact of active fracking wells on organic farming is not evident especially for a long-distance bandwidth choice.

Drilling activities appear to discourage organic farming in Colorado. Identifying the direct causes of organic certification loss can support policy development that regulates negative externalities on organic farms caused by drilling. While farmers with mineral ownership directly benefit from mineral rights leasing and royalty payments (Weber, Brown, and Pender 2013), most organic farmers without mineral ownership will be damaged as a result of hydraulic fracturing. High organic premiums encourage organic farmers to maintain the organic certification, but the cost of maintaining the certification itself would increase due to the threat of the nearby drilling pads. Intensive fracking activities can be a strong deterrent to many conventional farmers considering a conversion to organic farming. For those conventional farmers who are willing to convert to organic production, a higher opportunity cost may be present with the close proximity to drilling pads. For example, the waiting period of three years might extended during the initial transition period for farms near fracking wells. Federally subsidized crop insurance could be one of the solutions to lowering the opportunity cost during the initial transition of those farmers without mineral ownership.

Due to the restriction of dataset size, confidence intervals are quite wide in this paper. As the fracking boom only started around 2004, the fracking impact time is still relatively short. This problem might be solved in the future by collecting longer time period data and adding more observations into the dataset.

Table 4.1. Correlations and IV strengthen/significance

State	d1	d3	d5	d10
Colorado	0.195	0.217	0.237	0.299

Table 4.2. Summary of first stage model

	Active Well Number			
	1-mile	3-mile	5-mile	10-mile
	(1)	(2)	(3)	(4)
play	2.536***	25.232***	70.466***	313.419***
	-0.558	-4.962	-12.876	-47.705
aquifer	-0.001	-0.005	-0.01	0.035
	-0.001	-0.007	-0.018	-0.065
Constant	0.197	1.912	4.185	-6.275
	-0.452	-4.021	-10.435	-38.662
Observations	545	545	545	545
R2	0.039	0.048	0.057	0.09
Adjusted R2	0.035	0.045	0.053	0.087
Residual Std. Error (df = 542)	6.041	53.708	139.367	516.354
F Statistic (df = 2; 542)	10.951***	13.722***	16.261***	26.764***

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 4.3. Shale states and organic farms

States	Shales	Well Date Sources	Organic Farms	Uncertified farms
Texas	Permian Basin, Eagle Ford Shale, and Haynesville Shale	TX Railroad Commission	519	84
North Dakota	Bakken Shale	North Dakota Department of Mineral Resources	225	35
Colorado	Niobrara Shale	Colorado Oil and Gas Conservation Commission	546	121
Wyoming	Niobrara Shale	Wyoming Oil and Gas Conservation Commission	102	22
Louisiana	Haynesville Shale	Louisiana Department of Natural Resources Strategic Online Natural Resources Information System	45	8
Pennsylvania	Utica Shale and Marcellus Shale	Pennsylvania Spatial Data Access	1673	168
West Virginia	Utica Shale and Marcellus Shale	West Virginia Department of Environmental Protection	39	5
Ohio	Utica Shale and Marcellus Shale	Ohio Department of Natural Resources	1194	223

Note: This table is modified from website: https://www.fool.com/investing/2017/03/25/which-us-states-produce-the-most-shale-oil.aspx. Eight states, Texas, North Dakota, Colorado, Wyoming, Louisiana, Pennsylvania, West Virginia, and Ohio are included in the research. Because Permian Basin in New Mexico only produce a small portion of shale oil, I am not including New Mexico in my research region. I cut Montana in Bakken Shale by the same reason. The uncertified farm includes suspended and surrendered organic farms.

Table 4.4. Summary statistics of active wells

Statistic	N	Mean	St. Dev.	Min	Max
Within 1 miles	545	1	6	0	54
Within 3 miles	545	12	55	0	467
Within 5 miles	545	36	143	0	1,121
Within 10 miles	545	171	540	0	4,517
Duration	545	6	5	1	22

Note: This table depict the summary statistics of active fracking wells around all the organic farms in Colorado with four different bandwidth choices.

Table 4.5. Summary statistics of active wells around uncertified farms

Statistic	N	Mean	St. Dev.	Min	Max
Within 1 miles	139	4	11	0	54
Within 3 miles	139	34	91	0	467
Within 5 miles	139	91	237	0	1,120
Within 10 miles	139	349	873	0	4,517
Duration	139	4	4	1	22

Note: This table depict the summary statistics of active fracking wells around all the organic farms and uncertified organic farms in Colorado with four different bandwidth choices. There are 139 uncertified organic farms losing their organic certification until year 2018. Comparing the mean of surrounding active fracking well number, the data shows that more active wells are around the uncertified organic farms.

Table 4.6. Parameter of interest, α

	1-mile	exp(-α)		3-mile	exp(-α)		5-mile	exp(-α)		10-mile	exp(-α)	
Well Numbers	0.077	92.64%	**	0.007	99.32%	**	0.002	99.76%	**	0.001	99.93%	**
	(0.033)			(0.003)			(0.001)			(0.000)		
Production Days	0.007	99.35%	***	0.001	99.92%	***	0.000	99.97%	***	0.000	99.99%	***
	(0.002)			(0.000)			(0.000)			(0.000)		
Weighted Distance	0.073	93.01%	**	0.012	98.84%	**	0.007	99.35%	**	0.003	99.67%	**
	(0.030)			(0.006)			(0.003)			(0.002)		
Indicator	0.004	99.65%	**	0.002	99.84%	***	0.001	99.92%	***	0.000	99.96%	***
	(0.001)			(0.001)			(0.000)			(0.000)		

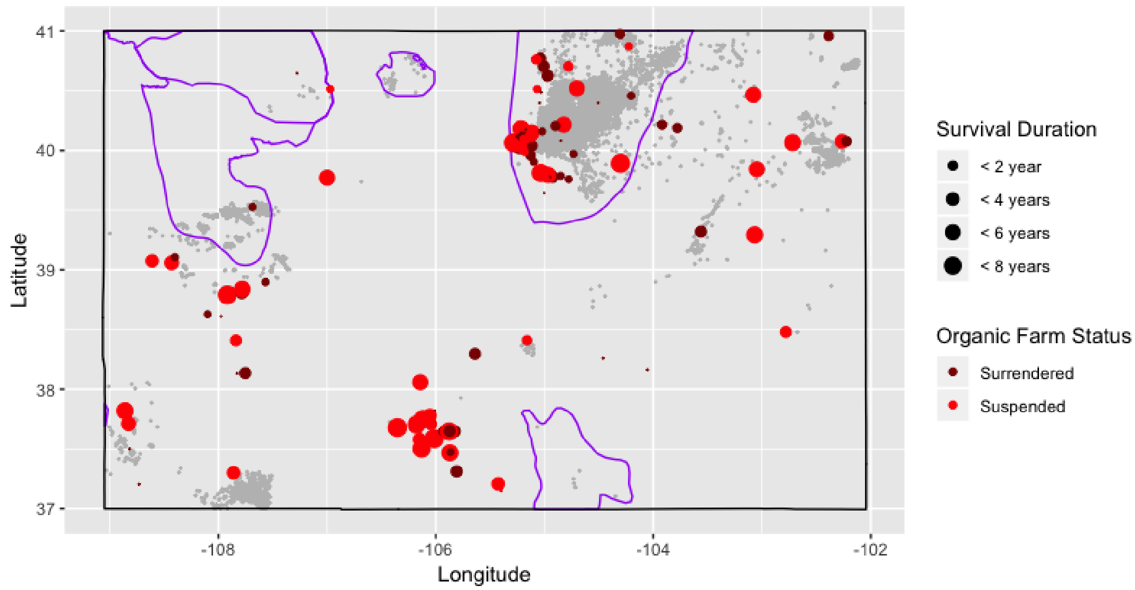


Figure 4.1. Location map of Colorado

Note: The dark red and bright red dots represent uncertified organic farms. Purple line represents the shale plays. Grey dots represent all the active wells in Colorado.

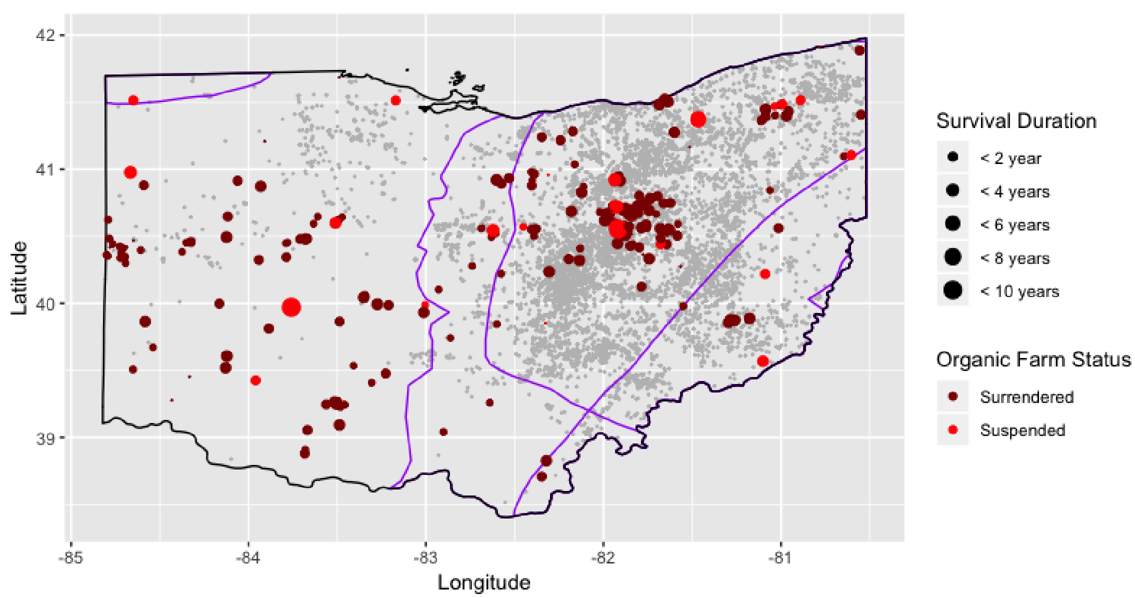


Figure 4.2. Location map of Ohio

Note: The dark red and bright red dots represent uncertified organic farms. Purple line represents the shale plays. Grey dots represent all the active wells in Ohio.

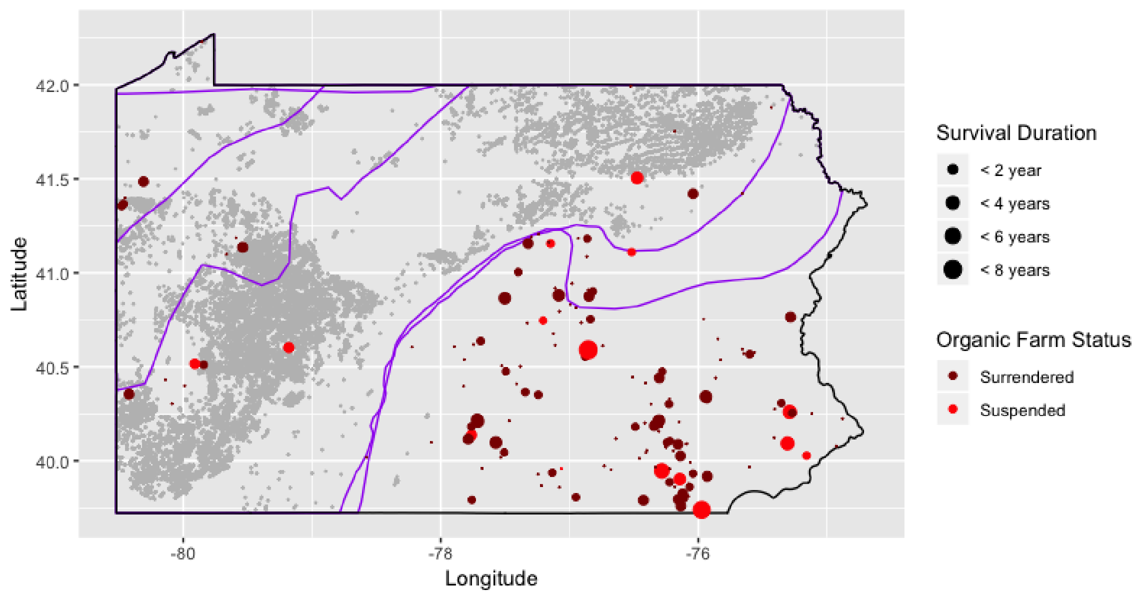


Figure 4.3. Location map of Pennsylvania
Note: The dark red and bright red dots represent uncertified organic farms. Purple line represents the shale plays. Grey dots represent all the active wells in Pennsylvania.

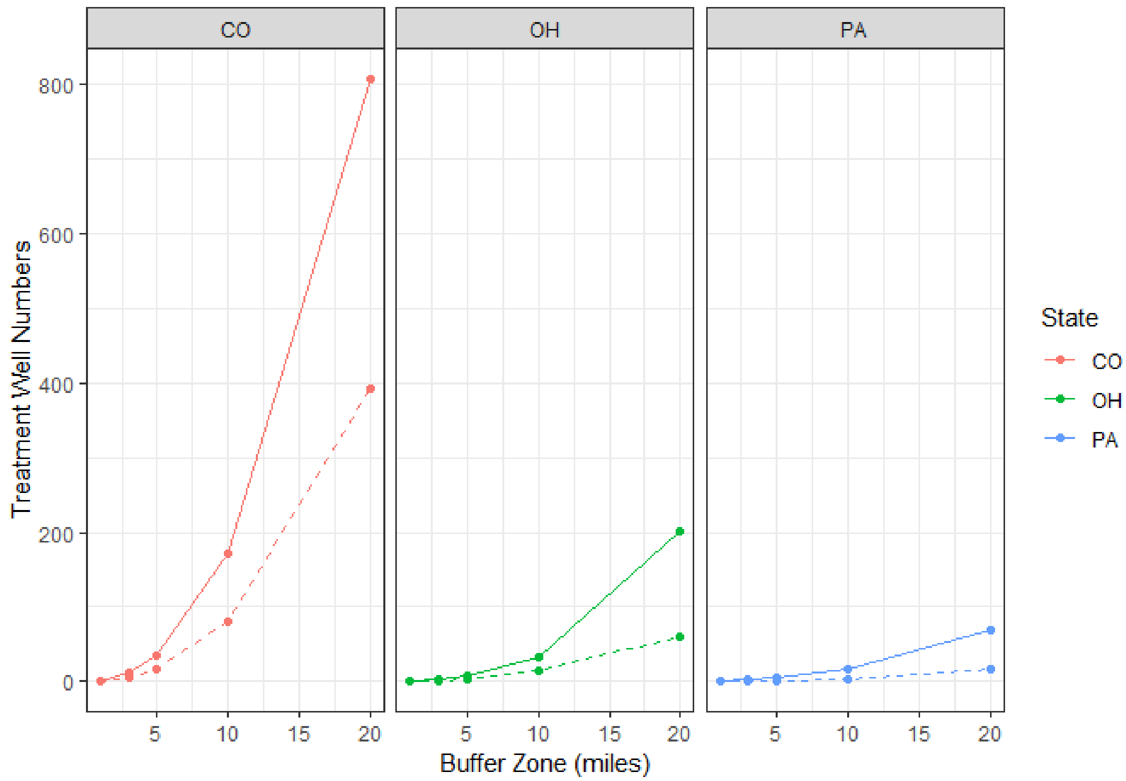


Figure 4.4. Fracking wells around organic farms in the target states
Note: The solid line represents total active wells around both certified and uncertified organic farms. The dashed line represents all the active wells around only uncertified organic farms. Red represents Colorado, green represents Ohio, and blue represents Pennsylvania.

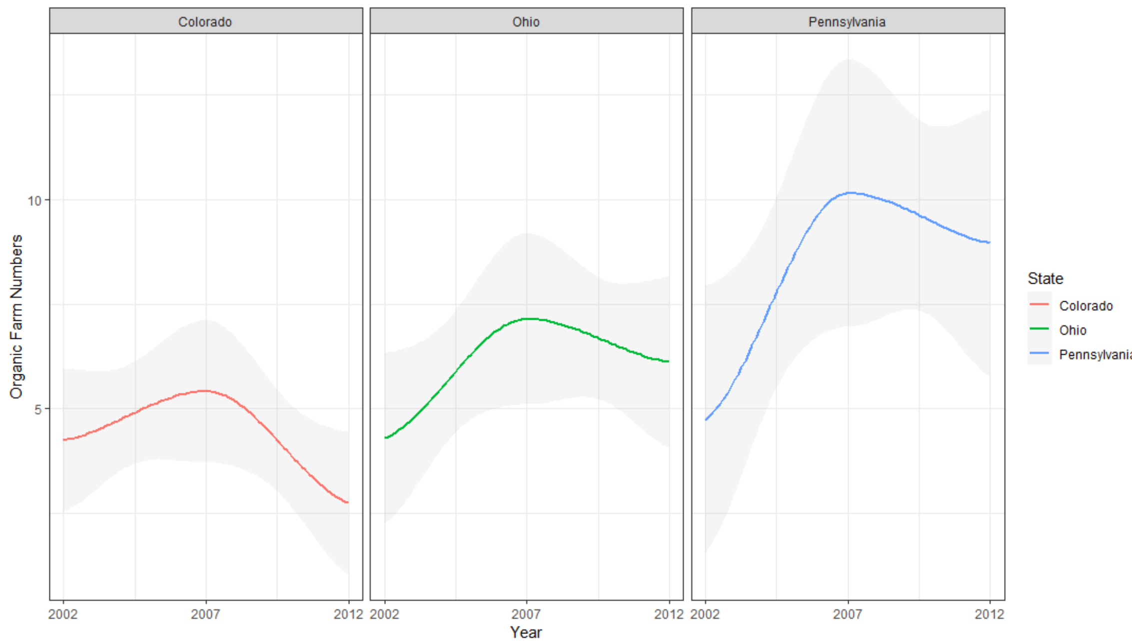


Figure 4.5. Number of organic farms in Colorado, Ohio and Pennsylvania
Note: Red represent Colorado, green represents Ohio and blue represents Pennsylvania. Grey shade represents the 95% confidence interval.

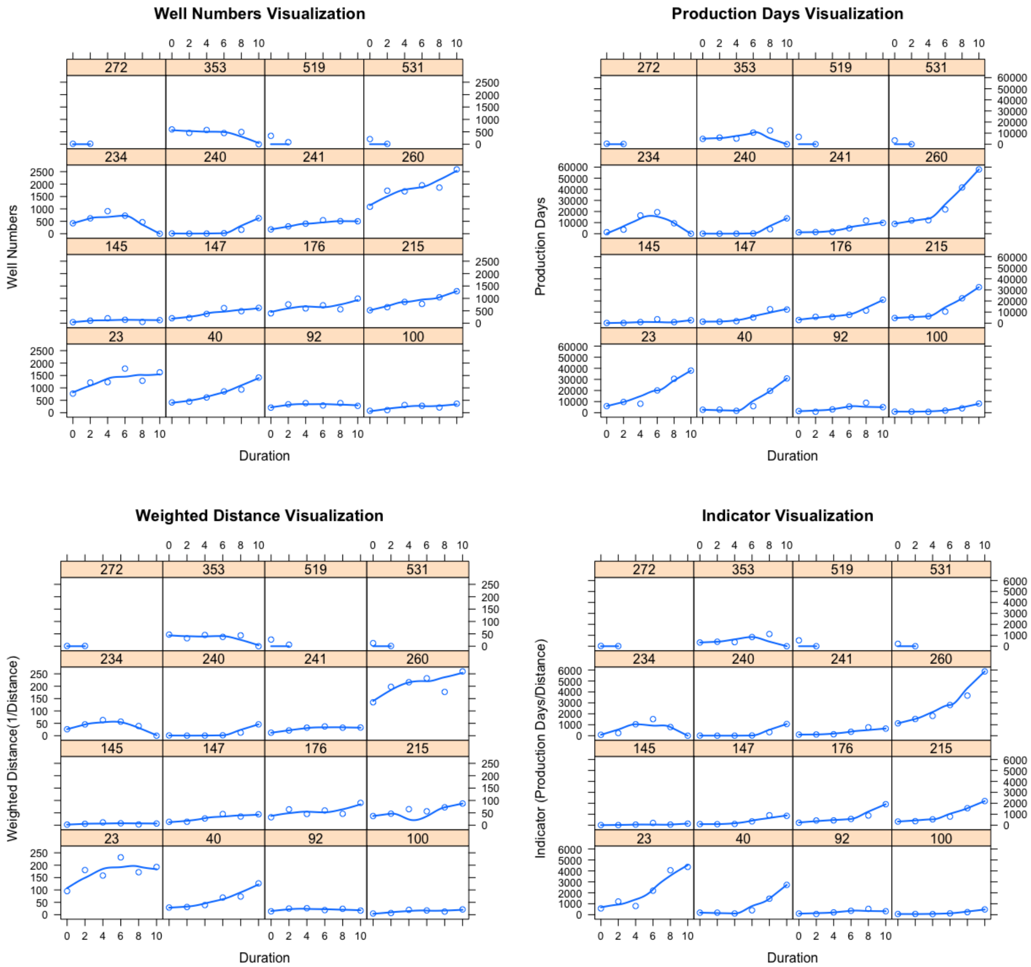


Figure 4.6. Selected active well characteristics visualization
Note: Top left panel shows how well number changes over time for sixteen different organic farms. Top right panel shows the production day visualization. The Bottom left shows the weighted distance visualization and the bottom right shows the indicator visualization. Among the sixteen farms, the farm No.23, No.260, No.519, No.531 lose their certification in their lifetime.

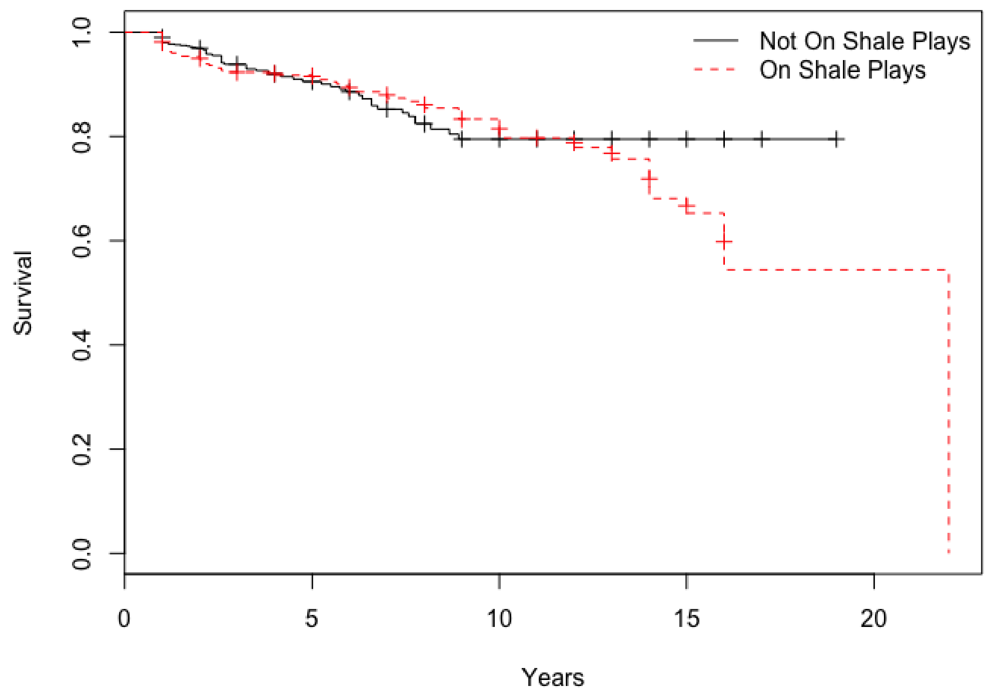


Figure 4.7. Kaplan-Meier estimates of the probability of survival for the control group (Not on Shale Plays) and treatment groups (On shale Plays)

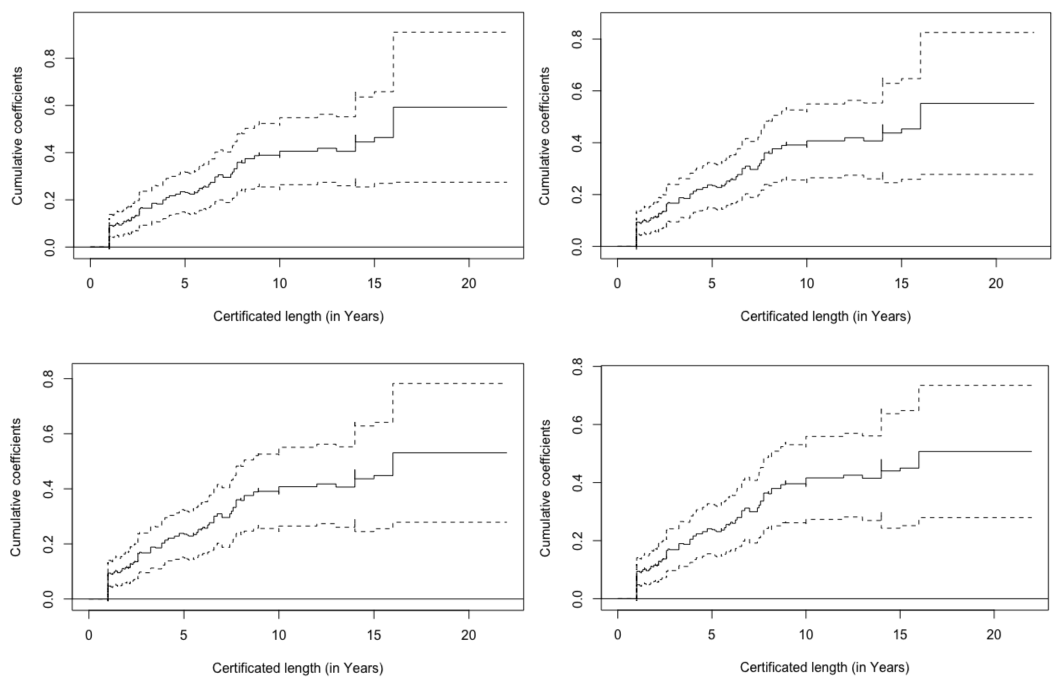


Figure 4.8. Estimated cumulative baseline hazard
Note: The top left panel shows the estimated cumulative baseline hazard of all the organic farms accounting for all the active wells within 1-mile radius. The top right represents the cumulative baseline hazard of 3-mile. The bottom left plot represents the cumulative baseline hazard of 5-mile and the bottom right represents the plot of 10-mile. The time duration is counted in years. The results are displayed together with 95% confidence intervals in dashed line.

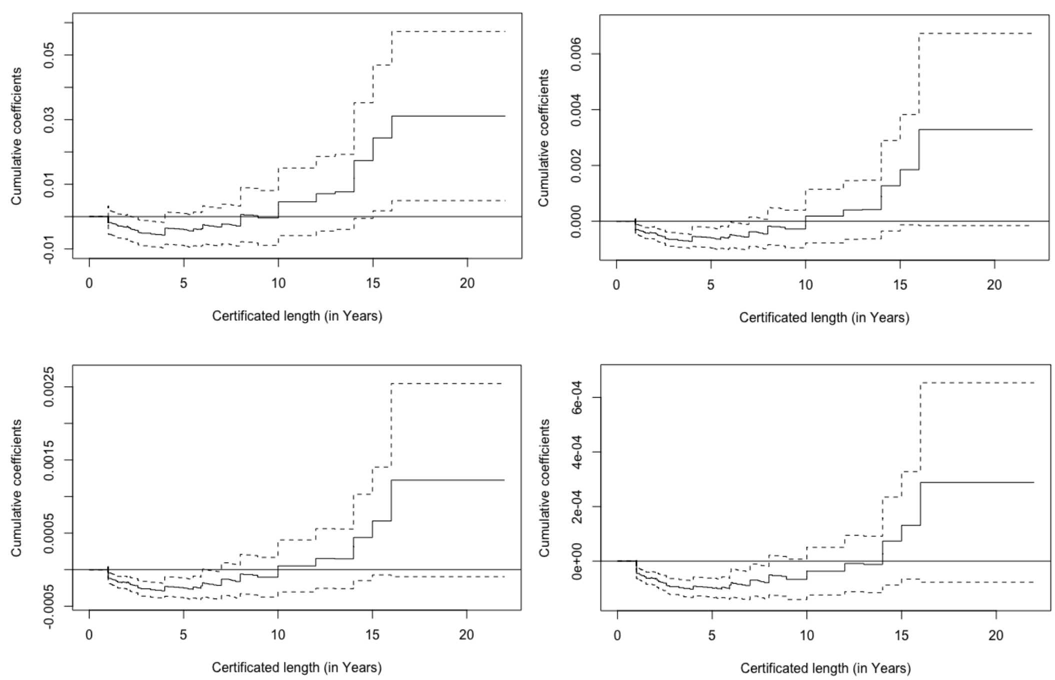


Figure 4.9. Estimated cumulative parameters
Note: The four figures from the top left panel to bottom right panel show the estimated cumulative parameters with 1-mile, 3-mile, 5-mile and 10-mile bandwidth choice. The time duration is counted in years. The results are displayed together with 95% confidence intervals in dashed line.

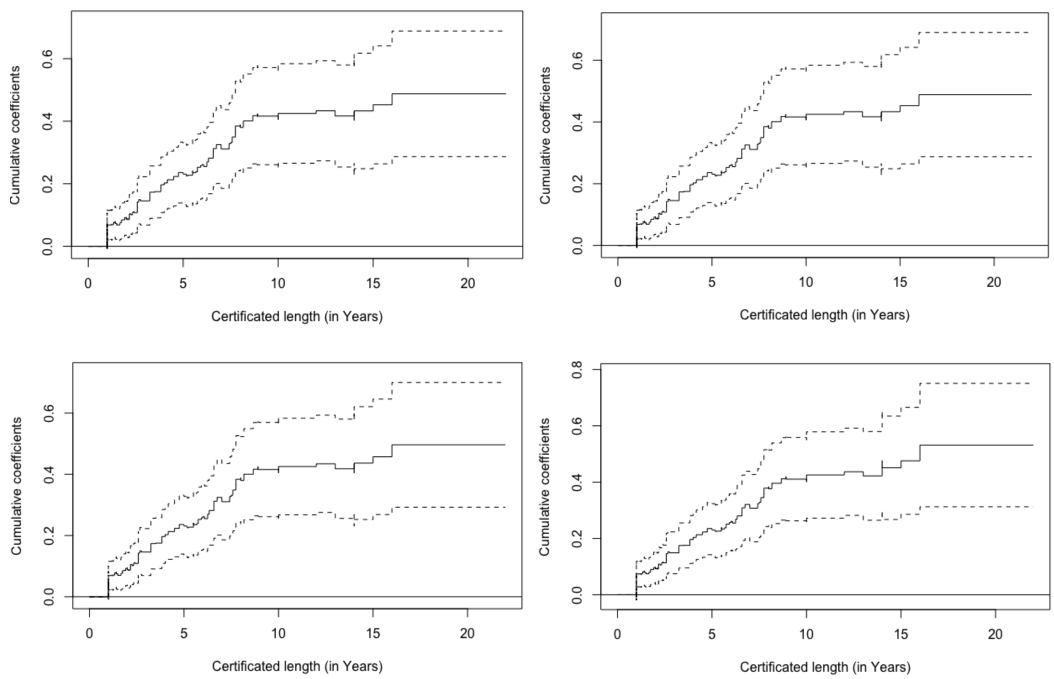


Figure 4.10. Estimated cumulative baseline hazard (two-stage approach)
Note: The top left panel shows the estimated cumulative baseline hazard of all the organic farms accounting for all the active wells within 1-mile radius. The top right represents the cumulative baseline hazard of 3-mile. The bottom left plot represents the cumulative baseline hazard of 5-mile and the bottom right represents the plot of 10-mile. The time duration is counted in years. The results are displayed together with 95% confidence intervals in dashed line.

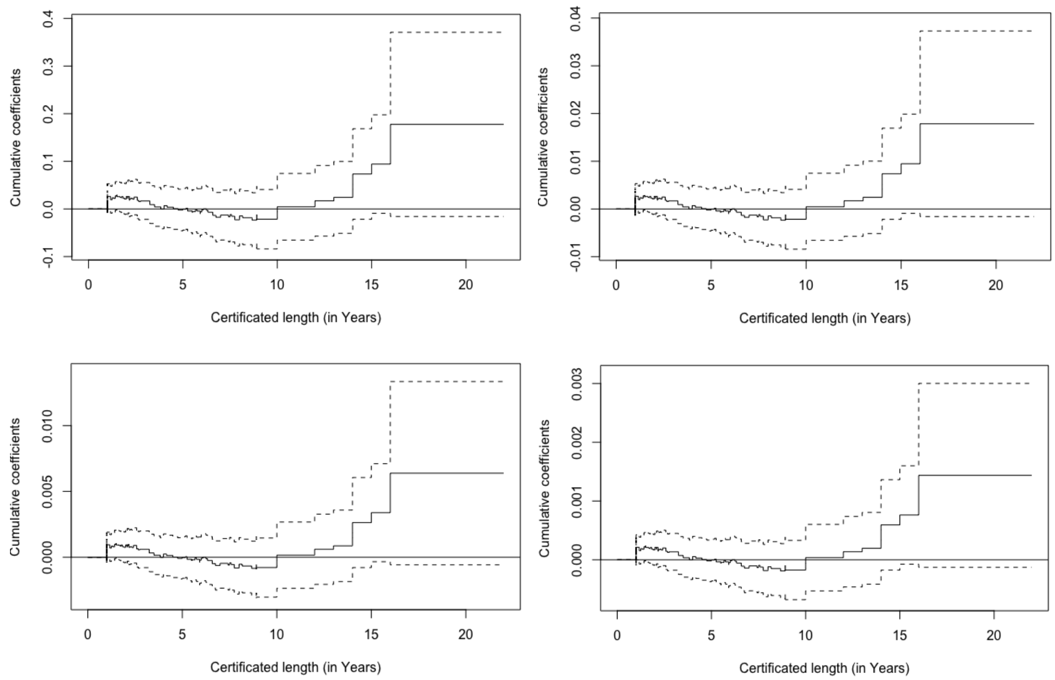


Figure 4.11. Estimated cumulative parameters (two-stage approach)
Note: The four figures from the top left panel to bottom right panel show the estimated cumulative parameters with 1-mile, 3-mile, 5-mile and 10-mile bandwidth choice. The time duration is counted in years. The results are displayed together with 95% confidence intervals in dashed line.

CHAPTER 5

SUMMARY AND CONCLUSIONS

This dissertation contributes to a better understanding of the externalities caused by incomplete ownership of natural resources, such as land, water, oil and gas, and provides three applications of novel econometric techniques to help explain the causal relationships between incomplete ownership and changes in regional development.

The first essay analyzes the relationship between insecure land ownership and lag of agricultural development on American Indian land. Our findings include that tribal lands have lower irrigation rates and significantly less investment in capital-intensive irrigation systems, such as sprinkler irrigation systems. The tribal land is also less likely to be used in high-value crop production compared to their non-tribal neighbors. These results suggest that tribal trust land ownership creates difficulties for tribal farmers in acquiring capital for agricultural investment. A key solution to improving agricultural development on tribal trust land appears to be improving tribal farmers' access to capital, allowing tribal farmers to invest in irrigation systems at the level of their fee-simple neighbors.

The central contribution of this paper is a causal estimate of the effect of tribal ownership on irrigation using a spatial RD approach and a new micro-level dataset linking agricultural irrigation choice, land ownership data, and historic land allocation. Previous studies have linked the limited tribal agricultural development to weak institutions, but, because of data limitation, none of the studies answer this research question using the spatial RD approach with a micro-level dataset. In this study, both sharp and fuzzy RD approaches were used to test the hypothesis, with the results of the two approaches complementing each other. The usage of Public Land Survey System to construct the unit of observation in the data construction procedure can be widely applied to construct micro-level land ownership datasets in the future.

The second essay presents a general equilibrium representation of a hydrologic-ecological-economic system together with synthetic control methodology to explore the impact of the Qualification Settlement Agreement between agricultural and urban regions in California. First, a water trading model between two small open economies is constructed. There are three separate sectors in each economy: agricultural sector, manufacturing sector, and ecosystem sector. One of the economies is rural, with the agricultural sector as the domain sector and the biggest share of labor. The rural economy exports water to urban economy. The theoretical model predicts that after exporting a large amount of water, the return to water in the water exporting region will increase, while the return to skilled and unskilled labor in the water exporting region will decrease. A synthetic control methodology is used in the empirical framework. Counties within California are divided into two groups, a treatment group with only the water exporting county and control group encompassing all other counties. Imperial County was selected as the treatment county in this study. Consistent with the theoretical model, the empirical analysis found similar results: the water transfer agreement in California caused a decline of both skilled and unskilled agricultural labor in Imperial County. The increased water value of the post-trade period can be seen through increased crop yields, but this result is not statistically significant. Adding more pre-trade sample years or extending the research region to the entire western United States may increase the significance of crop yields in the future.

The voluntary fallowing program increasingly conserves water from the agricultural sector. The conserved water is exported to San Diego County and the increased volume of water trading reduces water availability in Imperial County. One of the goals of QSA is to restore the Salton Sea ecosystem with the mitigation water saved from the voluntary fallowing program, however, the reduced water availability may cause Salton Sea to decline further. The water inflow decreases because of the limited water availability. The salinity problem of Salton Sea will become worse as a result of reduced inflow, as there is no outlet for Salton Sea and water is lost only through evaporation.

These results of this paper imply that more job opportunities should be created in

the water exporting region, such as Imperial County, to mitigate the rapid decrease of both the skilled and unskilled labor force. Due to the voluntary fallowing program, on-farm agricultural labor will face a significant drop in fallowing months. Instead of letting unemployed agricultural labor seek employment opportunities out of Imperial County, more job opportunities in other sectors within Imperial County should be created. Another alternative policy could focus on targeting low agricultural and environmental productivity water for transfers to minimize the aforementioned impacts while still creating gains from trade.

To our knowledge, this is the first research study that links an analytic general equilibrium analysis with causal empirical results in studying the performance of water transfers. This work suggests that managing water trading program more effectively could increase their adoption, potentially by determining policies to support the local labor market.

The third essay explores the impact of shale development on organic farming. This study provides empirical evidence of the potential impacts of hydraulic fracturing on the length of time in maintaining organic farming certification in Colorado. Two econometric models, survival analysis with an instrumental variable and a joint model with time-dependent variables, are used. The central contribution of this essay is that it is among the first research studies to use a farm-level geospatial dataset to explore the impact of hydraulic fracturing on farming, and, to my knowledge, the first to examine organic farming.

The paper starts with a discussion about distribution relationships between fracking wells and organic farms. Among all shale states, Colorado has the most exposure to fracking wells, as almost all fracking wells are clustered near organic farms. Hence, Colorado is selected as the target state in this paper. The empirical framework can be divided into three sections starting with a simple Kaplan-Meier estimate of the survival curve of the time until an organic farm fails to be certified. Kaplan-Meier estimators imply that the organic farms on the shale plays have a higher risk of organic certification loss than the farms not on the shale plays. In the second section of empirical framework, an instrumental variable estimation in a survival analysis context is used to solve the endogeneity problem

due to a lack of correspondence between oil and gas deposits and suitable agricultural land. A statistically insignificant negative impact of maintaining organic certification in both a short-distance and a long-distance bandwidth choice is observed for all four bandwidth choices. The impact of exposure to fracking is then treated as a time-dependent variable. Findings suggest that both statistically significant short-distance and long-distance impacts of fracking exist, but the overall impact of active fracking wells on organic farming is not evident, especially for a long-distance bandwidth choice.

This paper suggests that hydraulic fracturing appears to discourage organic farming in Colorado. The impact is quite small, however, due to the short impact time as the fracking boom only started around 2004. Direct benefits from mineral rights leasing and royalty payments would increase farmers' wealth, but these benefits are only limited to the organic farmers with the mineral ownership. The opportunity costs of maintaining organic certification would increase due to the surrounding drilling pads. The intensive fracking activities would discourage many conventional farmers from converting to organic farms. The average waiting time during the transition period would be extended, due to drilling activities around organic farms. Federally subsidized crop insurance could be one of the solutions to lowering the opportunity cost during the initial transition for those farmers without mineral ownership.

REFERENCES

Abadie, A., A. Diamond, and J. Hainmueller. 2015. “Comparative politics and the synthetic control method.” *American Journal of Political Science* 59:495–510.

—. 2010. “Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program.” *Journal of the American statistical Association* 105:493–505.

Adgate, J.L., B.D. Goldstein, and L.M. McKenzie. 2014. “Potential public health hazards, exposures and health effects from unconventional natural gas development.” *Environmental science & technology* 48:8307–8320.

Alchian, A.A., and H. Demsetz. 1973. “The property right paradigm.” *The journal of economic history* 33:16–27.

Allred, B.W., W.K. Smith, D. Twidwell, J.H. Haggerty, S.W. Running, D.E. Naugle, and S.D. Fuhlendorf. 2015. “Ecosystem services lost to oil and gas in North America.” *Science* 348:401–402.

Andersen, P.K., and R.D. Gill. 1982. “Cox’s regression model for counting processes: a large sample study.” *The annals of statistics*, pp. 1100–1120.

Anderson, T.L., and D. Lueck. 1992. “Land tenure and agricultural productivity on Indian reservations.” *The Journal of Law and Economics* 35:427–454.

Anderson, T.L., and D.P. Parker. 2008. “Sovereignty, credible commitments, and economic prosperity on American Indian reservations.” *The Journal of Law and Economics* 51:641–666.

Auch, T. 2015. “Organic farms near drilling activity in the U.S. and Ohio.” Available at: https://www.fractracker.org/2015/03/organic-farms-near-drilling/ [Accessed April 10, 2019].

Bayer, P., F. Ferreira, and R. McMillan. 2007. “A unified framework for measuring preferences for schools and neighborhoods.” *Journal of political economy* 115:588–638.

Besley, T. 1995. “Property rights and investment incentives: Theory and evidence from Ghana.” *Journal of Political Economy* 103:903–937.

Brooks, R., and E. Harris. 2008. “Efficiency gains from water markets: empirical analysis of Watermove in Australia.” *Agricultural Water Management* 95:391–399.

Cai, X. 2008. “Water stress, water transfer and social equity in Northern China—Implications for policy reforms.” *Journal of Environmental Management* 87:14–25.

Calonico, S., M.D. Cattaneo, and R. Titiunik. 2014. “Robust nonparametric confidence intervals for regression-discontinuity designs.” *Econometrica* 82:2295–2326.

Card, D., and L. Giuliano. 2016. "Can tracking raise the test scores of high-ability minority students?" *American Economic Review* 106:2783–2816.

Cornell, S., and J.P. Kalt. 2000. "Where's the glue? Institutional and cultural foundations of American Indian economic development." *The Journal of Socio-Economics* 29:443–470.

Cox, D. 1975. "Partial likelihood. Biometrika, 62, 269-276." *Mathematical Reviews (MathSciNet): MR400509 Zentralblatt MATH* 312.

Craven, J. 2013. "Fracking secrets: the limitations of trade secret protection in hydraulic fracturing." *Vand. J. Ent. & Tech. L.* 16:395.

Cuch, F.S. 2000. *History of Utah's American Indians*. University Press of Colorado.

Dachis, B., G. Duranton, and M.A. Turner. 2011. "The effects of land transfer taxes on real estate markets: evidence from a natural experiment in Toronto." *Journal of economic Geography* 12:327–354.

De la Cuesta, B., and K. Imai. 2016. "Misunderstandings about the regression discontinuity design in the study of close elections." *Annual Review of Political Science* 19:375–396.

Debatin, B.S. 2014. "The Dislocation of Side Effects: Fracking Wastewater Disposal in Appalachian Ohio." Unpublished.

Dell, M. 2010. "The persistent effects of Peru's mining mita." *Econometrica* 78:1863–1903.

—. 2015. "Trafficking networks and the Mexican drug war." *American Economic Review* 105:1738–79.

Demsetz, H. 1974. "Toward a theory of property rights." In *Classic papers in natural resource economics*. Springer, pp. 163–177.

Dumler, T.J., D.H. Rogers, and D.M. O'Brien. 2007. *Irrigation capital requirements and energy costs*. Agricultural Experiment Station and Cooperative Extension Service, Kansas State University.

Farah, N. 2017. "The Effects of Hydraulic Fracturing on Agricultural Productivity." Working paper, Available at: https://drive.google.com/file/d/0Bxp0L418FENjWklucjI3RzFvVkE/view. [Accessed November 10, 2017].

Feder, G., and D. Feeny. 1991. "Land tenure and property rights: Theory and implications for development policy." *The World Bank Economic Review* 5:135–153.

Finkel, M.L., and A. Law. 2011. "The rush to drill for natural gas: a public health cautionary tale." *American Journal of Public Health* 101:784–785.

Fleming, T.R., and D.P. Harrington. 2011. *Counting processes and survival analysis*, vol. 169. John Wiley & Sons.

GAO. 2006. "Additional Opportunities to Reduce Fragmentation, Overlap, and Duplication and Achieve Other Financial Benefits." Working paper No. 22, General Accounting Office.

Gelman, A., and G. Imbens. 2018. "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business & Economic Statistics*, pp. 1–10.

Grafton, R.Q., G. Libecap, S. McGlennon, C. Landry, and B. O'Brien. 2011. "An integrated assessment of water markets: a cross-country comparison." *Review of Environmental Economics and Policy* 5:219–239.

Greene, C., G. Ferreira, A. Carlson, B. Cooke, and C. Hitaj. 2017. "Growing Organic Demand Provides High-Value Opportunities for Many Types of Producers." Available at: https://www.ers.usda.gov/amber-waves/2017/januaryfebruary/growing-organic-demand-provides-high-value-opportunities-for-many-types-of-producers/ [Accessed November 7, 2018].

Grout, C.A., W.K. Jaeger, and A.J. Plantinga. 2011. "Land-use regulations and property values in Portland, Oregon: A regression discontinuity design approach." *Regional Science and Urban Economics* 41:98–107.

Gsell, K. 2012. "Hydraulic Fracturing and Food Security: Can We Have Our Cake and Eat it Too?" Available at: https://blogs.ei.columbia.edu/2012/03/21/hydraulic-fracturing-and-food-security-can-we-have-our-cake-and-eat-it-too/ [Accessed March 5, 2019].

Hadjigeorgalis, E., and J. Lillywhite. 2004. "The impact of institutional constraints on the Limarí River Valley water market." *Water Resources Research* 40.

Hanson, J., R. Dismukes, W. Chambers, C. Greene, and A. Kremen. 2004. "Risk and risk management in organic agriculture: Views of organic farmers." *Renewable agriculture and food systems* 19:218–227.

Hitaj, C., A. Boslett, and J. Weber. 2017. "Fracking, farming, and water." Unpublished.

Hitaj, C., A. Boslett, J.G. Weber, et al. 2014. "Shale development and agriculture." *Choices* 29:1–7.

Holcombe, R.G., and R.S. Sobel. 2001. "Public policy toward pecuniary externalities." *Public Finance Review* 29:304–325.

Howe, C.W., J.K. Lazo, and K.R. Weber. 1990. "The economic impacts of agriculture-to-urban water transfers on the area of origin: a case study of the Arkansas River Valley in Colorado." *American Journal of Agricultural Economics* 72:1200–1204.

Imbens, G.W., and T. Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of econometrics* 142:615–635.

Juana, J.S., K.M. Strzepek, and J.F. Kirsten. 2011. "Market efficiency and welfare effects of inter-sectoral water allocation in South Africa." *Water Policy* 13:220–231.

Kuhlman, M. 2014. "Does Fracking Threaten Future of Ohio Organic Farms?" Available at: https://www.publicnewsservice.org/2014-10-13/environment/does-fracking-threaten-future-of-ohio-organic-farms/a42209-1 [Accessed April 7, 2018].

Lergenmuller, S. 2017. "Two-stage predictor substitution for time-to-event data." MS thesis.

Mann, S., and H. Wüstemann. 2008. "Multifunctionality and a new focus on externalities." *The Journal of Socio-Economics* 37:293–307.

Marie, C. 2017. "Shale Law in the Spotlight: Chemical Disclosure." Available at: http://www.pennstateshalelaw.com/2017/07/shale-law-in-spotlight-chemical.html [Accessed August 5, 2018].

McBride, W.D., and C.R. Greene. 2008. "The Profitability of Organic Soybean Production.", pp. 26.

McBride, W.D., C.R. Greene, M.B. Ali, and L.F. Foreman. 2012. "The Structure and Profitability of Organic Field Crop Production: The Case of Wheat.", pp. 26.

McCrary, J. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of econometrics* 142:698–714.

McKenzie, L.M., R.Z. Witter, L.S. Newman, and J.L. Adgate. 2012. "Human health risk assessment of air emissions from development of unconventional natural gas resources." *Science of the Total Environment* 424:79–87.

McPherson, R.S. 2000. "Setting the stage: Native America revisited." *A History of Utah's American Indians*, pp. 3–24.

Muehlenbachs, L., E. Spiller, and C. Timmins. 2015. "The housing market impacts of shale gas development." *American Economic Review* 105:3633–59.

Munoz-Price, L.S., J.F. Frencken, S. Tarima, and M. Bonten. 2016. "Handling time-dependent variables: antibiotics and antibiotic resistance." *Clinical Infectious Diseases* 62:1558–1563.

Murrill, B.J., and A. Vann. 2012. "Hydraulic Fracturing: Chemical Disclosure Requirements." Available at: https://fas.org/sgp/crs/misc/R42461.pdf [Accessed April 5, 2019].

Ong, B. 2014. "The potential impacts of hydraulic fracturing on agriculture." *European Journal of Sustainable Development* 3:63–72.

Ortiz, C., and J. Dessert. 2017. "2016 Imperial County Agricultural Crop and Livestock Report." Working paper, County of Imperial, August.

Pan, Y., S.C. Smith, and M. Sulaiman. 2018. "Agricultural extension and technology adoption for food security: evidence from Uganda." *American Journal of Agricultural Economics* 100:1012–1031.

Pulwarty, R.S., K.L. Jacobs, and R.M. Dole. 2005. "The hardest working river: drought and critical water problems in the Colorado River Basin." *Drought and water crises: Science, technology, and management issues*, pp. 249–285.

Rakitan, T. 2018. "The impact of the US shale boom on agriculture: Evidence from North Dakota." Available at SSRN 3135519.

Reisner, M. 1993. *Cadillac desert: The American West and its disappearing water*. Penguin.

Rizopoulos, D. 2012. *Joint models for longitudinal and time-to-event data: With applications in R*. Chapman and Hall/CRC.

Royte, E. 2012. "Fracking our food supply." Available at: http://www.hpaf.co.uk/wp-content/uploads/2018/09/Fracking-Our-Food-Supply-_-The-Nation.pdf [Accessed April 5, 2019].

San Diego County Water Authority. 2019. "Quantification Settlement Agreement for the Colorado River." Working paper, San Diego County Water Authority, March.

Schaetzl, R.J., F.J. Krist Jr, and B.A. Miller. 2012. "A taxonomically based ordinal estimate of soil productivity for landscape-scale analyses." *Soil Science* 177:288–299.

Schewe, J., J. Heinke, D. Gerten, I. Haddeland, N.W. Arnell, D.B. Clark, R. Dankers, S. Eisner, B.M. Fekete, F.J. Colón-González, et al. 2014. "Multimodel assessment of water scarcity under climate change." *Proceedings of the National Academy of Sciences* 111:3245–3250.

Shoemaker, J.A. 2006. "Farm and Ranch Issues in Indian Country." Available at: http://www.flaginc.org/wp-content/uploads/2013/03/CLE_JS.pdf [Accessed April 5, 2017].

Sustainable Agriculture Research & Education. 2012. "Economics of Organic Production." Available at: https://www.sare.org/Learning-Center/Bulletins/Transitioning-to-Organic-Production/Text-Version/Economics-of-Organic-Production [Accessed April 5, 2019].

U.S. Department of Agriculture. 2019. "Organic Enforcement." Available at: https://www.ams.usda.gov/services/enforcement/organic. [Accessed April 5, 2019].

—. 2009. "Organic Production and Handling Requirements." Available at: https://www.law.cornell.edu/cfr/text/7/205.202 [Accessed April 5, 2019].

Vidic, R.D., S.L. Brantley, J.M. Vandenbossche, D. Yoxtheimer, and J.D. Abad. 2013. "Impact of shale gas development on regional water quality." *science* 340:1235009.

Vörösmarty, C.J., P. Green, J. Salisbury, and R.B. Lammers. 2000. "Global water resources: vulnerability from climate change and population growth." *science* 289:284–288.

Weber, J., J. Brown, and J. Pender. 2013. "Rural wealth creation and emerging energy industries: lease and royalty payments to farm households and businesses." Unpublished.

Wimmer, F., E. Audsley, M. Malsy, C. Savin, R. Dunford, P.A. Harrison, R. Schaldach, and M. Flörke. 2015. "Modelling the effects of cross-sectoral water allocation schemes in Europe." *Climatic Change* 128:229–244.

APPENDICES

APPENDIX A

Chapter 2 Appendix

Table A.1. Crops value classification

Crops	Crop Value	Crop Value Indicator
Alfalfa	Low	0
Beans	High	1
Berries	High	1
Corn	High	1
Dry Alfalfa	Low	0
Dry Beans	High	1
Dry Grain	High	1
Dry Grain/Seeds	High	1
Dry Oats	High	1
Dry Safflower	High	1
Fallow-Irrigated Ag	Low	0
Fallow-Irrigated Land	Low	0
Grain	High	1
Grass Hay	Low	0
Grass Hay-sub-irrigated	Low	0
Idle-Irrigated Ag	Low	0
Idle-Irrigated Land	Low	0
Idle-Irrigated Pasture	Low	0
Melon/Pumpkin/Squash	High	1
Oats	High	1
Onions	High	1
Orchard	High	1
Other Horticulture	High	1
Other Vegetables	High	1
Pasture	Low	0
Pasture-sub-irrigated	Low	0
Potatoes	High	1
Safflower	High	1
Sorghum	High	1
Tomatoes	High	1
Turf Farms	High	1
Vineyard	High	1

Table A.2. Sharp RD results of soil productivity index

Sample Within	Soil Productivity index						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Allotment1905	0.063	0.083	0.056	0.051	0.026	0.063	1.281
	(0.098)	(0.104)	(0.111)	(0.117)	(0.120)	(0.098)	
Second order polynomial							
Allotment1905	0.139	0.117	0.103	0.084	0.067	0.139	1.289
	(0.115)	(0.118)	(0.120)	(0.122)	(0.127)	(0.115)	
Third order polynomial							
Allotment1905	0.137	0.107	0.090	0.036	-0.010	0.137	1.732
	(0.122)	(0.123)	(0.124)	(0.129)	(0.138)	(0.122)	
Fourth order polynomial							
Allotment1905	0.143	0.134	0.068	0.005	-0.318 **	0.143	2.168
	(0.125)	(0.126)	(0.132)	(0.138)	(0.147)	(0.125)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(8424/6991)	(7185/6542)	(5870/5999)	(4555/5416)	(3062/4714)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.3. Sharp RD results of agricultural rate (CDL dataset)

Sample Within	Agricultural Rate						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Allotment1905	0.010	0.006	0.003	0.008	0.009	0.004	0.611
	(0.015)	(0.016)	(0.018)	(0.019)	(0.019)	(0.018)	
Second order polynomial							
Allotment1905	0.004	0.004	0.007	0.008	0.015	0.003	1.328
	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)	(0.018)	
Third order polynomial							
Allotment1905	0.005	0.008	0.011	0.015	0.020	0.003	2.222
	(0.019)	(0.020)	(0.020)	(0.020)	(0.021)	(0.018)	
Fourth order polynomial							
Allotment1905	0.009	0.012	0.014	0.020	0.022	0.004	2.070
	(0.020)	(0.020)	(0.020)	(0.021)	(0.022)	(0.019)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(3398/3133)	(2970/3087)	(2521/2998)	(2077/2893)	(1486/2714)		

Note: this table uses cropscape dataset. Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.4. Sharp RD results of agricultural rate (WRL dataset)

Sample Within	Agricultural Rate						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Allotment1905	0.026 *	0.033	0.042	0.045	0.046 *	0.040	0.731
	(0.020)	(0.021)	(0.023)	(0.025)	(0.026)	(0.022)	
Second order polynomial							
Allotment1905	0.040	0.044	0.041 *	0.040	0.037	0.037	1.699
	(0.024)	(0.025)	(0.026)	(0.026)	(0.027)	(0.023)	
Third order polynomial							
Allotment1905	0.043	0.039	0.039	0.036 *	0.031	0.038 *	1.140
	(0.026)	(0.026)	(0.026)	(0.027)	(0.028)	(0.026)	
Fourth order polynomial							
Allotment1905	0.037	0.036	0.033 *	0.034	0.030 *	0.039	1.603
	(0.026)	(0.027)	(0.027)	(0.028)	(0.030)	(0.026)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(3337/2623)	(2922/2596)	(2475/2545)	(2008/2479)	(1415/2347)		

Note: This table uses Uath GIS Portal dataset. Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.5. Sharp RD results of irrigation rate

Sample Within	Irrigation Rate						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Allotment1905	-0.020	-0.009	0.007	0.019	0.026	0.019	0.483
	(0.020)	(0.021)	(0.023)	(0.025)	(0.026)	(0.025)	
Second order polynomial							
Allotment1905	0.005	0.015	0.017	0.018	0.016	-0.002	1.595
	(0.024)	(0.025)	(0.025)	(0.026)	(0.027)	(0.023)	
Third order polynomial							
Allotment1905	0.017	0.015	0.016	0.014	-0.001	0.012	1.375
	(0.026)	(0.026)	(0.026)	(0.027)	(0.028)	(0.026)	
Fourth order polynomial							
Allotment1905	0.012	0.012	0.008	0.002	-0.020	0.012	1.567
	(0.026)	(0.027)	(0.027)	(0.028)	(0.030)	(0.026)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(3692/3000)	(3230/2963)	(2738/2906)	(2218/2906)	(1567/2661)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.6. Sharp RD results sprinkle-irrigated rate

Sample Within	Sprinkle-irrigated Rate											
	<1.5 Miles		<1.25 Miles		<1 Miles		<0.75 Miles		<0.5 Miles		Optimal Bandwidth		
First order polynomial												Optimal Miles
Allotment1905	-0.129	***	-0.140	***	-0.136	***	-0.123	***	-0.115	***	-0.125	***	0.475
	(0.021)		(0.023)		(0.025)		(0.027)		(0.028)		(0.027)		
Second order polynomial													
Allotment1905	-0.146	***	-0.133	***	-0.124	***	-0.119	***	-0.102	***	-0.120	***	0.954
	(0.026)		(0.027)		(0.028)		(0.028)		(0.029)		(0.028)		
Third order polynomial													
Allotment1905	-0.131	***	-0.124	***	-0.118	***	-0.104	***	-0.093	***	-0.117	***	0.968
	(0.028)		(0.028)		(0.028)		(0.029)		(0.030)		(0.028)		
Fourth order polynomial													
Allotment1905	-0.123	***	-0.118	***	-0.108	***	-0.096	***	-0.100	***	-0.099	***	0.794
	(0.028)		(0.029)		(0.029)		(0.030)		(0.031)		(0.029)		
Control Variables													
Townships	Ã		Ã		Ã		Ã		Ã		Ã		
Elevation	Ã		Ã		Ã		Ã		Ã		Ã		
Soil Productivity	Ã		Ã		Ã		Ã		Ã		Ã		
Observations (out/in)	(2819/1871)		(2467/1956)		(2083/1926)		(1698/1881)		(1203/1797)				

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.7. Sharp RD results of high-value crops rate (CDL dataset)

Sample Within	High-value Crops Rate (2012)												
	<1.5 Miles		<1.25 Miles		<1 Miles		<0.75 Miles		<0.5 Miles		Optimal Bandwidth		
First order polynomial													Optimal Miles
Allotment1905	-0.046	***	-0.045	***	-0.045	***	-0.041	***	-0.037	***	-0.041	***	0.498
	(0.011)		(0.012)		(0.013)		(0.014)		(0.014)		(0.014)		
Second order polynomial													
Allotment1905	-0.046	***	-0.046	***	-0.042	***	-0.038	***	-0.037	**	-0.045	***	0.972
	(0.013)		(0.014)		(0.014)		(0.014)		(0.015)		(0.014)		
Third order polynomial													
Allotment1905	-0.045	***	-0.042	***	-0.038	***	-0.035	**	-0.038	**	-0.046	***	1.814
	(0.014)		(0.014)		(0.014)		(0.015)		(0.016)		(0.014)		
Fourth order polynomial													
Allotment1905	-0.042	***	-0.039	***	-0.038	**	-0.040	***	-0.044	***	-0.052	***	2.171
	(0.014)		(0.015)		(0.015)		(0.016)		(0.016)		(0.014)		
Control Variables													
Townships	Ã		Ã		Ã		Ã		Ã		Ã		
Elevation	Ã		Ã		Ã		Ã		Ã		Ã		
Soil Productivity	Ã		Ã		Ã		Ã		Ã		Ã		
Observations (out/in)	(3398/3133)		(2970/3087)		(2521/2998)		(2077/2893)		(1486/2714)				

Note: This table use Cropscape dataset. Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.8. Sharp RD results of high-value crops rate (CDL dataset)

Sample Within	High-value Crops Rate (2012)						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Allotment1905	-0.008	-0.008	-0.010	-0.012	-0.012	-0.005	0.919
	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.006)	
Second order polynomial							
Allotment1905	-0.012	-0.013 *	-0.014 *	-0.014 *	-0.013	-0.009	2.162
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.006)	
Third order polynomial							
Allotment1905	-0.015 ***	-0.015 *	-0.014 *	-0.013	-0.015 *	-0.011	2.262
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.007)	
Fourth order polynomial							
Allotment1905	-0.016 **	-0.016 *	-0.015 *	-0.015	-0.015	-0.016 **	1.643
	(0.008)	(0.008)	(0.009)	(0.009)	(0.010)	(0.008)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(3337/2623)	(2922/2596)	(2475/2545)	(2008/2479)	(1415/2347)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.9. Fuzzy RD results of soil productivity index

Sample Within	Soil Productivity index						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Tribal2017	1.015 ***	0.967 ***	0.956 ***	1.014 ***	1.057 ***	0.939 ***	0.681
	(0.314)	(0.334)	(0.358)	(0.385)	(0.408)	(0.364)	
Second order polynomial							
Tribal2017	1.011 ***	1.060 ***	1.111 ***	1.123 ***	1.143 ***	1.025 ***	1.521
	(0.385)	(0.400)	(0.414)	(0.422)	(0.435)	(0.359)	
Third order polynomial							
Tribal2017	1.071 **	1.079 **	1.083 **	1.096 **	1.208 ***	1.001 **	1.730
	(0.419)	(0.424)	(0.427)	(0.437)	(0.451)	(0.406)	
Fourth order polynomial							
Tribal2017	1.108 **	1.110 **	1.043 **	1.033 **	0.764	1.044 **	2.505
	(0.432)	(0.434)	(0.444)	(0.455)	(0.478)	(0.415)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(22882/20438)	(19598/19264)	(16177/17867)	(12787/16261)	(9116/14114)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.10. Fuzzy RD results of agricultural rate (CDL dataset)

Sample Within	Agricultural Rate						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Tribal2017	-0.010	-0.029	-0.041	-0.035	-0.021	-0.043	0.796
	(0.025)	(0.027)	(0.030)	(0.033)	(0.036)	(0.032)	
Second order polynomial							
Tribal2017	-0.048	-0.043	-0.030	-0.019	-0.003	-0.048	1.323
	(0.033)	(0.034)	(0.036)	(0.036)	(0.037)	(0.034)	
Third order polynomial							
Tribal2017	-0.032	-0.022	-0.015	0.000	0.011	-0.048	1.963
	(0.036)	(0.036)	(0.036)	(0.037)	(0.038)	(0.034)	
Fourth order polynomial							
Tribal2017	-0.025	-0.015	-0.009	0.002	0.006	0.022	1.674
	(0.036)	(0.036)	(0.037)	(0.037)	(0.038)	(0.040)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(5905/4879)	(5318/4826)	(4649/4744)	(3921/4646)	(3063/4467)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.11. Fuzzy RD results of agricultural rate (WRL dataset)

Sample Within	Agricultural Rate						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Tribal2017	-0.010	-0.011	-0.018	-0.024	-0.023	-0.033	0.879
	(0.031)	(0.033)	(0.036)	(0.040)	(0.042)	(0.042)	
Second order polynomial							
Tribal2017	-0.020	-0.026	-0.027	-0.025	-0.024	-0.024	1.355
	(0.039)	(0.041)	(0.042)	(0.043)	(0.044)	(0.040)	
Third order polynomial							
Tribal2017	-0.028	-0.028	-0.025	-0.021	-0.027	-0.026	1.704
	(0.043)	(0.043)	(0.043)	(0.044)	(0.045)	(0.042)	
Fourth order polynomial							
Tribal2017	-0.029	-0.024	-0.022	-0.027	-0.031	-0.027	1.603
	(0.043)	(0.043)	(0.044)	(0.045)	(0.050)	(0.043)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(4954/3540)	(4462/3531)	(3922/3505)	(3345/3470)	(2603/3399)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.12. Fuzzy RD results of irrigation rate

Sample Within	Irrigation Rate						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Tribal2017	-0.191 ***	-0.192 ***	-0.195 ***	-0.192 ***	-0.182 ***	-0.188 ***	1.186
	(0.033)	(0.035)	(0.038)	(0.042)	(0.044)	(0.042)	
Second order polynomial							
Tribal2017	-0.195 ***	-0.193 ***	-0.187 ***	-0.182 ***	-0.177 ***	-0.192 ***	1.396
	(0.042)	(0.043)	(0.044)	(0.045)	(0.046)	(0.043)	
Third order polynomial							
Tribal2017	-0.189 ***	-0.186 ***	-0.182 ***	-0.175 ***	-0.188 ***	-0.190 ***	1.179
	(0.045)	(0.045)	(0.045)	(0.046)	(0.048)	(0.045)	
Fourth order polynomial							
Tribal2017	-0.185 ***	-0.178 ***	-0.176 ***	-0.191 ***	-0.214 ***	-0.180 ***	1.142
	(0.045)	(0.045)	(0.046)	(0.047)	(0.052)	(0.045)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(5503/4141)	(4944/4124)	(4344/4090)	(3707/4042)	(2902/3953)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.13. Fuzzy RD results of sprinkle-irrigation rate

Sample Within	Sprinkle-irrigated Rate												
	<1.5 Miles		<1.25 Miles		<1 Miles		<0.75 Miles		<0.5 Miles		Optimal Bandwidth		
First order polynomial													Optimal Miles
Tribal2017	-0.300	***	-0.311	***	-0.318	***	-0.316	***	-0.314	***	-0.312	***	1.108
	(0.033)		(0.036)		(0.040)		(0.044)		(0.046)		(0.042)		
Second order polynomial													
Tribal2017	-0.323	***	-0.319	***	-0.313	***	-0.309	***	-0.305	***	-0.320	***	1.319
	(0.044)		(0.046)		(0.047)		(0.047)		(0.048)		(0.045)		
Third order polynomial													
Tribal2017	-0.316	***	-0.312	***	-0.309	***	-0.304	***	-0.289	***	-0.323	***	1.641
	(0.047)		(0.047)		(0.047)		(0.048)		(0.049)		(0.047)		
Fourth order polynomial													
Tribal2017	-0.309	***	-0.304	***	-0.301	***	-0.298	***	-0.303	***	-0.312	***	1.624
	(0.047)		(0.047)		(0.048)		(0.049)		(0.054)		(0.047)		
Control Variables													
Townships	Ã		Ã		Ã		Ã		Ã		Ã		
Elevation	Ã		Ã		Ã		Ã		Ã		Ã		
Soil Productivity	Ã		Ã		Ã		Ã		Ã		Ã		
Observations (out/in)	(4197/2566)		(3775/2565)		(3317/2560)		(2834/2547)		(2216/2510)				

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.14. Fuzzy RD results of high-value crops rate (CDL dataset)

Sample Within	High-value Crops Rate (2012)						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Tribal2017	-0.140 ***	-0.142 ***	-0.144 ***	-0.147 ***	-0.143 ***	-0.144 ***	0.820
	(0.018)	(0.020)	(0.021)	(0.024)	(0.026)	(0.023)	
Second order polynomial							
Tribal2017	-0.147 ***	-0.147 ***	-0.147 ***	-0.143 ***	-0.144 ***	-0.153 ***	1.316
	(0.024)	(0.025)	(0.026)	(0.026)	(0.027)	(0.026)	
Third order polynomial							
Tribal2017	-0.149 ***	-0.148 ***	-0.145 ***	-0.142 ***	-0.150 ***	-0.144 ***	2.072
	(0.026)	(0.026)	(0.027)	(0.027)	(0.028)	(0.024)	
Fourth order polynomial							
Tribal2017	-0.148 ***	-0.144 ***	-0.144 ***	-0.157 ***	-0.173 ***	-0.150 ***	2.515
	(0.027)	(0.027)	(0.028)	(0.028)	(0.030)	(0.026)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(5905/4879)	(5318/4826)	(4649/4744)	(3921/4646)	(3063/4467)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

Table A.15. Fuzzy RD results of high-value crops rate (WRL dataset)

Sample Within	High-value Crops Rate (2012)						
	<1.5 Miles	<1.25 Miles	<1 Miles	<0.75 Miles	<0.5 Miles	Optimal Bandwidth	
First order polynomial							Optimal Miles
Tribal2017	-0.027 ***	-0.034 ***	-0.039 ***	-0.044 ***	-0.044 ***	-0.045 ***	1.121
	(0.010)	(0.011)	(0.012)	(0.013)	(0.014)	(0.013)	
Second order polynomial							
Tribal2017	-0.043 ***	-0.044 ***	-0.045 ***	-0.044 ***	-0.042 ***	-0.044 ***	1.168
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.013)	
Third order polynomial							
Tribal2017	-0.047 ***	-0.047 ***	-0.047 ***	-0.043 ***	-0.040 ***	-0.048 ***	1.443
	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.014)	
Fourth order polynomial							
Tribal2017	-0.046 ***	-0.046 ***	-0.043 ***	-0.044 ***	-0.048 ***	-0.048 ***	1.772
	(0.014)	(0.014)	(0.015)	(0.015)	(0.018)	(0.014)	
Control Variables							
Townships	Ã	Ã	Ã	Ã	Ã	Ã	
Elevation	Ã	Ã	Ã	Ã	Ã	Ã	
Soil Productivity	Ã	Ã	Ã	Ã	Ã	Ã	
Observations (out/in)	(4954/3540)	(4462/3531)	(3922/3505)	(3345/3470)	(2603/3399)		

Note: Coefficients significantly different from zero are denoted by the following system: * 10%, ** 5%, and *** 1%. Robust standard errors are provided in the parenthesis.

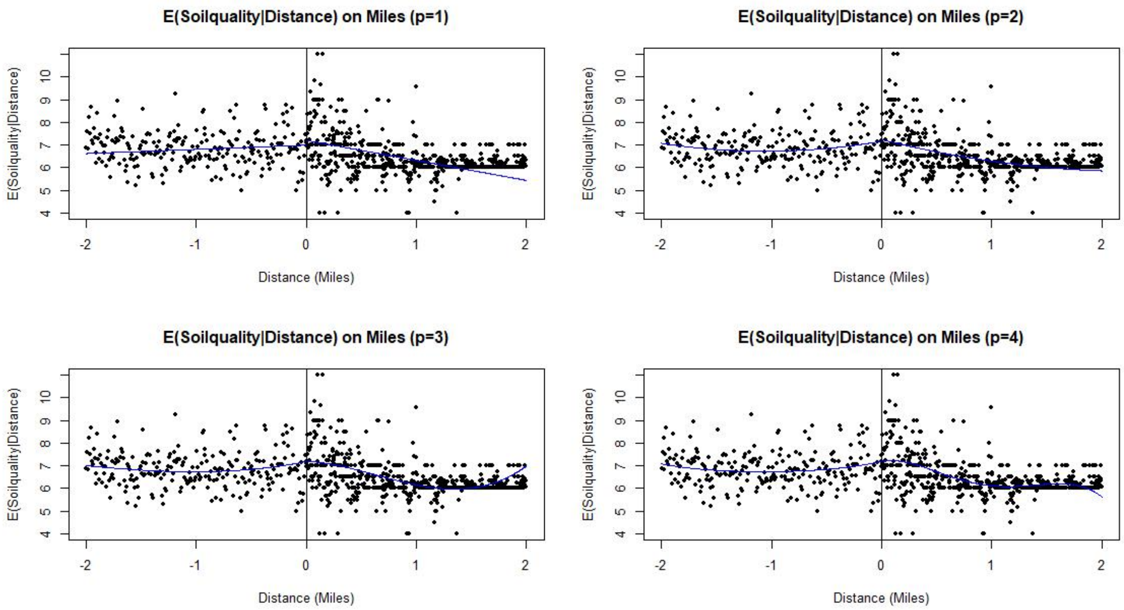


Figure A.1. RD plots of soil quality with 1st, 2nd, 3rd, and 4th order polynomial using 1905 Allotment Boundary

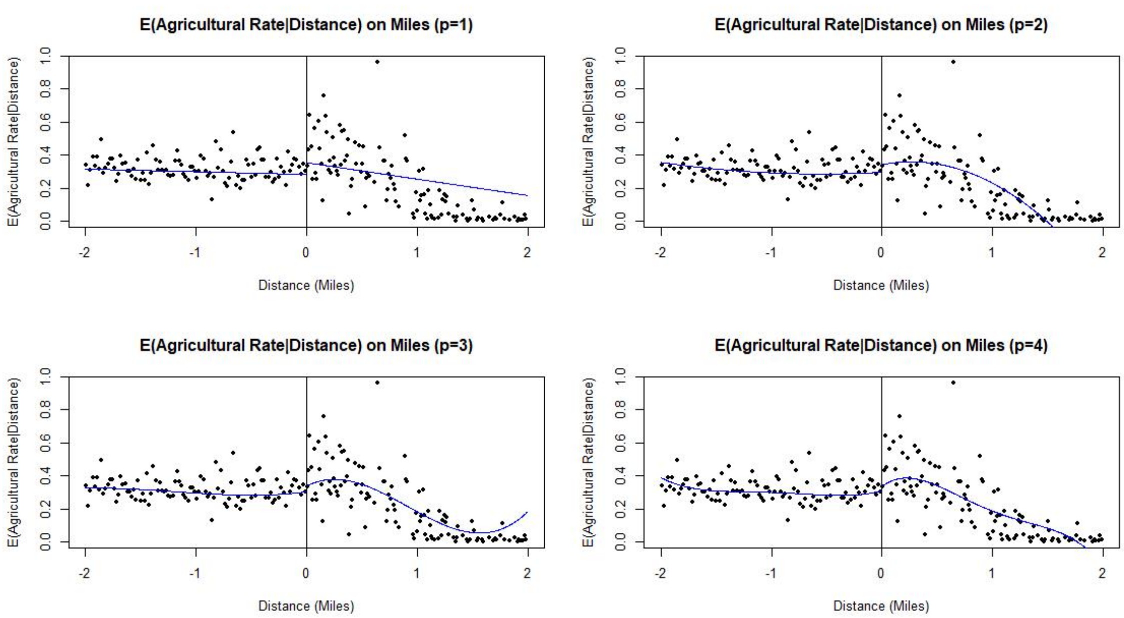


Figure A.2. RD plots of agricultural rate with 1st, 2nd, 3rd, and 4th order polynomial using 1905 Allotment Boundary using WRL dataset

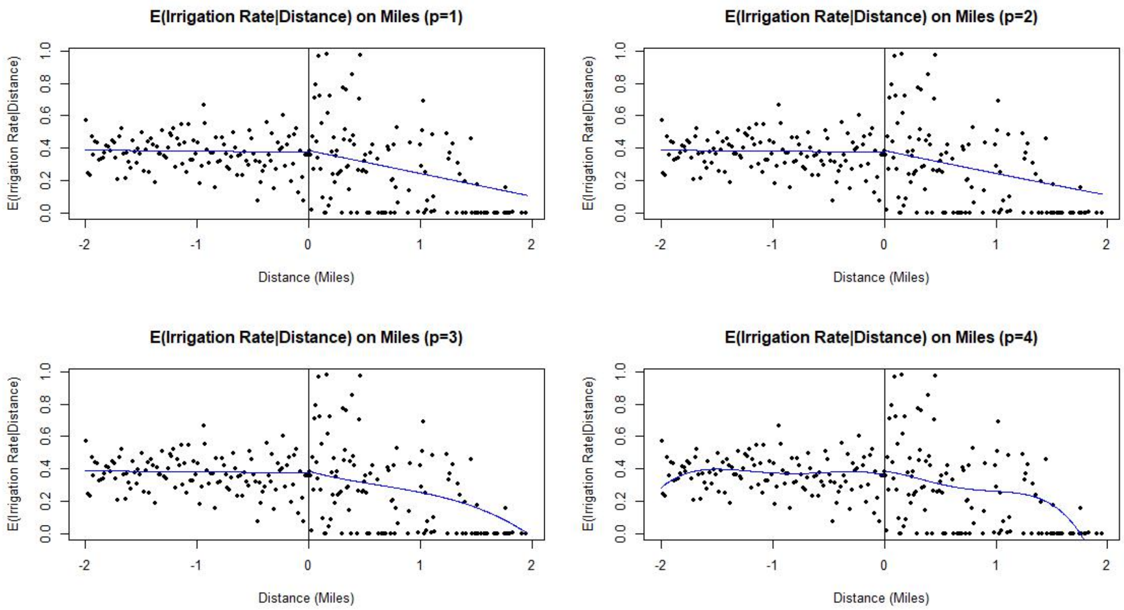


Figure A.3. RD plots of irrigation rate with 1st, 2nd, 3rd, and 4th order polynomial using 1905 Allotment Boundary

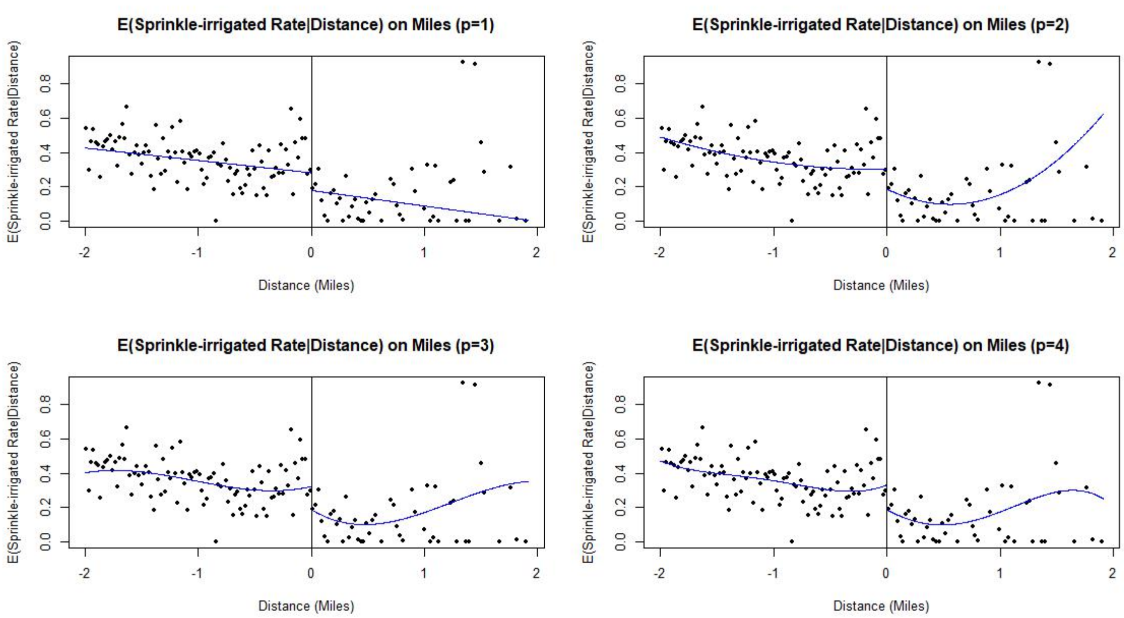


Figure A.4. RD plots of sprinkler-irrigated rate with 1st, 2nd, 3rd, and 4th order polynomial using 1905 Allotment Bundary

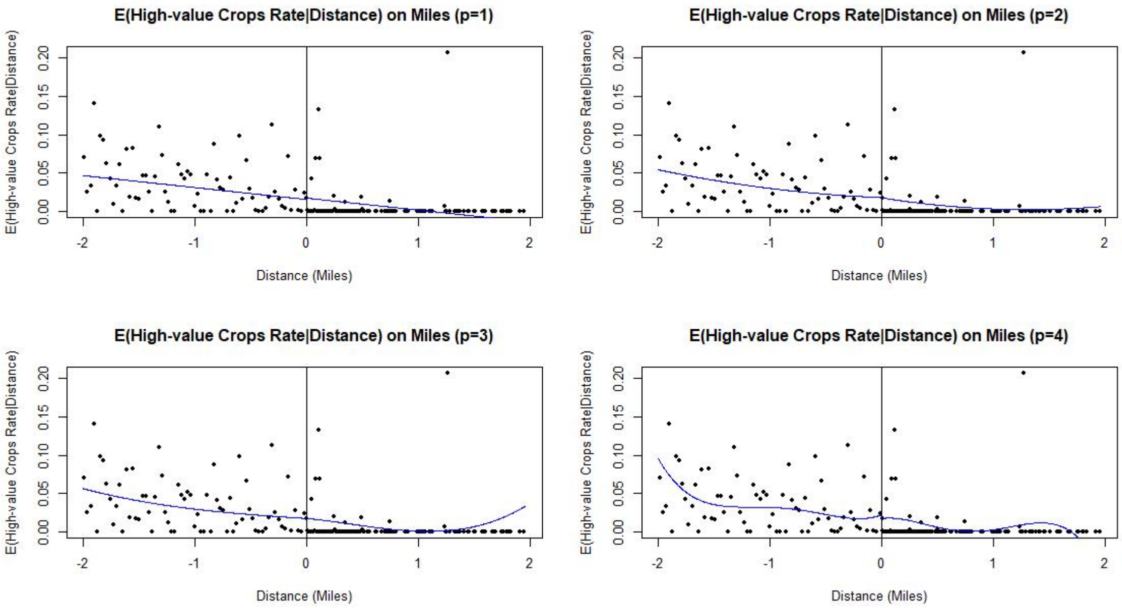


Figure A.5. **RD plots of high-value crops rate with 1st, 2nd, 3rd, and 4th order polynomial using 1905 Allotment Boundary using WRL dataset**

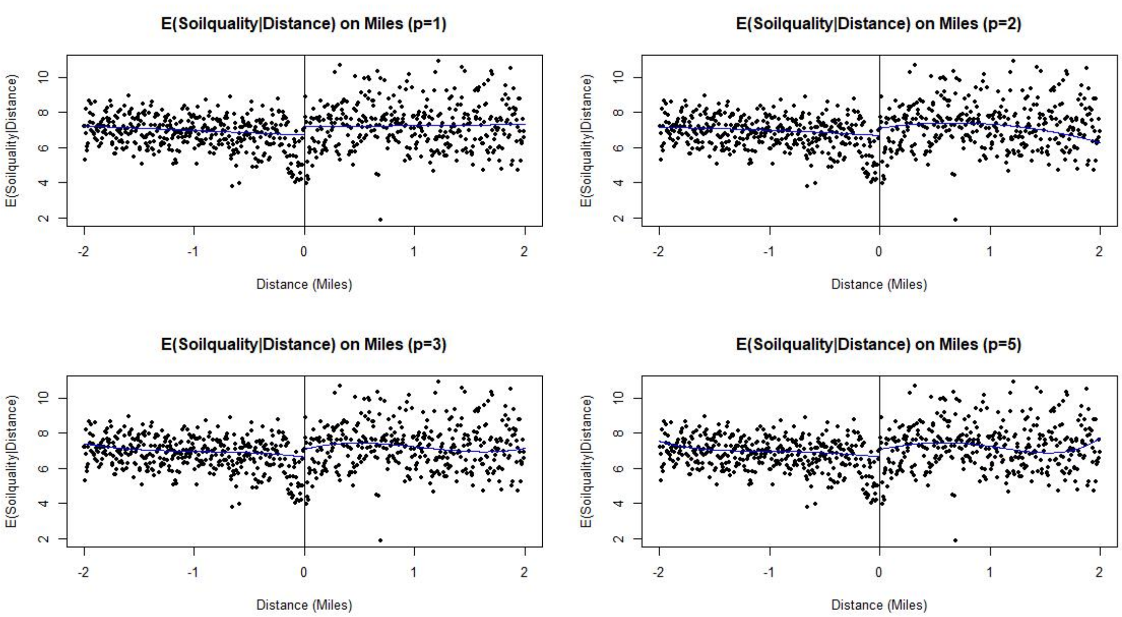


Figure A.6. **RD plots of soil quality with 1st, 2nd, 3rd, and 4th order polynomial using 2017 Tribal Boundary**

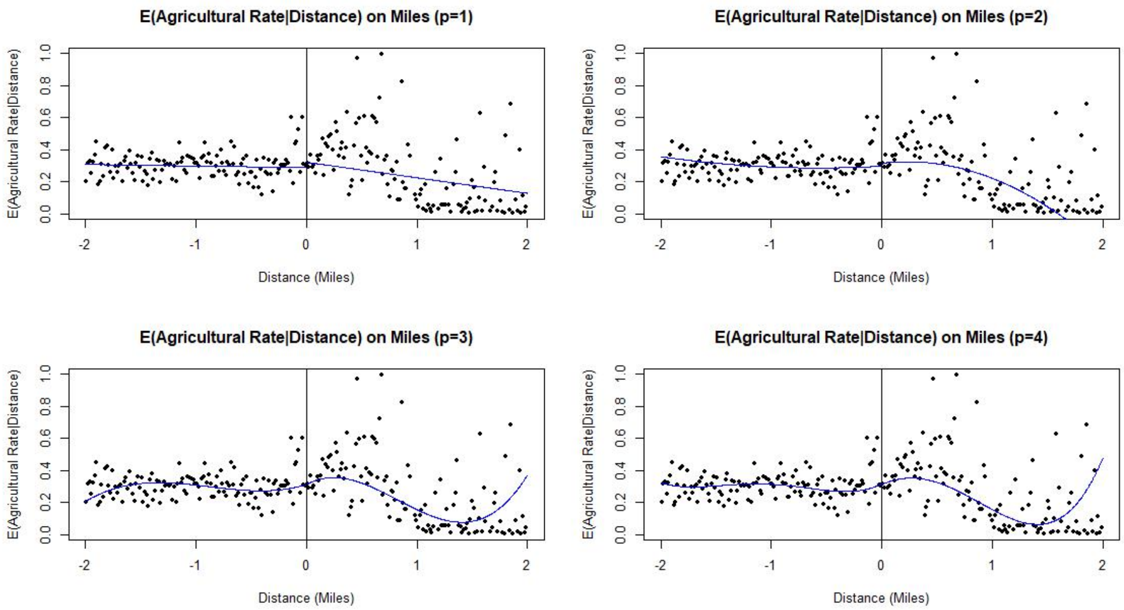


Figure A.7. RD plots of agricultural rate with 1st, 2nd, 3rd, and 4th order polynomial using 2017 Tribal Boundary using WRL dataset

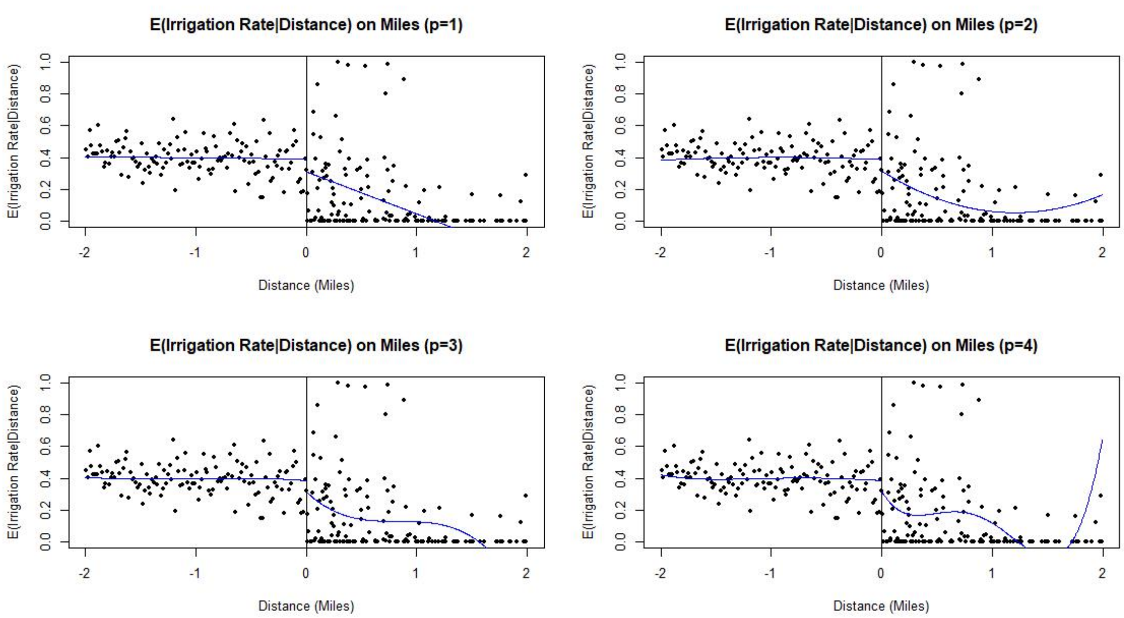


Figure A.8. RD plots of irrigation rate with 1st, 2nd, 3rd, and 4th order polynomial using 2017 Tribal Boundary

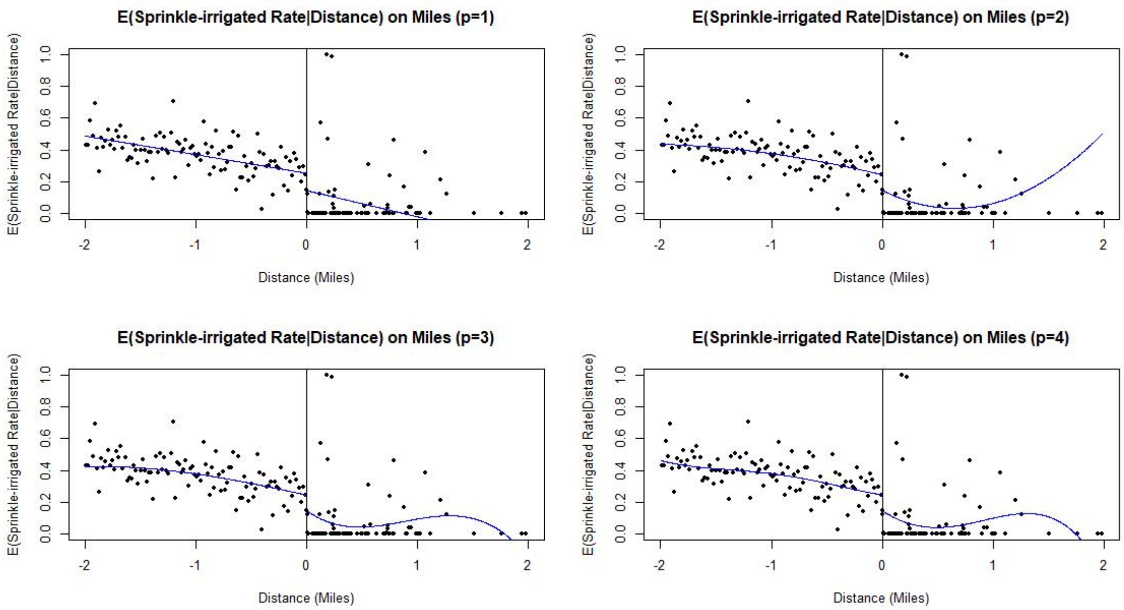


Figure A.9. RD plots of sprinkler-irrigated rate with 1st, 2nd, 3rd, and 4th order polynomial using 2017 Tribal Boundary

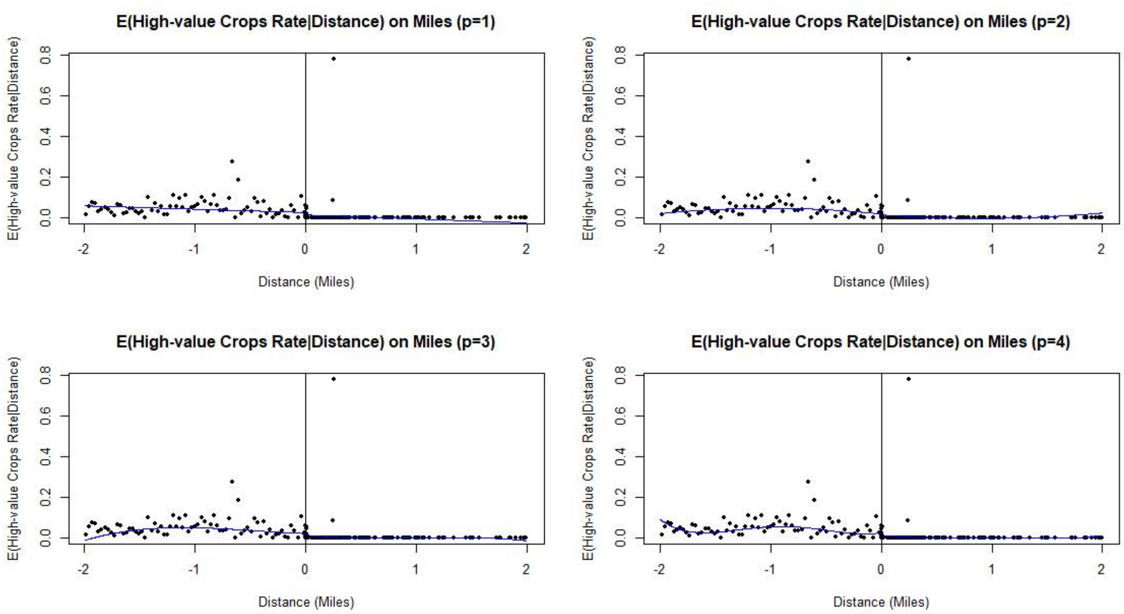


Figure A.10. RD plots of high-value crops rate with 1st, 2nd, 3rd, and 4th order polynomial using 2017 Tribal Boundary using WRL dataset

APPENDIX B

Chapter 3 Appendix

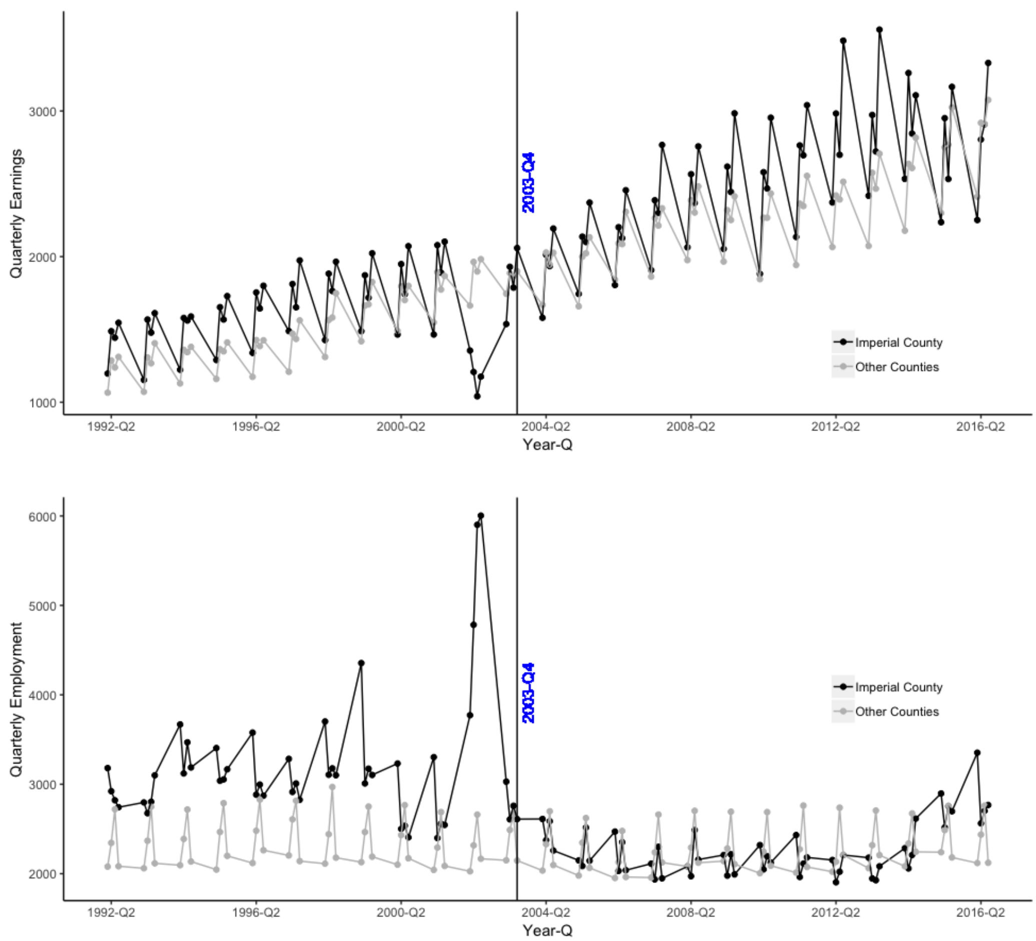


Figure B.1. Quarterly labor statistics (before imputation)
Note: Top panel represents the quarterly earnings and the bottom panel represents quarterly employment. The vertical line represents the QSA effected quarter, the fourth quarter of 2003.

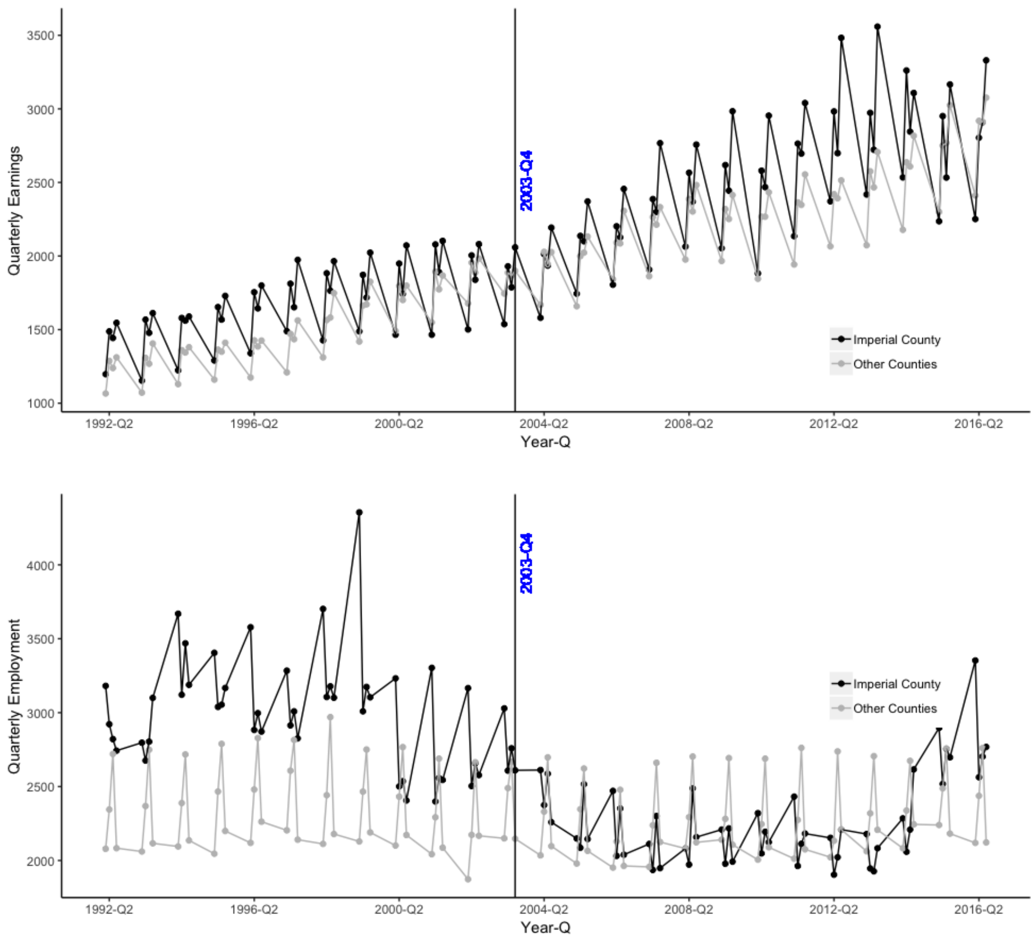


Figure B.2. Quarterly labor statistics (after imputation)

Note: Top panel represents the quarterly earnings and the bottom panel represents quarterly employment. The vertical line represents the QSA effected quarter, the fourth quarter of 2003. We notice an unusual spike of both earnings and employment in 2002 but it goes back to usual in the following year. This indicates a data issue of the 2002 labor data. We use the quarterly data of 2001 and 2003 to impute the labor data of 2002.

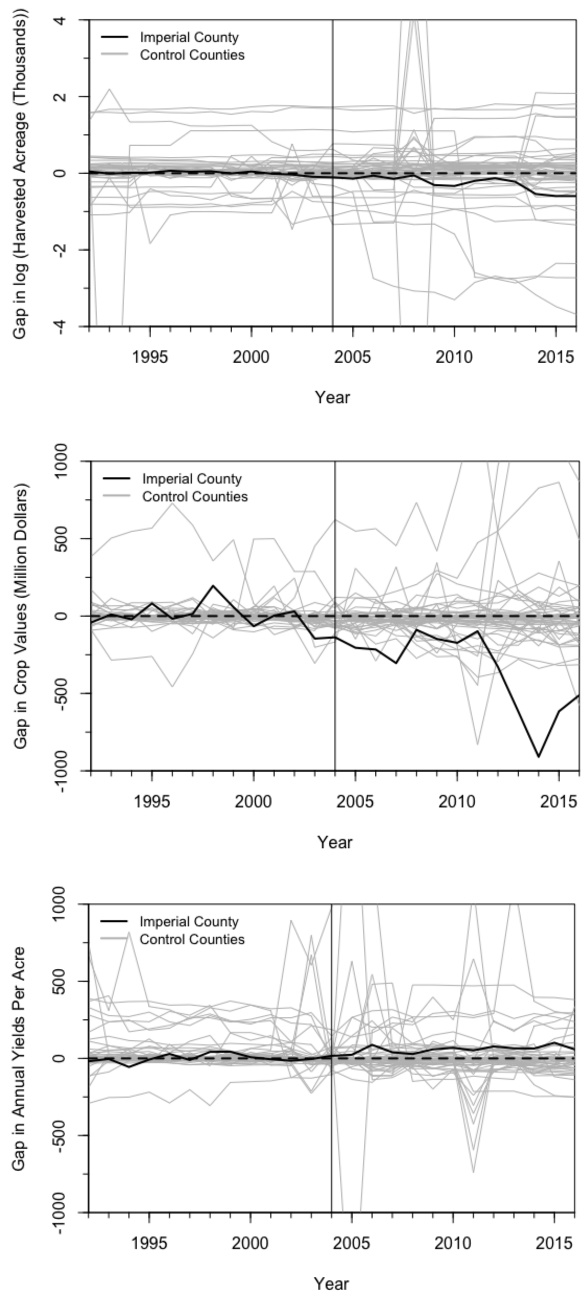


Figure B.3. Placebo test plots in crop production statistics
Note: The crop value is in millions of dollars. The vertical line represents the QSA effective year, 2004.

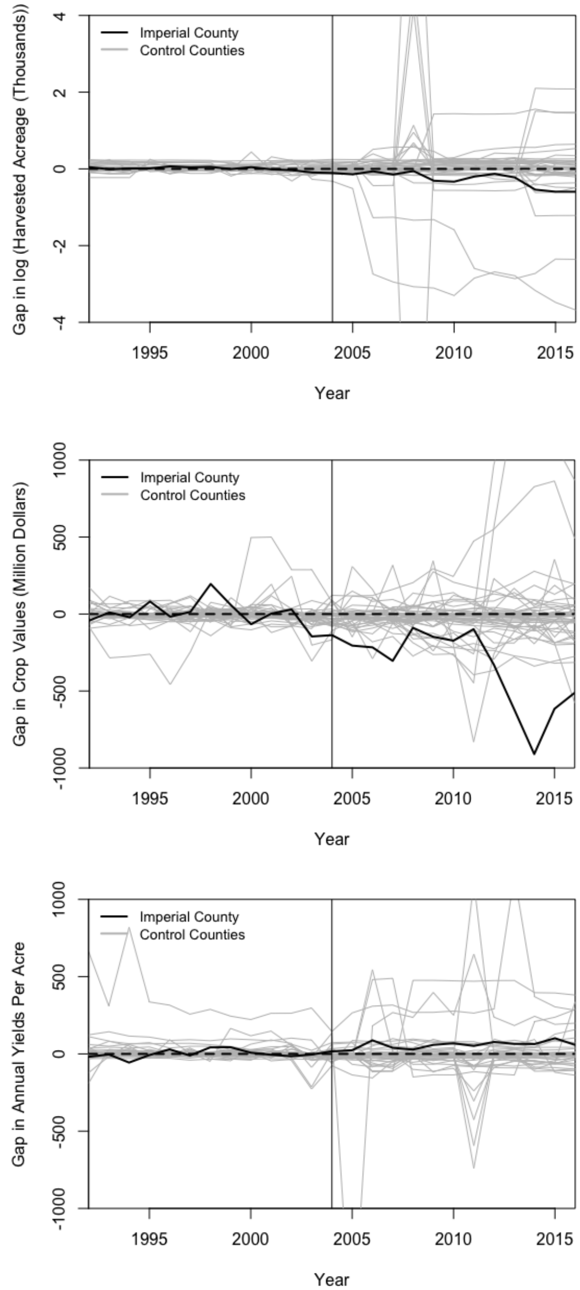


Figure B.4. Restricted sample size placebo test plots crop production statistics (20 times)

Note: The county with pre-QSA MSPE greater than 20 times Imperial County pre-QSA MSPE are excluded from the sample. The crop value is in millions of dollars. The vertical line represents the QSA effective year, 2004.

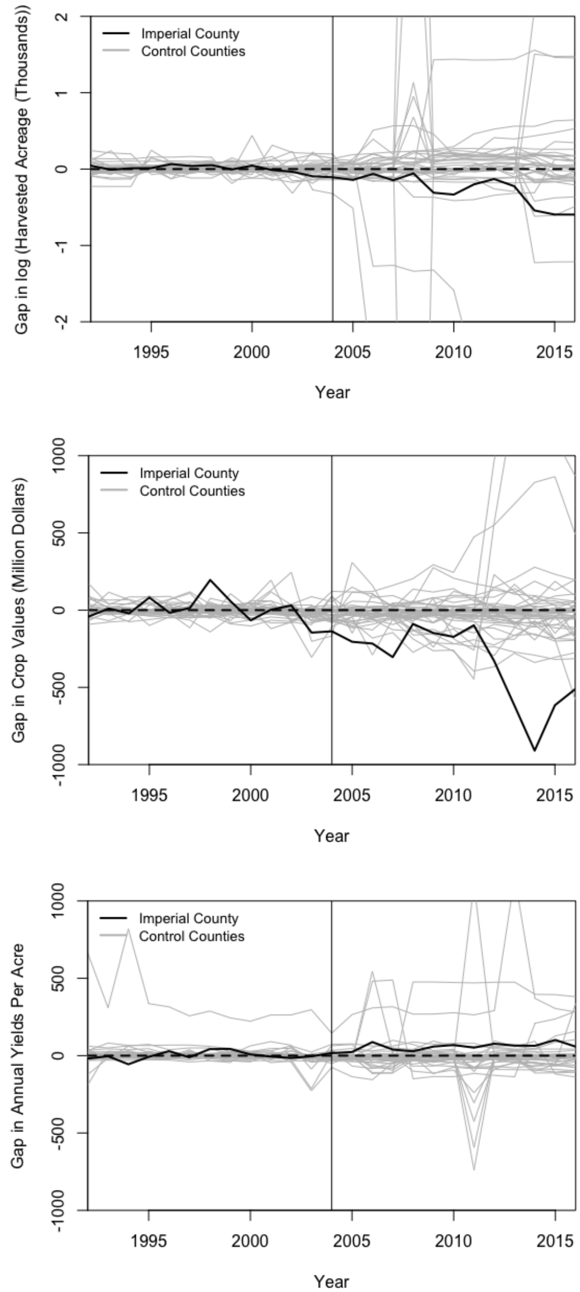


Figure B.5. Restricted sample size placebo test plots in crop production statistics (10 times)

*Note:*The county with pre-QSA MSPE greater than 10 times Imperial County pre-QSA MSPE are excluded from the sample. The crop value is in millions of dollars. The vertical line represents the QSA effective year, 2004.

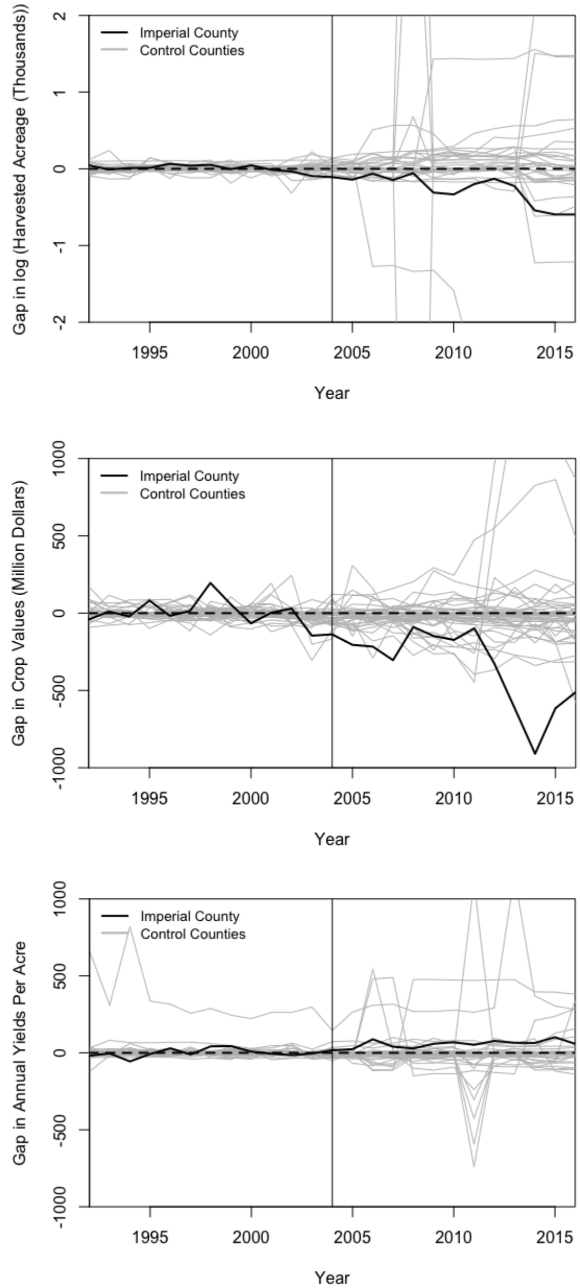


Figure B.6. Restricted sample size placebo test plots in crop production statistics (5 times)

Note: The county with pre-QSA MSPE greater than 5 times Imperial County pre-QSA MSPE are excluded from the sample. The crop value is in millions of dollars. The vertical line represents the QSA effective year, 2004.

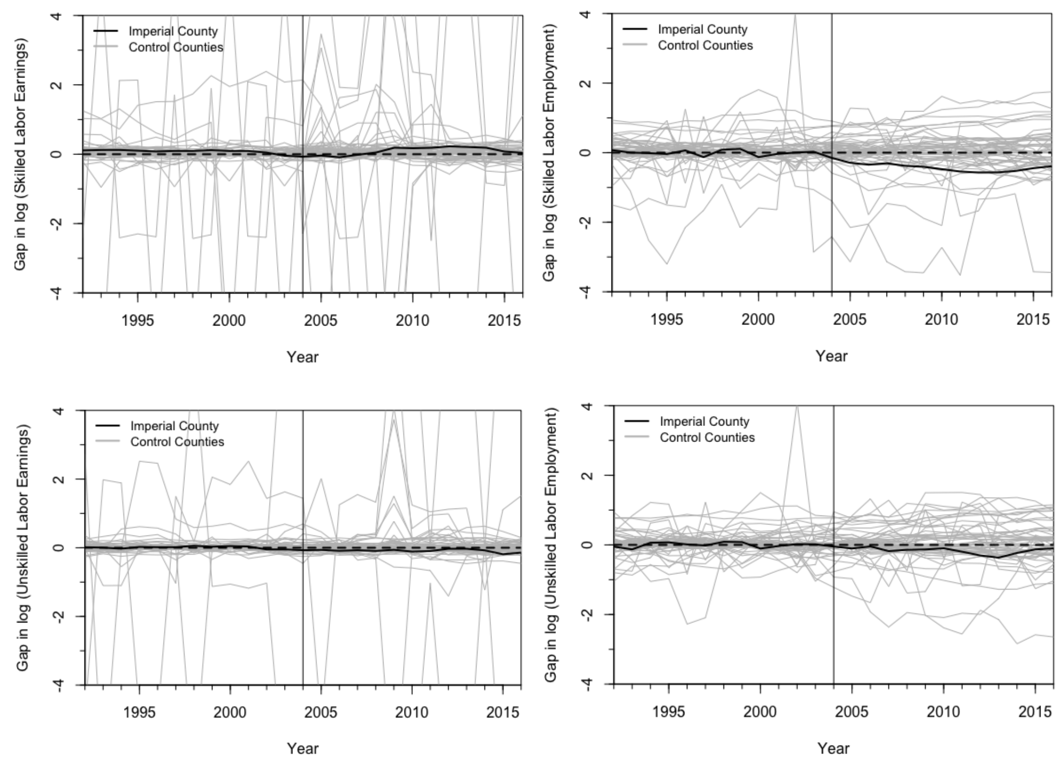


Figure B.7. Placebo test plots in labor statistics in crop production sector

Note: All the employment and earnings data are calculated by the quarterly average. Earnings is in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

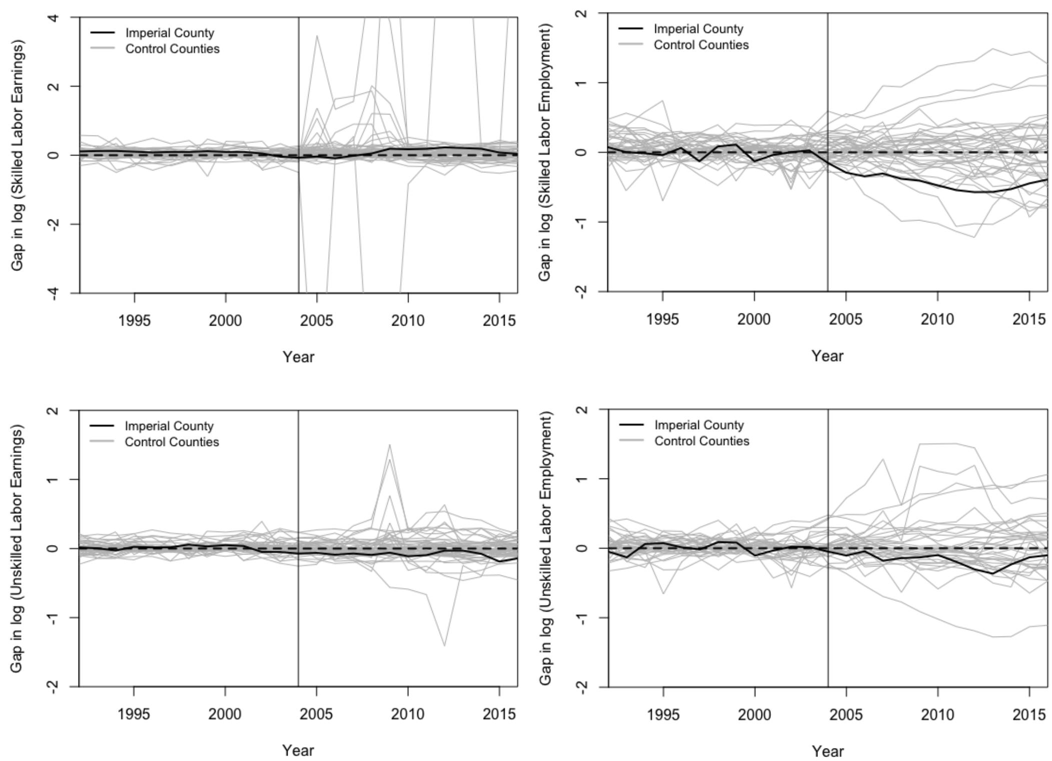


Figure B.8. Restricted sample size placebo test plots in labor statistics in crop production sector(20 times)

Note: The county with pre-QSA MSPE greater than 20 times Imperial County pre-QSA MSPE are excluded from the sample. All the employment and earnings data are calculated by the quarterly average. Earnings is in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

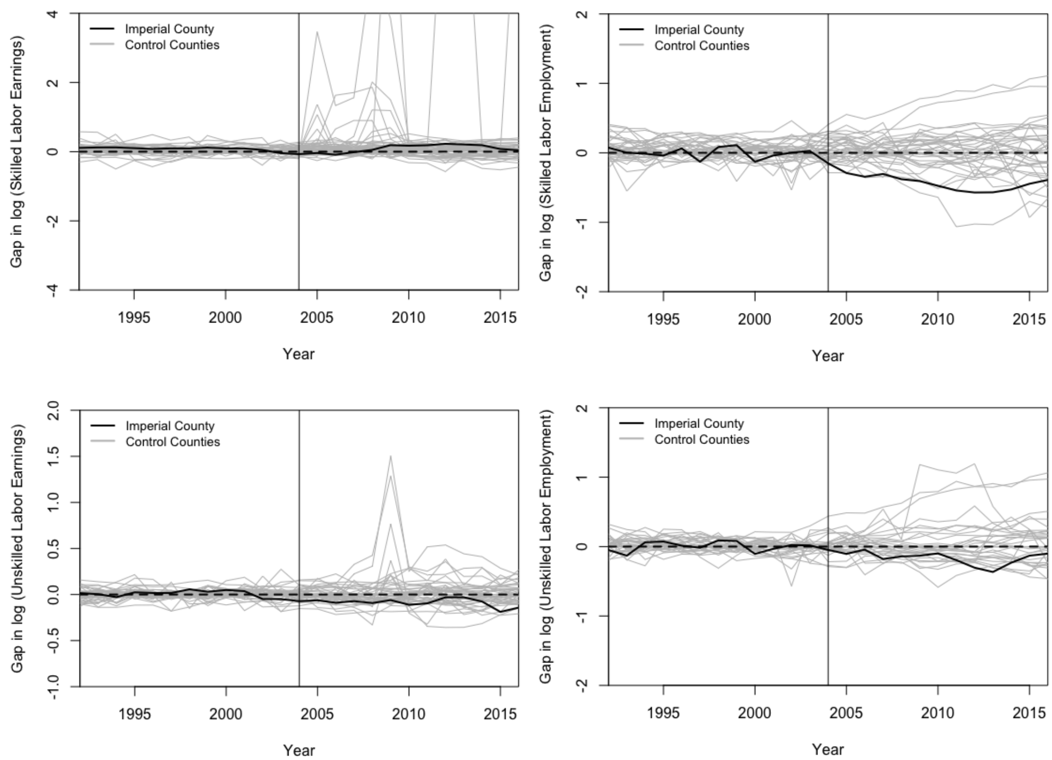


Figure B.9. Restricted sample size placebo test plots in labor statistics in crop production sector (10 times)

Note: The county with pre-QSA MSPE greater than 10 times Imperial County pre-QSA MSPE are excluded from the sample. All the employment and earnings data are calculated by the quarterly average. Earnings is in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

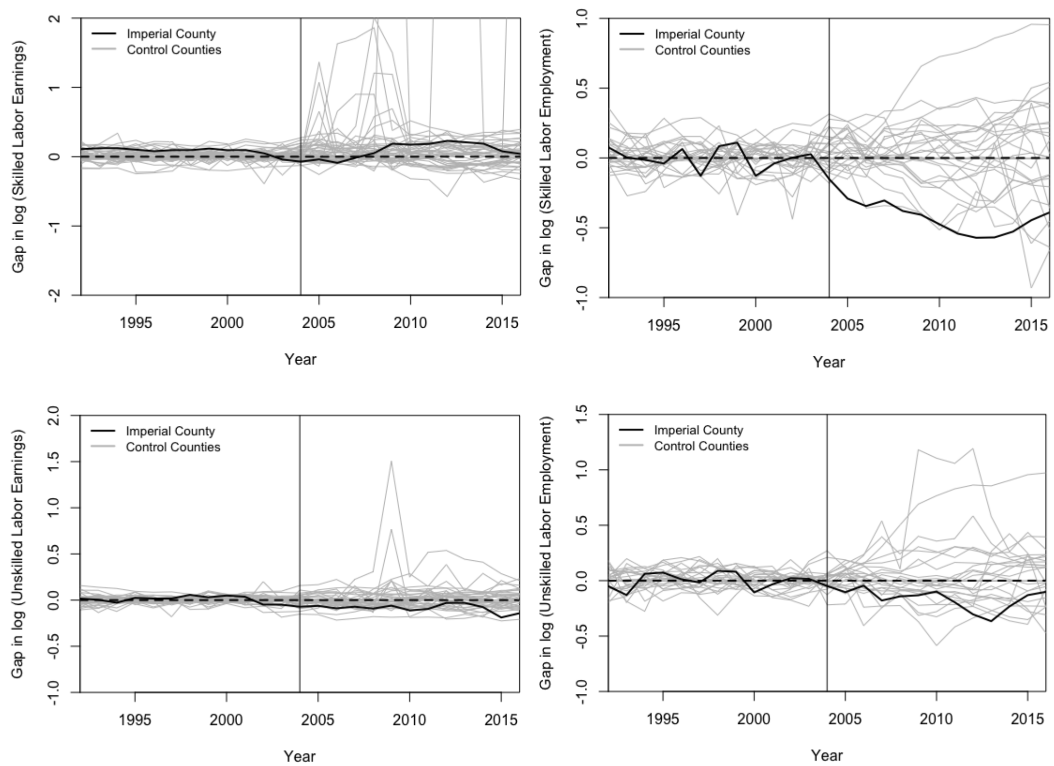


Figure B.10. Restricted sample size placebo test plots in labor statistics in crop production sector (5 times)

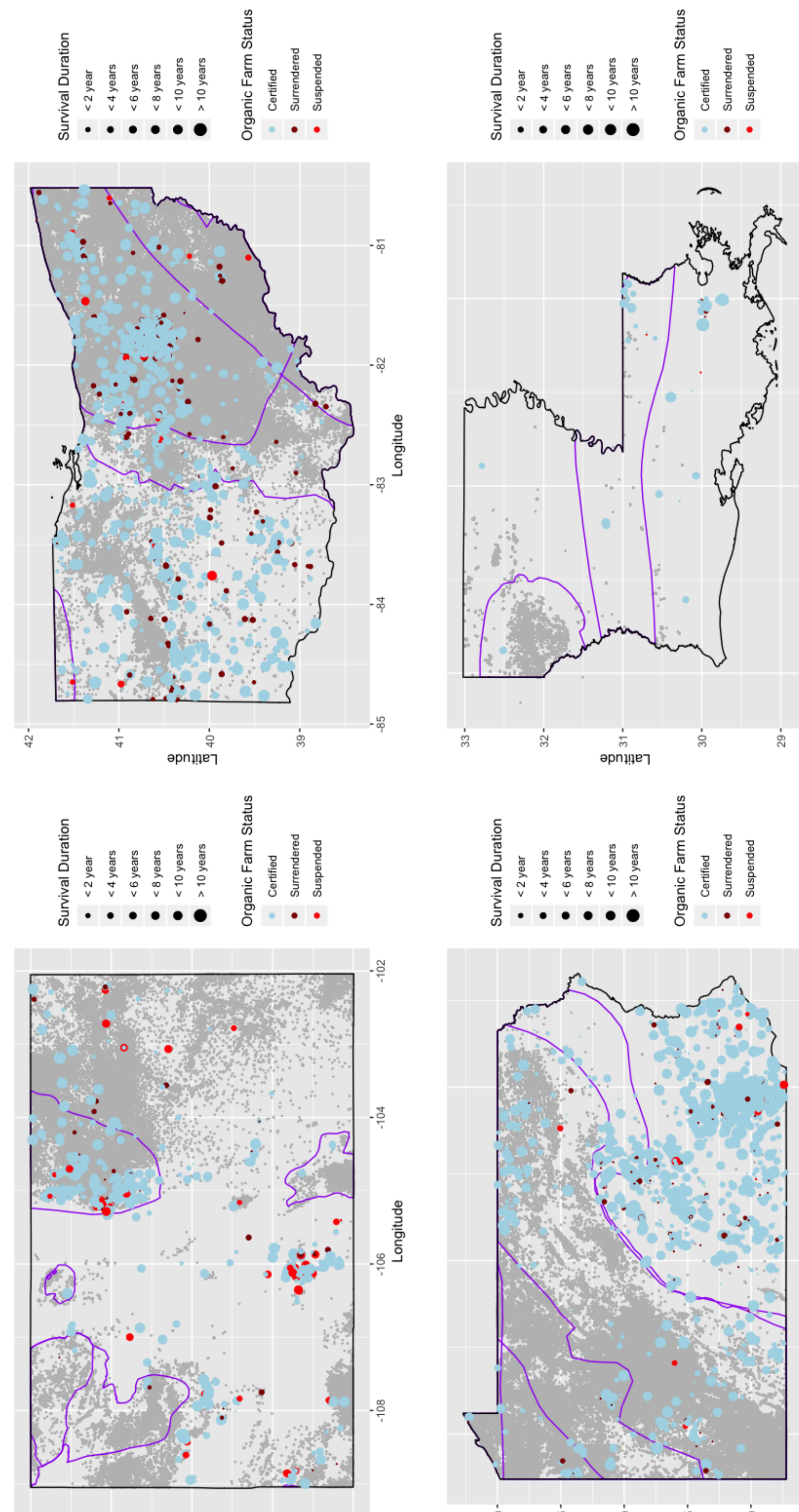
Note: The county with pre-QSA MSPE greater than 5 times Imperial County pre-QSA MSPE are excluded from the sample. All the employment and earnings data are calculated by the quarterly average. Earnings is in dollars and Employment is in the number of jobs. The vertical line represents the QSA effective year, 2004.

APPENDIX C

Chapter 4 Appendix

Table C.1. Likelihood ratio test

	Df	Sum of Sq	RSS	AIC	Pr(>Chi)	
1-mile			1.98E+04	1,963.37		
play	1	753.64	2.05E+04	1,981.75	6.34E-06	***
aquifer	1	15.93	1.98E+04	1,961.81	0.51	
3-mile			1.56E+06	4,345.06		
play	1	7.46E+04	1.64E+06	4,368.46	4.66E-07	***
aquifer	1	1,542.57	1.56E+06	4,343.60	0.46	
5-mile			1.05E+07	5,384.44		
play	1	5.82E+05	1.11E+07	5,411.75	6.16E-08	***
aquifer	1	6,016.74	1.05E+07	5,382.75	0.58	
10-mile			1.45E+08	6,812.00		
play	1	1.15E+07	1.56E+08	6,851.76	1.03E-10	***
aquifer	1	7.56E+04	1.45E+08	6,810.28	0.59	


Survival Duration
< 2 year
< 4 years
< 6 years
< 8 years
< 10 years
> 10 years
Organic Farm Status
Certified
Surrendered
Suspended
Longitude
Latitude

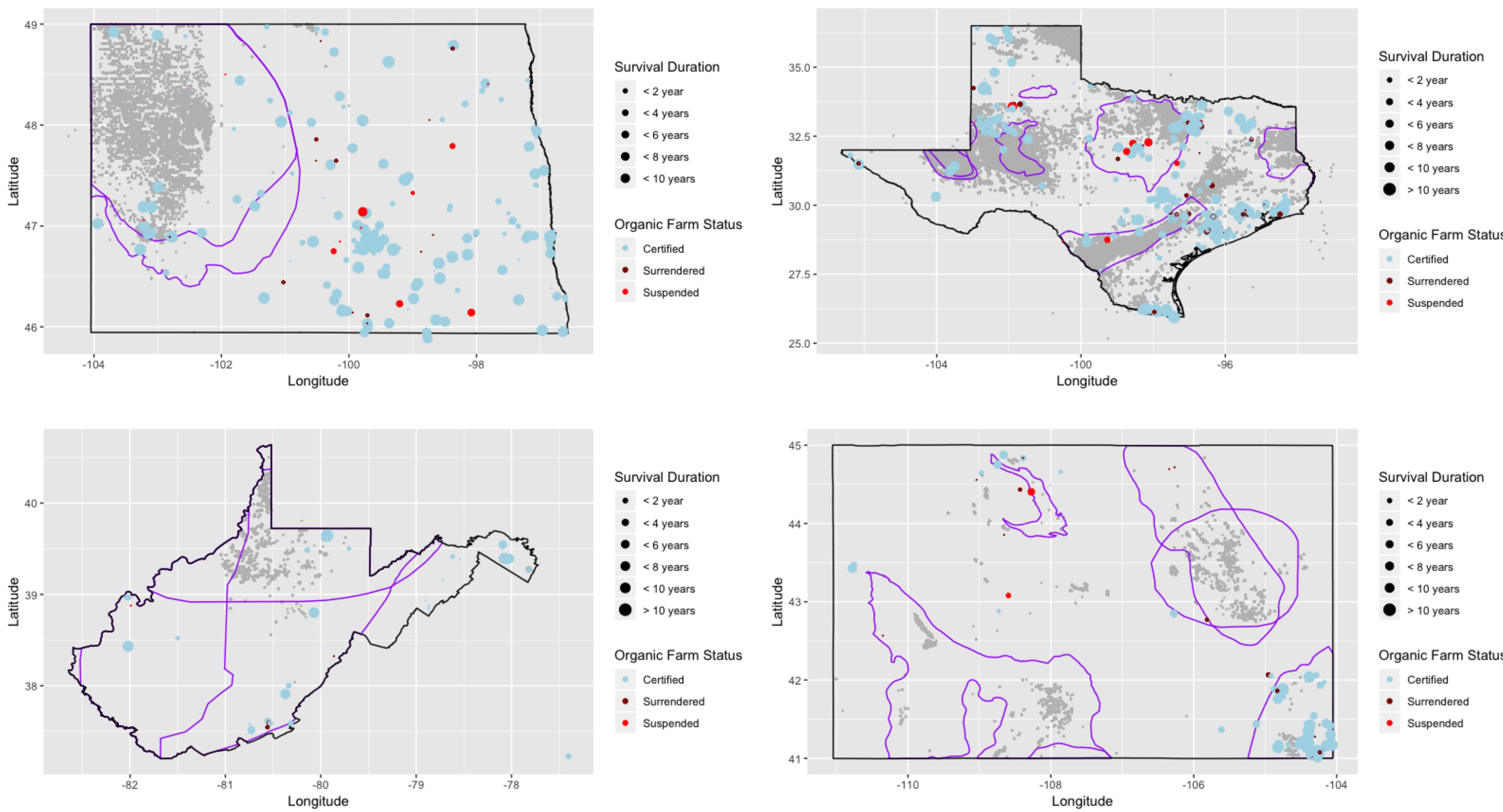


Figure C.1. Location map of eight shale states
Note: Blue dots represent certified organic farms, while dark red and bright red dots represent uncertified organic farms. Purple line represents the shale plays. Panel (A) to (H) represent Colorado, Ohio, Pennsylvania, Louisiana, North Dakota, Texas, West Virginia and Wyoming separately. Fracking wells data is from FracFocus: http://fracfocus.org/data-download.

CURRICULUM VITAE

MUYANG GE

Personal Cell: +1 217-819-7749
muyang.ge@aggiemail.usu.edu
https://sites.google.com/site/gemuyang/

EDUCATION AND TRAINING

Ph.D., Economics — May 2019(Expected)
Utah State University (USU)

M.S, Economics — December 2014
University of Illinois at Urbana-Champaign (UIUC)

B.S, Environmental Science — June 2013
Xiamen University (XMU)

Training:
NSF Traineeship Program, Climate Adaptation Science (CAS) — December 2018
http://climateadaptation.usu.edu/

RESEARCH INTERESTS

Environmental and Resource Economics
Applied Microeconomics
Climate Adaptation

PUBLICATION

Prudencio, Liana; Choi, Ryan; Esplin, Emily; Ge, Muyang; Gillard, Natalie; Haight, Jeffrey; Belmont, Patrick; Flint, Courtney. 2018. ”The Impacts of Wildfire Characteristics and Employment on the Adaptive Management Strategies in the Intermountain West.” *Fire* 1, no. 3: 46.
doi: 10.3390/fire1030046

RESEARCH PAPER

Ge M., Edwards, E.C., and Akhundjanov, S.B. 2018. Land Ownership and Irrigation on American Indian Reservations. (Revision requested at *American Journal of Agricultural Economics*)

WORK IN PROGRESS

Ge M. 2018. Organic farming in shale states: An instrumental variable estimation in survival analysis context

Regional water trade: the benefits of water market (With Eric C. Edwards, Reza Oladi and Dong-hun Go)

TEACHING EXPERIENCE

Instructor, Department of Applied Economics, USU — Fall 2018
APEC 3012: Natural Resource Economics

RESEARCH EXPERIENCE

Research Intern Summer 2017, Summer 2018
U.S. Geological Survey Moab, UT

- Worked with a team of experienced ecologists and soil scientists to understand social-ecological systems of the Colorado Plateau
- Examined the socioeconomic impact resulting from advanced shale development technology, fracking, using geospatial analysis and instrumental variable estimation in a survival analysis in the state of Colorado

Researcher January 2017 - present
Climate Adaptation Science Program, USU Logan, UT

Projects:
Assessing Fire Trends, Economic Effects, and Adaptive Management Strategies in the Intermountain West

- Work in an interdisciplinary team of academic and non-academic researchers
- Built Event Study model studying socioeconomic impacts of urban and rural fire on local labor market within the Intermountain West
- Communicated with non-scientist audiences such as policy-makers and shareholders about the broader impacts of climate change

Research Assistant August 2015 - present
Department of Applied Economics, USU Logan, UT

Projects:
Agriculture, Water, and Climate Response on American Indian Reservation (USDA-funded)
iUTAH-innovative Urban Transitions and Aridregion Hydro-sustainability (NSF-funded)

- Developed general-equilibrium representation of a simple coupled hydrologic-ecological-economic framework to explore the economic and environmental performance of a large–scale water transfer agreement signed in California
- Developed sythentic control methodology to test the trade theory in regional setting empirically
- Digitized maps and analyzed spatial data using GIS and R
- Developed Regression Discontinuity model to explain lagging agricultural development on American Reservations by examining the cross-border effects of Uintah-Ouray Indian Reservation boundary on agricultural irrigation

Undergraduate Researcher December 2011 - May 2013
Department of Environmental Science, XMU Xiamen, China

Project:
Effects of N_2O Emission on the River Reservoir Eutrophication, National Innovation Fund (NIF-funded)

- Wrote grant proposal to NIF and got fully funded

PRESENTATIONS AT PROFESSIONAL MEETINGS

Prudencio, L., Choi, R., Esplin, E., Ge, M., Gillard, N., Haight, J., Belmont, P., Flint, C. The Impacts of Wildfire Characteristics and Employment on the Adaptive Management Strategies in the Inter-mountain West. International Symposium on Society and Resource Management, Snowbird, Utah, June 18, 2018.

Ge M., Edwards, E.C., and Akhundjanov, S.B. Land Ownership and Irrigation on American Indian Reservations. Heartland Environmental and Resource Economics Workshop, Urbana, Illinois, September 29-30, 2018. - Poster

COMPETITIVE GRANT

National Innovation Fund, China 2011 - 2013

Projects: "Effects of N_2O Emission on the River Reservoir Eutrophication", PI, $5,000

HONORS AND SCHOLARSHIPS

Don & Ming Wang Graduate Fellowship Scholarship, 2018 - 19
USU Graduate Travel Award, 2018
Co Bank Economics Fellowship, 2017 - 18
Top 10 Percent Scholarship Awards, 2012 - 13

RELEVANT SKILL SETS

Programming Skills:
Proficient in: R, Stata, GIS, LaTeX; Also basic ability with: Python, GAMS. Power user of Microsoft packages (Word, PowerPoint, and Excel)

Language Skills: English (fluent), Chinese (native)

MUYANG GE
Personal Cell: +1 217-819-7749
gemuyang@gmail.com
https://sites.google.com/site/gemuyang/

REFERENCES

Professor Eric C. Edwards
Major Professor, USU
+1-208-301-0247
eric.edwards@ncsu.edu

Professor Ryan Bosworth
Committee Member, USU
+1-435-797-0594
ryan.bosworth@gmail.com

Professor Reza Oladi
Committee Member, USU
+1-435-797-8196
reza.oladi@usu.edu

Professor Courtney G Flint
CAS Program Supervisor, USU
+1-435-797-8635
courtney.flint@usu.edu