2017

Three essays on environmental and resource economics

Kevin Michael Meyer
Iowa State University

Three essays on environmental and resource economics

by

Kevin Michael Meyer

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Catherine L. Kling, Co-major Professor
David Keiser, Co-major Professor
Quinn Weninger
John Beghin
GianCarlo Moschini

Iowa State University

Ames, Iowa

2017

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vi

ACKNOWLEDGMENTS vii

ABSTRACT viii

CHAPTER 1. ADAPTING TO CLIMATE CHANGE THROUGH TILE DRAINAGE: EVIDENCE FROM MICRO LEVEL DATA 1

Abstract 1

Introduction 2

Conceptual Model 5

Data 8

Land Rents 8

Tile Drainage 8

Climate 9

Soil and Demographics 10

Summary Statistics 11

Empirical Model 11

Regression Estimates 13

Climate Change Simulations 16

Damage Estimates 19

Conclusion 21

CHAPTER 2. THE IMPACT OF LAND USE ON LAKE WATER QUALITY: EVIDENCE FROM IOWA 24

Abstract 24

Introduction 25

Background 26

Data 30

Water Quality Model 32

Cropland Response Model 34

Estimation 34

Results 36

Application: The Renewable Fuel Standard 38

Conclusion 39

CHAPTER 3. OPTIMAL DEER MANAGEMENT IN THE MIDWEST . . 41

Abstract 41

Introduction 42

Model 45

Regulation and License Purchase 47

Deer Herd Benefits and Costs 47

Management goal 48

Functional Form 49

Application to Iowa 50

Numerical Results 53

Simulation 54

Conclusion 55

APPENDIX A. CHAPTER ONE APPENDIX 60

Functional Form 60

Tile Drainage: Historical Estimates 62

STATSGO Variables 63

Regression Results 64

APPENDIX B. CHAPTER TWO APPENDIX 72

Econometrics Criteria 72

BIBLIOGRAPHY 75

LIST OF TABLES

Table 1.1 Descriptive Statistics for Primary Data Set 12
Table 1.2 Regressions Results 15
Table 1.3 Summary Statistics for Climate Projections 18
Table 1.4 Predicted Impact of climate change on farmland values 20

Table 2.1 Summary Statistics 32
Table 2.2 Regression Results: Water Quality Model 37
Table 2.3 Regression Results: Cropland Model 38

Table 3.1 Iowa Parameters 52

Table A.1 STATSGO variables 63
Table A.2 Regressions Results: Quadratic Model 64
Table A.3 Regressions Results: Seasonal Model 66
Table A.4 Regressions Results: Linear Spline Model 68
Table A.5 Regressions Results: Step-Function Model 70

LIST OF FIGURES

Figure 1.1 U.S. tile drainage as a percentage of cropland 9

Figure 1.2 Estimated quadratic response of farmland values to precipitation . . . 16

Figure 1.3 Spatial distribution of climate change impacts, 2020-2049 21

Figure 1.4 Spatial distribution of the difference in climate change impacts between the Structural Ricardian model and the pooled model, 2020-2049. . . . 22

Figure 2.2 Iowa Lakes, HUC8, and HUC12 watersheds 31

Figure 2.3 Crop Price Index, 2001-2016 35

Figure 2.4 Estimated Change in CTSI due to the RFS 40

Figure 3.2 Iowa Land Uses Over Time 43

Figure 3.3 The economic value of deer hunting. Reprinted from Aiken (2006). . . 51

Figure 3.5 Numerical Approximation of the Value Function 57

Figure 3.7 Numerical Approximation of the Policy Functions 58

Figure 3.9 Simulation results 59

Figure A.1 Plot of Coefficients for Precipitation Using a Step-Function 61

ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere thanks to those who helped me with various aspects of conducting research and writing this thesis.

Thank you to all the people and organizations who helped fund my graduate research. Thanks to the USDA for supporting me through both the NIFA and NEEDS fellowship programs, and to the EPA for awarding me the NARS campus challenge scholarship. Thank you to Jeri Prescott for awarding me the James R. Prescott Scholarship. Finally, thanks to Iowa State's economics department for supporting me through teaching assistantships.

I would also like to thank my major professors, Dr. Catherine Kling and Dr. David Keiser. Through good advice, and donuts, you guided me through a difficult yet incredibly rewarding period of my life.

Thanks to Dr. Weninger, Dr. Moschini, and Dr. Beghin for being a part of my committee. All were integral to my success at Iowa State and in writing this dissertation.

Finally, thank you to my wife and daughter for lifting my spirits every day.

ABSTRACT

Environmental issues in modern day Iowa are a perfect example of the indirect consequences of exploiting natural resources. Over 12,000 years ago, glaciers left the state with rich and fertile soil, perfect for agriculture. Over the course of more than one hundred years, Iowa's landscape has been cleared and drained to gain access to this valuable farmland. While the economic benefits of agriculture are clear, it is important to understand the environmental consequences of this transformation. This dissertation uses economic tools and analysis to investigate three environmental and resource issues related to the complex interplay between Iowa agriculture and the environment.

The first chapter examines the relationship between an important adaptive tool, tile drainage, and climate. Tile drainage is largely responsible for transforming Iowa from mostly wetlands into prime farmland. It fundamentally changes the relationship between land, climate, and soil, by draining away excess water, allowing crops to grow. This chapter uses observations from over 800,000 farms across the U.S. to estimate the relationship between farmland value and climate while explicitly incorporating tile drainage. We find fundamental differences in the relationship between tile drained and non-tile drained land, which has not been accounted for in previous research. Using climate projections, we estimate the impact of climate change on farmland and show how these estimates can be biased when tile drained and non-tile drained farms are pooled together.

The second chapter looks at the relationship between land change and lake water quality. While most of Iowa's lakes are artificial, many are popular destinations for fishing, boating, swimming, and other recreational activities. But their close proximity to farmland results in high nutrient levels and decreased water quality, which can reduce recreational and ecosystem benefits. This chapter combines fifteen years of water quality measurements with satellite images of land use to estimate the impact of land use change on water quality. These estimates

are used to assess the lake water quality impacts of the Renewable Fuel Policy, a government policy which has had a large impact on agriculture and land use in Iowa.

The third chapter is concerned with the optimal management of the Iowa deer population through hunting licenses. Although not all species have benefited from the transformation of Iowa's landscape, the deer population has thrived due to a lack of predators and an abundant new food source in crops. While deer hunters enjoy a large population of deer, farmers and drivers face costs due to crop depredation and deer vehicle collisions, creating a complex management problem. This chapter uses the tools of dynamic programming to solve for an optimal policy that balances these opposing interests.

Altering the natural landscape turned Iowa into one of the most productive farming regions in the world, but has also created the need to balance intensive farming practices with the impacts on the surrounding environment. The tools of economics provide an appealing framework to propose solutions to these problems. The goal of the following three chapters is to use these tools to shed some light on three such issues Iowa currently faces. The insight and results from this research will hopefully help inform future researchers and policymakers in Iowa and beyond.

CHAPTER 1. ADAPTING TO CLIMATE CHANGE THROUGH TILE DRAINAGE: EVIDENCE FROM MICRO LEVEL DATA

Abstract

Climate change is expected to alter precipitation patterns in the U.S., requiring adaptation to mitigate damages in areas that recieve high levels of precipitation. Nearly all of the previous literature has focused on irrigation, with little attention given to tile drainage, the key adaptive technology to reduce excess water. Using a simple conceptual model, we show that the value of precipitation should differ between drained and non-drained land. Thus, pooling lands that are tile drained and those that are not could bias estimates of the effects of climate change on land values. We [1] test this hypothesis by estimating a Structural Ricardian model using farm level data from the 2012 Census of Agriculture. Consistent with our theoretical model, our estimates show significant differences in the value of precipitation across tile drained and non-tile drained land. We calculate damages using the most recent, spatially detailed climate simulations available. We find that pooled models underestimate damages in the cornbelt region by as much as 15%, and underestimate benefits in northern counties by as much as 20%.

JEL codes: Q10, Q15, Q51, Q54

Keywords: Climate change; adaptation; agriculture; climate impacts; tile drainage

[1] Co-Author: Dr. David A. Keiser, Iowa State University

"Look at this flower pot. What is the hole at the bottom for? I ask you, because there is a complete agricultural revolution in that hole." - (Klippart, 1861, p. 3)

Introduction

Designing efficient policies to address climate change requires accurate damage estimates. Recent studies that quantify the value of climate to agriculture have provided key insights in this area (Fisher et al., 2012; Deschenes and Greenstone, 2007; Schlenker et al., 2006; Mendelsohn et al., 1994). However, much uncertainty still remains regarding the magnitude and distribution of expected damages. One critical piece to this puzzle is the role of adaptation (Burke et al., 2016; Auffhammer and Schlenker, 2014). Specifically, farmers may invest in technologies that mitigate the harmful effects of changes in climate. Understanding this adaptive behavior is fundamental to improving damage estimates.

In this paper we study a key adaptive technology available to farmers - subsurface drainage. This type of drainage, also known as tile drainage, reduces excess water stress in crops by lowering the water table, allowing rainfall to move more quickly through poorly drained soils.[2] Tile drainage was first introduced in the U.S. in 1835 and quickly experienced wide-spread adoption (Pavelis, 1987). From 1855 to 1985, approximately 43 billion dollars was invested in tile drainage infrastructure (Pavelis, 1987).[3] Today, nearly 48 million acres utilize tile drainage; this land represents approximately a quarter of the total cropland value in the country. Many have argued that U.S. agriculture as we know it today would not exist without this critical piece of infrastructure (Jaynes and James, 2007; Pavelis, 1987).

To assess the importance of this adaptation tool, we develop an economic model of drainage adoption and use newly available farm level data on the location of tile drainage from the 2012 Census of Agriculture to estimate the effects of climate on farmland values on tile drained and non-tile drained lands. To our knowledge, only Fezzi and Bateman (2007), Schlenker et al. (2007), and Kurukulasuriya et al. (2011) have used farm level data in a Ricardian analysis; these studies were based in Germany, California, and Africa, respectively. Combined with zip

[2] Subsurface drainage is often referred to as tile drainage because early drains were constructed from clay tiles. However, most modern drains are constructed from corrugated plastic.

[3] All dollar amounts represent 2012 dollars.

code level climate data, our use of micro-level data helps alleviate concerns of aggregation bias that has been raised by others (Fezzi and Bateman, 2007).

Our conceptual model illustrates that marginal increases in precipitation on tile drained land are less valuable than on non-tile drained land. Intuitively, a farm that installs tile drainage already has a sufficient if not excess supply of water for crops, so that the marginal value of an increase in precipitation is low. Conversely, a farm that does not install tile drainage does not have an excess supply of water and likely profits from increases in precipitation.

To test our model's predictions, we employ a Structural Ricardian model. The original Ricardian method, developed by Mendelsohn et al. (1994), utilizes a cross-sectional regression of farmland values on climate and control variables to recover implicit values for marginal changes in long-run averages of precipitation and temperature. The key insight of this method is that long-run climate effects should be capitalized into land values. The *Structural* Ricardian model, developed in Kurukulasuriya et al. (2008) and Seo and Mendelsohn (2008), builds on the original Ricardian model by explicitly modeling the adaptive choices of farmers. In the first stage, we estimate the probability of a farm investing in tile drainage as a function of exogenous land characteristics and climate. In the second stage, we estimate separate Ricardian functions for tile drained and non-tile drained farms. We use our estimates from both stages along with detailed, spatially explicit climate change simulations to estimate expected farmland damages from future climate change. The resulting damage estimates capture changes in farmland values as well as changes in the probability of being on tile drained land.

We find strong evidence that the marginal value of precipitation is higher on non-tile drained land. The result is robust to using a variety of functional forms for precipitation. Combined with differences in other climate variables, all results imply that pooling tiled and non-tiled land will bias the climate coefficients, and therefore expected climate change impacts. The differences are most noticeable for the coefficients on precipitation and harmful degree days. We compare our damage estimates with pooled estimates and find that, on average, damages from the Structural Ricardian model are higher than in the pooled model. While adaptation will generally lower damage estimates from climate change, we attribute the increase in damages to misspeficiation bias. That is, prior models that pool tile drained and non-tile drained farms are fundamentally misspecified and do not represent a clear baseline for comparison. There are also important geographic patterns in the bias. The potential damages from climate change are underestimated by as much as 15% in the Midwest, while the potential benefits are underestimated by as much as 20% in the northernmost counties.

Our paper builds on a strand of literature which investigates the role of adaptation in Ricardian models. To date, this work has focused almost exclusively on irrigation. As observed by Schlenker et al. (2005), a key assumption of the Ricardian method is that the coefficient on precipitation measures the supply of water for crops. In the Western half of the United States, however, much of the supply of water for crops is obtained through irrigation. Schlenker et al. (2005) show that Ricardian functions are fundamentally different on irrigated and non-irrigated land. We build on this idea by observing that the supply of water for crops on tile drained land is also not equal to precipitation, since excess water is drained away. This implies there is a fundamental difference in tile drained and non-tile drained land as well. While some studies have focused on non-irrigated land, we know of no other Ricardian analysis which accounts for tile drainage, including those based in countries with heavily drained areas such as the U.S. (Burke and Emerick, 2016; Deschenes and Greenstone, 2007; Schlenker et al., 2006; Mendelsohn et al., 1994), Canada (Reinsborough, 2003), Germany (Chatzopoulos and Lippert, 2015), and the United Kingdom (Fezzi and Bateman, 2007).[4] This paper also contributes to more recent Ricardian work that has tried to improve how econometricians properly account for the supply of water to crops (Hendricks, 2016).

The paper proceeds as follows. Section 2 builds a simple conceptual model of a representative farmer's optimal choice of tile drainage, which generates intuition for the empirical analysis. Section 3 provides a summary of the data, which includes a discussion on the availability and accuracy of the existing universe of tile drainage data. Sections 4 and 5 describes the econometric model used in this paper, followed by the coefficient estimates. Section 6 uses these coefficient estimates to simulate the effects of climate change on agriculture, and compares our results with previous estimates. Section 7 provides a brief conclusion.

[4]See Feick et al. (2005) for a digital global map of artificial drainage.

Conceptual Model

In this section, we present a simple economic model of tile drainage. Although our model uses profits as the outcome variable rather than land rents, profits can be converted into land rents using a capitalization ratio (Schlenker et al., 2006). Since Ricardian models rely on land rents, the results can be used to gain intuition on the effect of including tile drainage in a Ricardian analysis. Assume that there is a crop production function, $f(w)$, whose sole input is the amount of water supplied to crops, w. Let this function have an inverted-U shape, reflecting the fact that water is beneficial for crop growth up to a certain amount, $\overline{w}$, but is harmful beyond that point as the soil becomes saturated. This implies $f_w \geq 0$ when $w \leq \overline{w}$, $f_w \leq 0$ when $w \geq \overline{w}$, and $f_{ww} < 0$ throughout.

The supply of water for crops, $w(D; P)$, is a function of tile drainage, D, and the exogenous level of precipitation, P. Assume $w_D < 0$, since tile drainage is typically used to decrease the amount of water on cropland, and $w_{DD} > 0$, indicating that the amount of water that tile drainage is able to remove decreases as tile drainage increases.[5] Assume that $w_P > 0$, meaning that an increase in precipitation will increase the supply of water to crops. Let p_c and p_d represent the strictly positive, exogenous prices of crops and drainage, respectively. The farmer chooses D in order to maximize their profit:

$$\max_{D \geq 0} \pi = p_c f(w(D; P)) - p_D D \tag{1.1}$$

This gives the following first order condition:

$$p_c f_w w_D \quad \leq \quad p_D \tag{1.2}$$

Equation (1.2) holds with equality when the optimal amount of tile drainage is positive. Given the strictly positive prices, as well as the negative sign on w_D, this implies that tile drainage will only be used on the portion of f where the marginal product of water is negative

[5] Tile drainage is also used to reduce excess salinity in soils due to irrigation. Imperial County, California, for example, has a

($f_w < 0$). In other words, tile drainage will only be installed when the supply of water is so high that it is detrimental to crop growth.

On the other hand, when the optimal amount of tile drainage is zero, equation (1.2) becomes a strict inequality. This inequality is always satisfied when the available water supply is not harmful to crop growth ($f_w \geq 0$), indicating that farmers will not remove water from their land if it is beneficial to crops. However, depending on the relative prices of crops and tile drainage, it is also possible for the inequality to be satisfied when the amount of water is harmful to crop growth, $f_w < 0$. Specifically, if the marginal cost of tile drainage, p_D, is higher than the value of the marginal product, $p_c f_w w_D$, then farmers would rather accept the damage to their crops (or not grow crops at all) rather than install tile drainage.

More formally, the demand function for tile drainage is a function of the exogenous prices and precipitation, $D^*(\boldsymbol{p}, P)$, where $\boldsymbol{p} = (p_c, p_D)$. Substituting this demand function into equation (1.1) gives the profit function:

$$\pi = p_c f(w(D^*(\boldsymbol{p}, P), P)) - p_D D^*(\boldsymbol{p}, P) \tag{1.3}$$

When the optimal amount of tile drainage is zero, this simplifies to:

$$\pi = p_c f(w(0; P)) \tag{1.4}$$

Taking the derivative of equation (1.4) with respect to precipitation yields:

$$\frac{\partial \pi}{\partial P}|_{D^*=0} = \underset{(+)(+/-)(+)}{p_c f_w w_P} \tag{1.5}$$

The single term on the right hand side of equation (1.5) is simply the increase in profits that occur through an increase in crop productivity due to an increase in precipitation. In general, the sign of equation (1.5) is ambiguous, since optimal tile drainage may be zero even when $f_w < 0$, as noted above. If water is beneficial to crops on the average farm that does not install tile drainage, so that $f_w > 0$ when $D = 0$, then equation (1.5) predicts that precipitation increases profits on land without tile drainage.

When the optimal amount of tile drainage is positive, the derivative of equation (1.3) with respect to precipitation is:

$$\frac{\partial \pi}{\partial P}|_{D^*>0} \quad = \quad p_c f_w w_D D_p + p_c f_w w_p - p_d D_p \tag{1.6}$$

Grouping terms by D_p and using the fact that $p_c f_w w_D$ equals p_D when the optimal amount of tile drainage is positive:

$$\begin{aligned} \frac{\partial \pi}{\partial P}|_{D^*>0} &= D_p(p_c f_w w_D - p_D) + p_c f_w w_P \\ &= \underset{(+)(-)(+)}{p_c f_w w_P} \end{aligned} \tag{1.7}$$

Equation (1.7) shows that an increase in precipitation on tile drained lands will decrease profits. This makes sense, since farmers will either have to accept the crop damages due to an increase in precipitation, or pay the price of installing more tile drainage to mitigate damages.

If we assume that, on non-tile drained land, an increase in precipitation does not negatively effect crop productivity, so that $f_w \geq 0$, then:

$$\frac{\partial \pi}{\partial P}|_{D^*=0} \geq \frac{\partial \pi}{\partial P}|_{D^*>0} \tag{1.8}$$

The intuition behind this result is straightforward. In general, a farmer who hasn't installed tile drainage is still enjoying yield increases from increases in precipitation. On the other hand, we can infer that a farmer on tile drained land enjoys less benefits from increases in precipitation since they are already removing excess water.

To test the hypothesis provided by equation (1.8), we perform a Ricardian analysis which separates observations into tile drained and non-tile drained land. Land values are regressed on a set of explanatory variables, including precipitation. By comparing the coefficients on precipitation we are able to gain further insight into how its value differs between the two subsets.

Data

Land Rents

A Ricardian analysis assumes that current farmland values represent the sum of discounted land rents in equilibrium. Following the convention in the literature, we use reported farmland value as a proxy for land rents. Although most previous Ricardian analyses have used county averages, we use farm level data through a confidentiality agreement with USDA NASS. These data are from the 2012 U.S. Census of Agriculture and are reported in dollars per acre at the farm level. To focus on the effect of tile drainage, separate from irrigation, we follow Schlenker et al. (2006) and remove all counties west of the 100th meridian.[6]

Tile Drainage

The most current tile drainage data is the 2012 U.S. Census of Agriculture, which includes data on number of farms with tile drainage as well as the number of acres drained. These data represent the first attempt at assessing the extent of tile drainage in the U.S. in over 20 years.

Figure 1.1 shows the spatial distribution of tile drainage as a percentage of cropland acres across the entire U.S. Tile drainage is concentrated in the Midwest, with high concentrations in Iowa, Illinois, Indiana, Ohio, and southern Minnesota. Many of the counties in these states have tile drainage on over 70% of their cropland, with the highest being Henry, Ohio, with 84%. The map also shows that tile drainage is not exclusively used in the "wet" region of the U.S., as patches of tile drained counties appear throughout the West. However, tile drainage in these areas is primarily used to remove excess salinity from the soil that occurs due to irrigation.[7]

Through conversations with experts familiar with tile drainage data as well as our own analysis, we believe that the 2012 Census data on tile drainage is a reasonably accurate representation of the existing extent of U.S. tile drainage. However, due to the data limitations of past drainage data, as well as the large gap since the last set of estimates was released, we do not believe it is sensible to combine the 2012 data with historical data.[8] This limits the analysis

[6] We estimated the models using two alternative subsets of the data, which are discussed in the appendix.

[7] Imperial County, CA, shown as the southernmost dark blue county in California, has tile drainage in over 57% of its cropland, despite being one of the driest counties in the U.S.

[8] Appendix A provides a summary of historical tile drainage datasets.

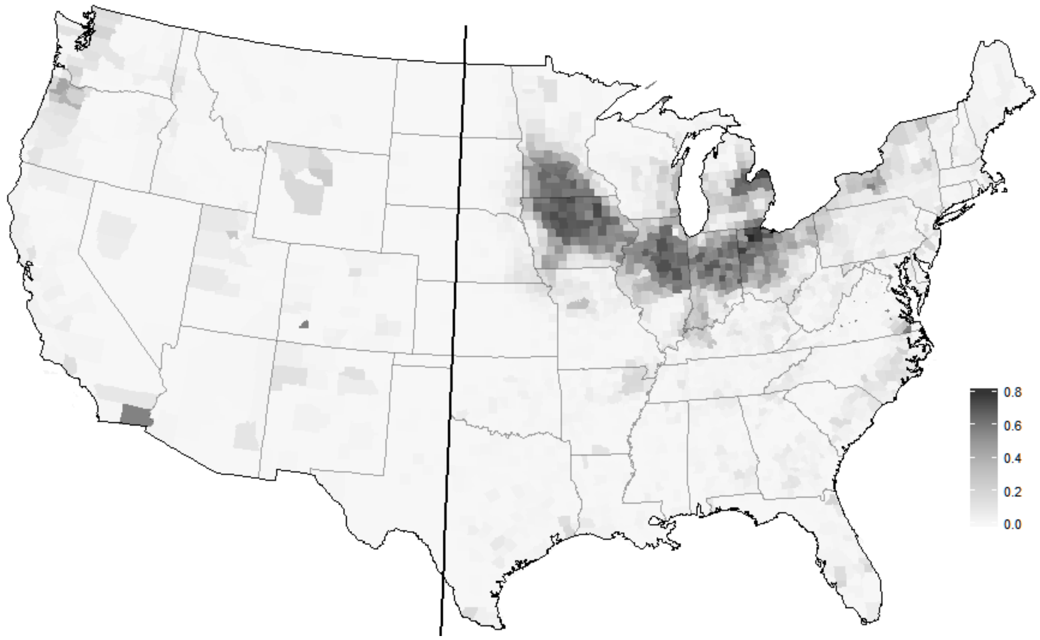


Figure 1.1: U.S. tile drainage as a percentage of cropland

Notes: Data on tile drainage and cropland come from the 2012 U.S. Census of Agriculture. Darker areas indicate a higher percentage of drained cropland. Vertical line represents the 100th meridian.

to a cross-sectional, Ricardian model, rather than alternative panel data methods (Deschenes and Greenstone, 2007; Massetti and Mendelsohn, 2011).

Climate

Climate variables are derived from Oregon State's PRISM datasets (Oregon State University, 2016). PRISM is a publicy available, high resolution dataset which has become ubiquitous in the literature on climate change and agriculture in the U.S. For each day since 1981, PRISM provides data on maximum and minimum temperatures, plus precipitation, for 481,631 grid cells which cover the contiguous United States. The size of each grid cell is 2.5 square miles.[9] We use daily data from 1982-2011 to calculate yearly variables, which are subset to April through September to approximate the growing season (Schlenker and Roberts, 2006). These yearly weather variables are then averaged over the 30 year period to represent the climate.

[9]Monthly values are available as far back as 1895.

Since we are only given the zip code of each farm, we construct our climate variables at the zip code level.

Temperature variables are derived from degree days to capture possible non-linear responses of crop growth to heat (Schlenker and Roberts, 2006).[10] These variables include growing degree days, defined as the number of degree days between 8 and 32, and degree days above 34, which is a threshold that has been used to indicate temperatures that are harmful to plant growth (Schlenker et al., 2006). These variables, along with daily precipitation, are summed over the growing season.

To isolate the effects of the climate variables over cropland within a county, rather than over all land (which can include developed land, water bodies, etc.), we weight each PRISM grid cell by its proportion of land in cropland. Using the approach of Schlenker et al. (2006) we intersect a shapefile of PRISM grid cells with a 2011 National Land Cover Database (NLCD) satellite image using ArcGIS.[11] To calculate the proportion of cropland, we sum the areas of "cropland" or "pasture" and divide by the area of a PRISM grid cell.

Soil and Demographics

To control for land quality across the U.S., we include a group of variables from the STATSGO soil data set. STATSGO divides the U.S. into over 10,000 "map units" whose soil characteristics are determined by statistically expanding the results of several samples in each unit. The variables taken from STATSGO include measures of water capacity, slope, k-factor, amount of clay, soil class, drainage class, and saturation.[12] Each variable is calculated as a weighted average of map units within a zip code, where the weight is the area of cropland within a map unit.

Several papers have shown that proximity to urban areas can influence farmland values (Zhang and Nickerson, 2015; Shi et al., 1997). To control for these influences we use demo-

[10] Daily Degree days are derived from daily minimum and maximum temperature using the procedure from Schlenker et al. (2006).

[11] The 2011 NLCD satellite data can be downloaded at www.mrlc.gov.

[12] The definition of these variables can be found in Appendix A.

graphic data from the 2010 U.S. Census. These variables are at the zip code level, and include population density and median income per capita.

Summary Statistics

Table (1.1) presents the summary statistics for the variables used in our estimation. The first two variables, farmland value per acre and percent of cropland tiled, come from the farm level data provide by NASS. Farmland values have an average value of $4,074 per acre and are skewed towards zero, with a maximum of $126,087. The percentage of cropland with tile drainage is skewed towards zero, with an average of 9.6%. Growing degree days and precipitation have roughly symmetric distributions, with means of 2 thousand degree days and 54.74 centimeters, respectively. The distribution for degree days above 34 is skewed towards zero, with a mean of 5.06 harmful degree days.

Empirical Model

A traditional Ricardian analysis regresses farmland values on climate variables and controls without distinguishing between adaptive technologies available to farms. In this paper we introduce tile drainage using a Structural Ricardian model similar to Kurukulasuriya et al. (2011). This two-stage model, closely related to the Heckman model (Heckman, 1977), explicitly models the choice of tile drainage, as well as changes in farmland values, on both tile drained and non-tile drained land. For the following models, the dependent variables are farm level and taken from the 2012 Census of Agriculture, while the explanatory variables are at the zip code level.

The first stage models tile drainage as a dichotomous choice, estimated using a probit regression. That is, we define a farm as tile drained ($TD = 1$) if it contains a non-zero amount of tile drainage according to the 2012 U.S. Census of Agriculture. Assume that farmer i in zip code z and state s decides whether or not to install tile drainage, TD, conditional on a vector of exogenous local variables X, which includes precipitation ($prec$), growing degree days (GDD), harmful degree days (HDD), as well as a vector soil and land characteristics ($\boldsymbol{X}$), state fixed effects (η_s) and an unobservable error term ϵ_{izs}:

Table 1.1: Descriptive Statistics for Primary Data Set

	Mean	SD	10th Percentile	90th Percentile
Farmland and tile drainage:				
Farmland value ($/acre)	4,074.31	4,353.78	1,766.60	6,724.40
Percent of cropland tiled	9.62	17.60	0.00	38.30
Climate data:				
Growing degree days (thousands)	2.24	0.52	1.60	3.01
Precipitation (cm)	54.74	7.92	44.76	63.87
Degree days above 34	5.06	8.97	0.10	14.99
Soil and land data:				
Slope	8.81	8.05	1.88	18.47
K-factor	0.11	0.13	0.00	0.33
Soil class (%)	56.01	14.19	39.11	76.63
Water capacity (in./in.)	7.84	1.81	5.52	10.19
Soil Permeability (in./h)	11.59	13.54	0.00	27.76
Percent clay	10.47	12.01	0.00	28.24
Latitude (degree)	38.34	4.48	31.86	43.45
Demographic Variables:				
Population density (hundreds per square mile)	5.97	15.66	0.10	18.80
Income per capita (thousand $)	11.31	16.77	2.81	22.65

Note: All dollar values are in 2012 U.S. dollars. Climate, soil, and demographic data are derived from zip code level data. Farmland value and percent of cropland tiled are farm level.

$$TD_{izs} = \beta_0 + \mathrm{f}(prec_{zs}; \beta_{prec}) + \mathrm{g}(GDD_{zs}; \beta_{GDD}) + \beta_{HDD} HDD_{z,s} + \beta' \boldsymbol{X}_{zs} + \eta_s + \epsilon_{izs}, \quad (1.9)$$

where $TD_{izs} = 1$ if that farm contains a positive amount of tile drainage.

Estimating equation (1.9) serves three main purposes. First, we can gain insight into the farmer's choice to install tile drainage. Specifically, we would like to know what effect changes in the climate have on the probability of adopting tile drainage. Second, estimates of (1.9) are used to contruct an inverse mills ratio, which is included in the second stage to control for selection (Heckman, 1977). Fianlly, we will use theses estimates to construct the damage estimates from climate change in a later section.

The second stage estimates two separate Ricardian functions for tile drained and non-tile drained land (Π^d_{izs} and Π^n_{izs}, respectively), where Π represents the value of cropland per acre.

$$
\begin{aligned}
\Pi^d_{izs} &= \beta_0 + \mathrm{f}(prec_{zs};\beta_{prec}) + \mathrm{g}(GDD_{zs};\beta_{GDD}) + \beta_{HDD}HDD_{z,s} + \beta' X_{zs} + \beta_\lambda \lambda_{izs} + \eta_s + \epsilon^d_{izs} \\
\Pi^n_{izs} &= \beta_0 + \mathrm{f}(prec_{zs};\beta_{prec}) + \mathrm{g}(GDD_{zs};\beta_{GDD}) + \beta_{HDD}HDD_{z,s} + \beta' X_{zs} + \beta_\lambda \lambda_{izs} + \eta_s + \epsilon^d_{izs}
\end{aligned}
$$

The functions f and g are unknown; for the main results in this paper we use a quadratic functional form, as this is a popular convention in the literature. Results for several alternative specifications can be found in the appendix. The unobservable variables that determine the farmer's choice of tile drainage in equation (1.9) are likely to be correlated with the unobservable variables that determine farmland values. To control for the resulting selection bias, we calculate the inverse mills ratio from equation (1.9) and include it as an additional explanatory variable (λ) in the Ricardian equations.

Regression Estimates

In this section, we present the first and second stage regression results. For clarity, we limit the displayed coefficients to the climate variables. The full results can be found in appendix (A).

Table (1.2) shows the results using a quadratic functional form for precipitation and growing degree days. We report the marginal effects for the probit model rather than the coefficients for easier interpretation. For the first stage probit model, the signs indicate that the probability of adopting tile drainage increases with precipitation, at a decreasing rate. A one centimeter increase in precipitation increases the probability of adopting tile drainage by 0.8% The same positive relationship is found for growing degree days (GDD). Since GDD is in thousands, an increase in one thousand degree days increases the probability of adopting tile drainage by 30%. On the other hand, an increase of one harmful degree day (HDD) decreases the probability of adopting tile drainage by 1.6%. These results are intuitive since a farmer should be more inclined to invest in the cost of tile drainage as farming conditions become more favorable. Coefficient estimates for soil and land characteristics, shown in the appendix, follow the intution from agronomics, based on conversations with experts. Slope, for example, has a strongly negative relationship with tile drainage, as an increase in slope will increase a farm's natural drainage, thus decreasing the need for artificial tile drainage.

Next, we turn to the estimates of the effects of climate on land values. Columns 2-4 of table (1.2) shows the OLS results for pooled, tile drained, and non-tile drained observations. The coefficients on the precipitation variables in the pooled model are relatively low in magnitude compared to the tile drained and non-tile drained estimates and indicate that a one centimeter increase in precipitation increases farmland value by 1.41%. Although the coefficient on precipitation in the pooled model is only significant at the 10% level, and its square is not significant, the two variables are jointly significant at the 1% level.

The coefficient estimates for precipitation on tile drained and non-tile drained are several times larger than the pooled model. In addition, there are clear differences between the two subsamples. To help visualize these differences, the shape of the quadratic curves for tile drained and non-tile drained land are shown in figure (1.2). The intercept has been normalized to zero for both curves to highlight the differences in curvature. Consistent with the prediction from the conceptual model, the curve for non-tile drained land is steeper than on tile drained land, which is relatively flat. The curves imply an optimal amount of precipitation of 57 and 59 centimeters for tile drained and non-tile drained land, respectively.

The coefficient estimates for GDD for all three models are strongly significant and indicate an inverted U-shape. The coefficient on GDD for the pooled model has a much smaller magnitude than the two subsamples and indicate 2,420 as the optimal number of growing degree days. The coefficient for GDD on non-tile drained is higher than on tile drained land, although the difference is not as stark as for precipitation. The curves imply an optimal number of growing degree days of 1,903 and 2,006 for tile drained and non-tile drained farms, respectively.

There are sharp differences in the effect of harmful degree days (HDD) on farmland values. The pooled estimate is relatively close to previous estimates in the literature (Schlenker et al., 2006) and indicate that an increase of one HDD decreases farmland value by 9.8%. On tile drained land, however, the coefficient is nearly twice as large as the pooled estimate. On non-tile drained land, the coefficient is statistically significant and indicates an increase of one HDD decreases farmland value by a little over 40%.

Table 1.2: Regressions Results

	Probit	OLS		
	(1)	(2)	(3)	(4)
		Pooled	Tile Drainage	No Tile Drainage
Prec.	0.00842***	0.0141*	0.0793***	0.172***
	(0.00205)	(0.00814)	(0.0193)	(0.0106)
Prec.^{2}	-0.0000836***	-0.0000753	-0.000698***	-0.00146***
	(0.0000188)	(0.0000726)	(0.000172)	(0.0000942)
GDD	0.317***	2.865***	6.700***	7.453***
	(0.0357)	(0.155)	(0.389)	(0.215)
GDD^2	-0.0908***	-0.592***	-1.760***	-1.858***
	(0.00835)	(0.0367)	(0.100)	(0.0561)
HDD	-0.0157***	-0.144***	-0.259***	-0.416***
	(0.00253)	(0.0124)	(0.0315)	(0.0165)
IMR			1.206***	2.332***
			(0.151)	(0.0931)
N	812,304	812,304	128,716	683,588
R^2		0.09	0.06	0.10

Notes: Marginal effects and coefficient estimates for probit and OLS models, respectively, along with standard errors. The dependent variable for the probit model is equal to 1 if a farm contains tile drainage, and zero otherwise. The dependent variable for the OLS model is the log of farmland value per acre. The "pooled" column pools all farms east of the 100th meridian, while the "Tile Drainage" and "No Tile Drainage" columns use only farms with tile drainage or no tile drainage, respectively. GDD stands for growing degree days, and HDD stands for harmful degree days. All models includes state fixed effects and standard errors clustered at the zip code level. See the appendix for full results.

The coefficients on the inverse Mills ratio are both positive and strongly significant for the tile drained and non-tile drained models, which is evidence that the pooled model suffers from selection bias. The positive sign indicates a positive covariance between the error terms in the first and second stage models.

The signs of the control variables, included in the appendix, largely follow intuition. Slope has a negative relationship with tile drainage adoption, since a higher slope creates more natural drainage and therefore less of a need for artificial drainage. The same argument can be used for the negative relationship with water permeability- if water easily moves through the soil there

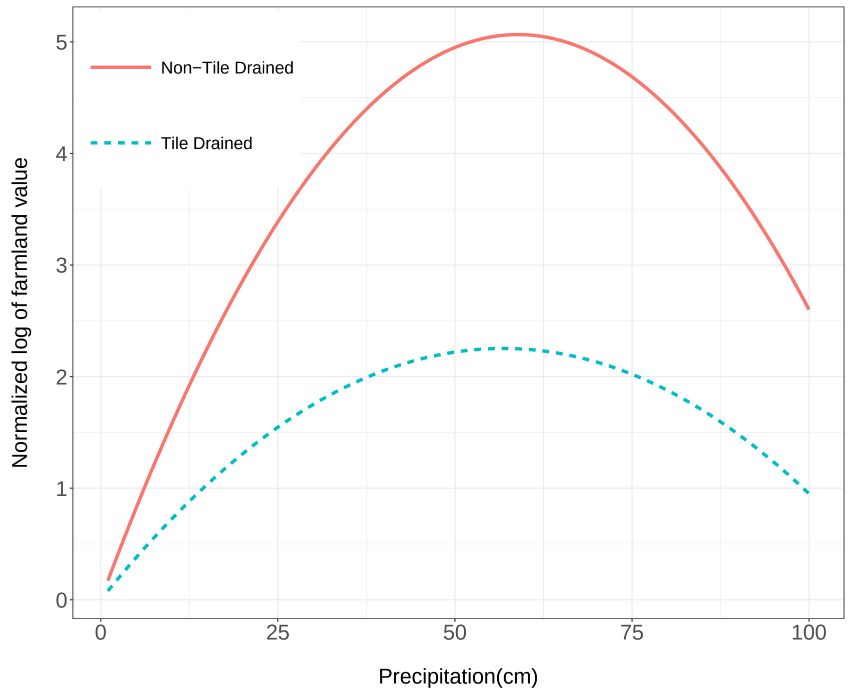


Figure 1.2: Estimated quadratic response of farmland values to precipitation

Notes: Plot of the quadratic curves for precipitation implied from an OLS regression of the log of farmland values on climate, soil, and demographic variables as well as state fixed effects. The intercepts have been normalized to zero to facilitate comparison of the shapes of the curves.

is less of a need for artificial drainage. Soil class has a positive relationship, since it implies the farmland (and therefore the choice to install tile drainage) is more profitable.

The following section applies these estimates to climate change simulations to estimate the damages to agriculture through climate change.

Climate Change Simulations

Data on climate change were derived from global climate models (GCM's) used in the fifth IPCC report, each of which produced daily weather simulations through 2100. Simulated climate data can vary widely between different GCM's, and can even give contradicting results for the same regions (Auffhammer et al., 2013). We limit our choice of models to the five models which performed the best at simulating historical data in the Eastern half of the United States. These rankings were derived from (Sheffield et al., 2013), who calculate the root mean squared error (RMSE) of historical forecasts for a variety of climate variables in this region. We use rankings of the RMSE for precipitation and temperature to compute an average rank, and choose the top five models available. These models were MIROC-ESM, IPSL-CM5A-LR, CCSM4, BCC-CMS1.1, and GFDL-ESM2M [13].

Simulated data from these models have been down scaled through NASA's NEX project to 25 by 25 km grid cells.[14] Climate normals for each NEX grid cell were calculated for two thirty year periods: 2020-2049 and 2070-2099. We then calculate the climate normals for each PRISM grid cell that covers the U.S. by taking a weighted average of the five closest NEX grid cells, using the inverse of the distances between the centroids as weights. To find the projected climate normals for each county, we average the PRISM grid cells over the portions of the county that contained cropland, using the same method described in the data section.

To establish a standard for model simulations, the GCM's used in the fifth IPCC report simulated data for the same four scenarios, known as Representative Concentration Pathways (RCP's). We calculate damages for RCP 4.5 and RCP 8.5. RCP 4.5 assumes emissions peak around 2040 and then decline and is often described as a moderate scenario. RCP 8.5, the most severe scenario, assumes emissions continue to rise throughout the 21st century.

To calculate the change in climate we compare the future climate scenarios with a historical climate scenario generated by each GCM, rather than PRISM data, since each model's historical simulations will differ from the actual data. We define the historical scenario as 1976-2005, which is the most recent 30 year period available for the NEX data. For our main results we use climate changes calculated from an ensemble average of the 5 GCM's. Summary statistics for this ensemble average are shown in table (1.3). Average temperature changes range from 1.11 to 1.34 in the short term and 2.12 to 4.59 in the long term. Precipitation changes range from 0.88 cm to 1.65 cm in the short term, and 0 cm to 2.14 cm in the long run. In contrast to temperatures, the average long run change in precipitation is actually projected to be lower for the severe scenario compared to the moderate scenario, and also has a wider range.

[13]The top ranked model, HadGEM2-ES, was not available for download through NASA's NEX website.
[14]https://cds.nccs.nasa.gov/nex-gddp/

Table 1.3: Summary Statistics for Climate Projections

	Mean	SD	10th Percentile	90th Percentile
RCP 4.5, 2020-2049:				
Temperature (C°)	1.11	0.09	0.99	1.21
Growing degree days (thousands)	0.18	0.02	0.15	0.19
Degree days above 34 C°	5.15	6.78	0.03	14.51
Precipitation (cm)	0.88	1.82	−0.96	3.04
RCP 8.5, 2020-2049:				
Temperature (C°)	1.34	0.15	1.11	1.49
Growing degree days (thousands)	0.22	0.03	0.17	0.25
Degree days above 34 C°	6.65	7.41	0.36	18.37
Precipitation (cm)	1.65	1.29	−0.00	3.09
RCP 4.5, 2070-2099:				
Temperature (C°)	2.12	0.21	1.89	2.39
Growing degree days (thousands)	0.34	0.04	0.28	0.39
Degree days above 34 C°	15.62	16.00	1.24	37.22
Precipitation (cm)	2.14	1.93	−0.07	4.62
RCP 8.5, 2070-2099:				
Temperature (C°)	4.59	0.34	4.16	4.90
Growing degree days (thousands)	0.69	0.09	0.56	0.77
Degree days above 34 C°	70.12	52.00	12.70	149.15
Precipitation (cm)	−0.00	4.04	−5.26	3.68

Notes: Sample statistics for changes in climate variables relative to simulated data for 1976-2005. These statistics are calculated from the ensemble average of five Global Climate Models.

Damage Estimates

Climate change will affect both the estimated probability of a land containing tile drainage, as well as the value of that land. We report the expected change in cropland value as a weighted average of the change in cropland value of observation i when it has tile drainage, and when it does not have tile drainage. The weights are the probability of being in each type of land following climate change, based on the first stage probit estimates:

$$EV_i = P(\widehat{Y_i = 1})(\widehat{\Pi_d}) + P(\widehat{Y_i = 0})(\widehat{\Pi_n}),$$

We assume that farmers who have tile drainage in 2012 do not remove tile drainage, which implies that all the weight goes on the first term in equation (1.10) for those farmers. Farmers who do not have tile drainage in 2012 will have some non-negative weight on both terms. For ease of reporting, we report the weighted percentage change in value rather than the weighted dollar value.

Table (1.4) shows the mean and standard deviation of the estimated percentage change in farmland values. Column (1) shows damage estimates using a quadratic functional form for precipitation and growing degree days. For comparison, columns (2) and (3) show the damage estimates using a step-function for precipitation and growing degree days, while column (2) uses a quadratic form for these variables but splits up the season into “Spring” (April through June) and “Summer” (July through September).[15]. For the moderate RCP 4.5 scenario, estimated damages range from -11.45% to -43.26% from 2020-2049, and from -26.78% to -68.99% from 2070-2099. For the severe RCP 8.5 scenario, estimates range from -13.25% to -49.65% from 2020-2049, and from -61.90% to -95.60% from 2070-2099. Interestingly, there is little difference in the damage estimates when using a seasonal model as compared to an annual model. The step-function model has less severe estimates because the damages are not as severe at the tails, although this result may be driven by fewer observations.

Figure (1.3) show the spatial distribution of total damages using estimates from the quadratic model. Climate change will have a positive benefit for the northernmost counties, but a neg-

[15] The full set of coefficient estimates for these models can be found in the appendix

Table 1.4: Predicted Impact of climate change on farmland values

	(1) Quadratic	(2) Step-Function	(3) Seasonal
RCP 4.5, 2020-2049:			
Mean	−41.06	−11.45	−43.26
Std. Deviation	29.78	41.79	36.51
RCP 4.5, 2070-2099:			
Mean	−68.99	−26.78	−65.36
Std. Deviation	28.39	50.41	40.76
RCP 8.5, 2020-2049:			
Mean	−48.38	−13.25	−49.65
Std. Deviation	28.79	42.90	32.98
RCP 8.5, 2070-2099:			
Mean	−94.34	−61.90	−95.60
Std. Deviation	11.98	30.53	18.91

Notes: Percentage change in farmland values due to climate changes across two different climate change scenarios (RCP 4.5 and RCP 8.5), and two different time periods (2020-2049 and 2070-2099). Column (1) uses a quadratic functional form for precipitation and growing degree days. Column (2) uses a step-function. Column (3) uses a quadratic specification in both the Spring and Summer. Climate projections were Calculated from an ensemble average of five GCMs.

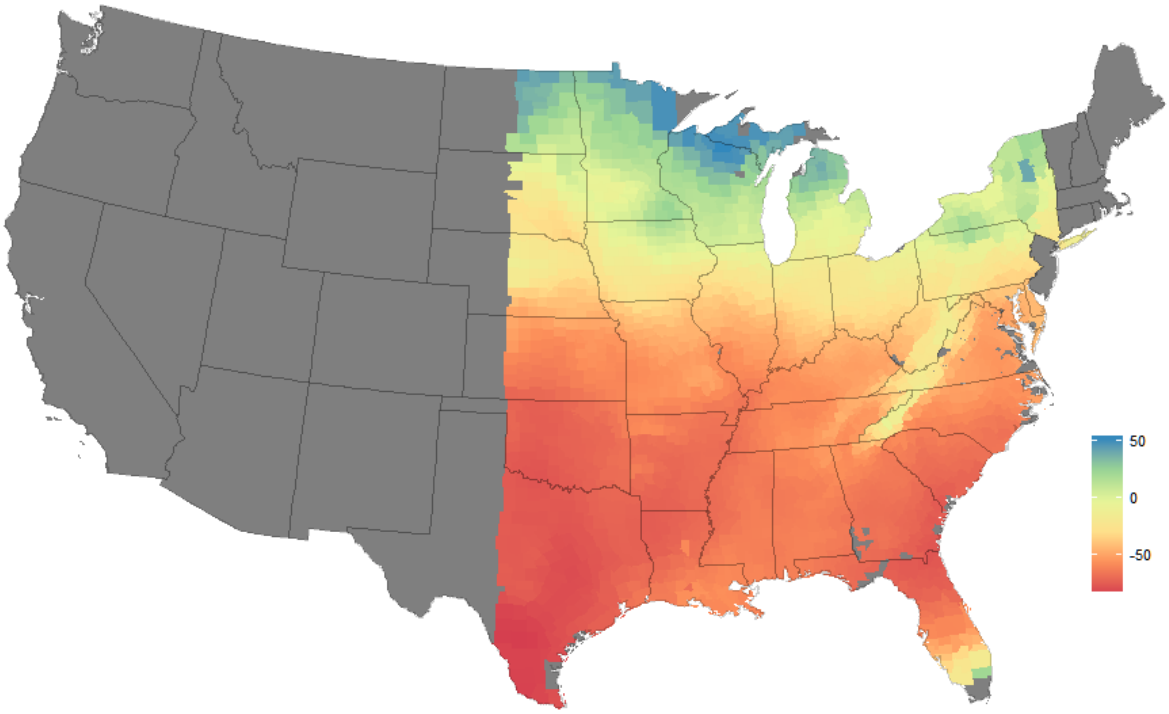


Figure 1.3: Spatial distribution of climate change impacts, 2020-2049

Notes: Percentage change in farmland values under the RCP 4.5 scenario from 2020-2049. Coefficient estimates come from the quadratic model using farm level observations. Climate simulation data comes from an average of five GCM models.

ative effect elsewhere. In particular, states in the middle of the country will suffer the most damages. Figure (1.4) shows the spatial distributions of the *difference* in climate change impacts between the Structural Ricardian model and the pooled models. Thus these are spatial distributions of the potential bias in pooled estimates. This figure indicates that the pooled model underestimates the benefits the northernmost counties receive from climate change. For all other regions, which will incur damages from climate change, the negative sign indicates that the pooled model underestimates the damages.

Conclusion

In this paper we study how climate variables are capitalized into farmland values while explicitly accounting for an important adaptation technology, tile drainage. We build intuition for the empirical analysis by formulating a simple profit maximization model that includes the

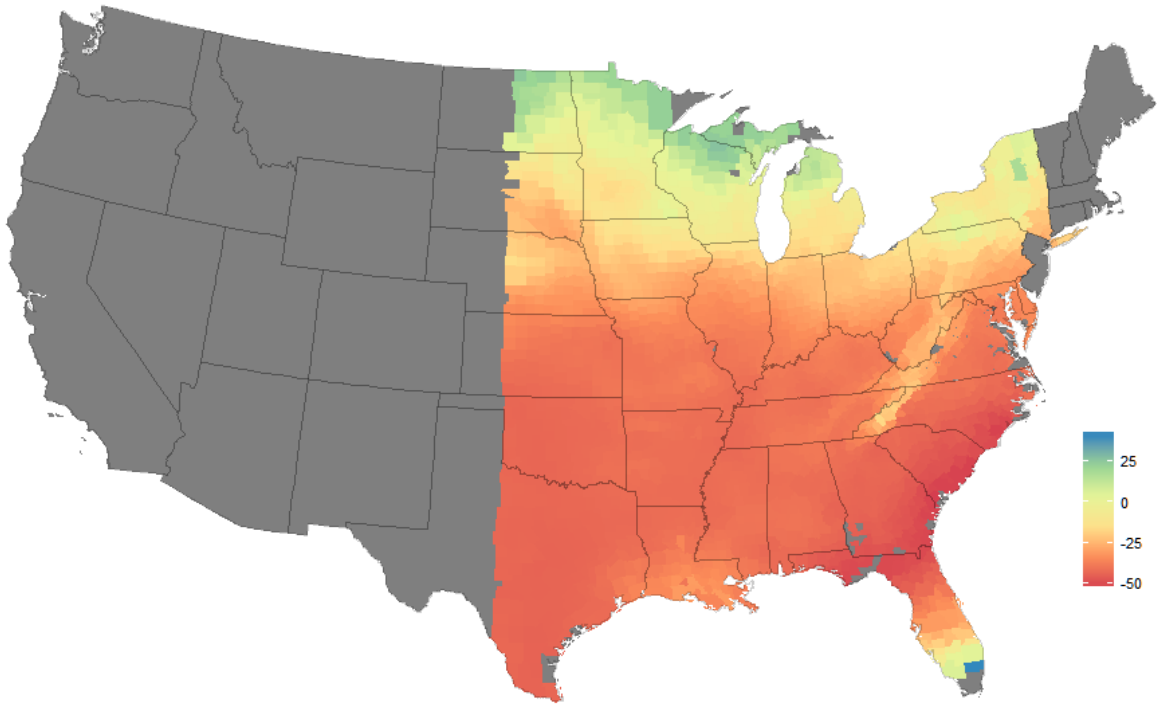


Figure 1.4: Spatial distribution of the difference in climate change impacts between the Structural Ricardian model and the pooled model, 2020-2049.

Notes: Difference in the percentage change in farmland values for the pooled and Structural Ricardian models using the RCP 4.5 scenario from 2020-2049. Coefficient estimates are from the quadratic model using farm level observations. Climate simulation data comes from an average of five GCM models.

farmer's choice of tile drainage. The model predicts that the marginal value of precipitation is higher in non-tile drained land versus tile drained land. We test this result by estimating a Structural Ricardian model, which accounts for the farmer's choice of tile drainage through a probit regression. We estimate this model using newly available, micro level data from the 2012 Census of Agriculture.

We find evidence that the marginal value of precipitation is higher on non-tile drained land, consistent with the model's predictions. An increase of one centimeter in precipitation increases farmland value by 7.93% on tile drained land, but by 17.2% on non-tile drained land. There are also key differences between the value of temperature increases in the two subsamples, in particular harmful degree days. These results indicate that models that pool together tile drained and non-tile drained land may result in biased coefficients.

Under a short term, moderate warming scenario, climate change decreases farmland values by an average of 41.06%. Damages are most severe in the lower half of the U.S., while most northern border states experience benefits due to climate change. There are important geographical patterns in the differences between damage estimates using a model that pools together all farms, and one that explicitly models tile drained farms. In particular, benefits to border countries may be underestimates, while damages to farms in the mid and southern U.S. may be overestimated.

.

CHAPTER 2. THE IMPACT OF LAND USE ON LAKE WATER QUALITY: EVIDENCE FROM IOWA

Abstract

I use a unique panel data set to estimate the effect of changes in cropland on lake water quality. Fifteen years of water quality measurements across over 100 lakes are combined with satellite imagery and weather data. Using a dynamic panel data model, I find that the elasticity of water quality to cropland is 0.0535. To understand the policy implications, I estimate a second model to find the elasticity of cropland to crop prices. I use these estimates to analyze the effect of the Renewable Fuel Standard on lake water quality. Due to the inelastic responses from both models, I find a negligible impact on lake water quality. [1]

JEL codes: Q15, Q18, Q51

Keywords: Water quality; agriculture

[1] I would like to thank the EPA for their support of related work as a graduate student.

Introduction

Understanding the effect of land use on water quality is an important question in environmental policy. Agriculture in particular is consistently identified by the EPA as a cause of water quality degradation due to fertilizer runoff (EPA, 2002). These effects have important economic and ecological consequences, such as adverse effects on recreation and drinking water quality. As has been known since the time of Pigou, these effects are a classic case of a market failure. Since water quality is not a market good, the free market will not provide the optimal level. Efficient government intervention requires understanding the benefits and costs of land use on water quality, but measuring these effects has been difficult.

In this paper, I use a unique panel data set to estimate how changes in cropland affects lake water quality to help inform the cost of policies which affect land use. The panel nature of the data allows me to control for fixed effects as well as the dynamic effects of changes in lake water quality. Controlling for fixed effects helps mitigate omitted variable bias, while controlling for the dynamic effects allows for both short and long run estimates of water quality changes (Montgomery and Reckhow, 1984). Both of these important issues are addressed using a dynamic panel data model. I estimate the model with high quality water measurements across 100 Iowa Lakes over 15 years, along with satellite data on cropland use. I find a statistically significant decrease in water quality due to an increase in cropland. I also find evidence that water quality effects persists over time.

Using the estimates, I analyze the effect of the Renewable Fuel Standard (RFS) on lake water quality. There is evidence that the RFS has had a significant effect on the price of corn and soybeans, which are the primary crops grown in Iowa (Hausman et al., 2012). If farmers are responsive to price changes then they will expand the cropland devoted to corn, which can increase nutrient runoff to nearby lakes. To examine this I estimate a secondary model of the elasticity of land use to price changes. I then construct a counter factual scenario where the biofuel mandate did not occur, and simulate the resulting water quality. Due to the relatively inelastic responses estimated from both models, I find a negligible effect of the RFS on lake water quality.

This paper contributes to the literature in several ways. First, it exploits 15 years of water quality measurements across Iowa to perform a statistical analysis of the effect of land use on water quality. Second, it provides strong statistical evidence of a persistent effect of water quality across time. Third, it adds to the literature on the response of cropland expansion to crop prices. Finally, it adds to the literature on the environmental effects of biofuel related policies.

The paper is organized as follows. The next section provides some background on the typical techniques used to assess the relationship between land use and water quality. This is followed by a description of the econometric model used in both the water quality and cropland response models. I then give a detailed description of the data set, followed by the results for both models. These results are then used in an application to estimate the effect of the RFS on lake water quality, followed by a brief conclusion.

Background

There is a long history of studies that attempt to identify the relationship between land use and the water quality of lakes, rivers, and streams. Most of these studies can be divided into two types- simulation models such as SWAT and BASINS [2], and econometric models. The former are able to model complex relationships between the climate, land use, and water quality to examine issues that might otherwise be intractable. For example, simulations from these types of programs have been used to examine the hypoxia "dead zone" in the Gulf of Mexico (Rabotyagov et al., 2014), the effect of corn-based ethanol on environmental quality (Secchi et al., 2009), and the potential for cropland to reduce flood risk (Schilling et al., 2014). Simulation models are invaluable for gaining insight into issues that may otherwise be too complicated for any one statistical model to capture, but they have drawbacks. On a practical level, the complexity of the simulated relationships requires a large number of parameters, and choosing these parameters requires a significant amount of expertise. Statistical models, on the

[2]SWAT stand for Soil and Water Assessment Tool; Basins stands for Better Assessment Science Integrating point & Non-point Sources.

other hand, are helpful in their ability to model relationships between variables in a relatively straightforward and transparent way.

Many statistical analyses in the literature rely on simple correlation coefficients between different land uses and a measure of water quality. For example, Tong and Chen (2002) found a statistically significant positive correlation of 0.1913 and 0.1563 between agriculture and total nitrogen and phosphorous, respectively, in 11-digit HUCS in Ohio. In fact, most studies find a positive correlation between the two variables (Meador and Goldstein, 2003; Dauer et al., 2000) While correlations are informative, they do not help isolate causal effects. In other words, does an increase in cropland cause the water quality to drop, or could it represent something else, such as the quality of the land? Answering this requires a model that controls for the quality of land, as well as other possible omitted variables.

A number of studies have used regression techniques to try to estimate the relationship. Tu (2011) uses geographically weighted regressions to estimate local effects in an area surrounding Boston. He estimates a separate univariate regression for 6 different water quality variables and 14 land uses, for a total of 84 regressions. The results showed little influence of agricultural land on water quality. A drawback of this study is that water quality measurements are averaged over time and estimated using only one year of observed land uses. In fact, cross-sectional regressions are common in water quality studies- possibly due to a lack of quality, publicly available time series data. Another technique often encountered in the literature are simple univariate regressions of land use on water quality, for example Lougheed et al. (2001). Limiting the model to one period, or not controlling for other factors that can affect water quality can potentially bias the coefficients of interest.

This omitted variable bias problem can potentially create misleading results. For example, Sprague and Gronberg (2012) study the effect of the Conservation Reserve Program (CRP) on total nitrogen and phosphorus loadings in rivers using a cross-sectional regression and found a marginally positive effect, indicating that CRP land increases nutrient levels; this is the opposite effect intended by the program and lacks a credible explanation. The key problem in this study is that, for the results to hold, the model must assume that CRP land is randomly distributed across space and uncorrelated with omitted factors that affect water quality. This is unlikely to be the case since profit maximizing landowners will choose to retire the least profitable farmland into the CRP first. In Iowa, for example, CRP land is concentrated in the south, where the soil quality is relatively low. Therefore it would not be surprising to find a negative correlation between CRP land and water quality, since lower quality soil typically means increased runoff.

The geographic characteristics of water bodies has led to some studies that use more complicated regression models. Atasoy et al. (2006) employs spatial econometric techniques to study the effect of urban land use on water quality. Their analysis uses monthly nutrient measurements over a four year period combined with monthly measures of urban development, weather variables, and a single year of satellite imagery to control for agricultural land. Their emphasis on rivers and streams is an example of how geography plays an important role in the specification of an appropriate econometric model when studying environmental issues. In their study, upstream river quality clearly affects downstream river quality as it is carried through a stream network, thus it makes sense to explicitly include a spatially-lagged dependent variable while allowing for temporal correlations in the error term. In this study, where the observed unit is lakes, it does not make sense conceptually to include a spatial lag, since lakes do not flow into each other. Instead, it is more appropriate to include a temporal lag of the dependent variable, since lake water remains relatively stationary over time. Spatial correlation is instead introduced through the errors, clustered by hydrological unit.

Existing lake water quality studies that attempt to include dynamics have typically been confined to one lake and its watershed. For example, Balkcom et al. (2003); Mankin et al. (2003) used multiple samples from a lake over time to calibrate an integrated assessment model, which was then used to analyze different land use scenarios. By contrast, this study uses data on over 120 lakes over 15 years, creating a rich variation in lake quality, geographical characteristics, and the characteristics of surrounding land use.

As one of the most productive farming states in the country, Iowa land use can be particularly sensitive to changes in farming policies. Therefore, given the evidence of the link between cropland and water quality degradation, government policies can directly and indirectly affect water quality. Two primary examples are the Conservation Reserve Program (CRP), and the Renewable Fuel Standard (RFS).

The CRP has evolved from its initial goals of removing cropland from production to focus more on maximizing the environmental impact of the program. Only land currently in production or expiring CRP land are eligible to be retired and receive CRP subsidies, and retired land must be planted with species that will improve environmental health and quality. Thus the possible water quality benefits of an acre of CRP land are 1) removing an acre of cropland, and therefore all related nutrient use, and 2) replacing it with an acre of plants that can help improve soil quality and reduce runoff of nutrients from the surrounding area. CRP land in Iowa began a major decline around 2007. It is likely that multiple factors contributed to this decline, especially rising crop prices (and thus profitability of land) and a decline in funding for the program.

The RFS, first established under the Energy Policy Act of 2005, mandated 7.5 billion gallons of ethanol be used by 2012. The scope of the biofuel mandate expanded significantly in 2007 by mandating 36 billion gallons of ethanol in the U.S. by 2022. The vast majority of the current biofuel supply comes from corn ethanol. Therefore, the biofuel mandate has and will continue to have significant economic and environmental impacts on Iowa, the nation's leading corn producer. In particular, researchers have identified water quality degradation as an important consequence of biofuel production (Simpson et al., 2008). Although corn cultivation requires a significant amount of water, water shortages are typically not a concern in Iowa. Rather, the increased use of nutrients from expanding corn production along both the intensive and extensive margin are of concern. In addition, an increase in the demand for corn can affect the price of other crops, such as soybeans, which can cause cropland expansion for those crops as well. As corn uses nitrogen relatively inefficiently (Balkcom et al., 2003), switching over to corn from other crops can potentially increase the amount of nitrogen in the soil. Finally, if we assume that farmers grow crops on the best farmland available, cropland expansion will likely occur in marginal, more environmentally sensitive areas, including CRP land (Secchi et al., 2009). Thus the two policies mentioned here are to some degree interdependent, as farmer's will look to maximize their profit by either accepting subsidies to retire the land into the CRP, or to farm the land and sell their crops.

Data

Water quality data were downloaded from Iowa State's Limnology Laboratory website [3]. The Iowa Lakes data uses consistent, scientifically based, and well documented hydrological sampling methods. In this paper I use the average of three annual measurements. Averaging three lake water samples over a year offers adequate precision for water quality indicators Downing et al. (2006). CTSI, a measure of lake water clarity, is used as the main water quality indicator because it summarizes the outcome of increased sediment or nutrient loadings, as opposed to a measurement of the inputs of sediments or nutrients into a lake. Most lakes have a CTSI between 0-100, with each increase of 10 units representing an approximate doubling in algal biomass. An intuitive way to think about the index is that a CTSI of 0 represents a visible depth of 64 meters, while a CTSI of 100 represents a visible depth of only 6.4 centimeters. The CTSI can be approximately divided into four trophic classes: oligotrophic (less than 30-40), mesotrophic (40-50), eutrophic (50-70), and hypereutrophic (70-100+). Figure (2.2) shows the locations of the 123 lakes used in this study.

Annual land use data comes from USDA NASS cropland data layers (CDL's), which are satellite images. Each pixel of a satellite image is assigned a land use based on color analysis. I find the total land use for a geographic region by summing the pixels assigned to each land use. Since the focus of the paper is water quality, land use was aggregated to the local watershed level, known as a HUC (hydrologic unit code). Aggregating to a watershed captures drainage characteristics more accurately than aggregating to an arbitrary governmental boundary. HUC's differ in size and are nested within each other- a HUC12 is located within a HUC10, which itself is within a HUC8, and so on. The right hand panel of figure (2.2) shows the size of typical HUC12 watersheds, as well as the larger HUC8 which contains them.

[3] www.limnology.eeob.iastate.edu/lakereport/default.aspx

(a) (b)

Figure 2.2: Iowa Lakes, HUC8, and HUC12 watersheds

Note: Panel (2.1a) shows the location of the 123 Iowa lakes used in the analysis, along with a HUC8 watershed. Panel (2.1b) shows the HUC12 watersheds and lakes contained within the same HUC8 watershed.

For this paper I focus on land devoted to corn, soybeans, and grassland [4]. An issue with using cropland devoted to corn and soybean use is they are highly correlated (.80). To avoid multicollinearity I sum the two land uses into a single variable labeled *crops*. Since official CRP enrollment numbers are only available at the county level, I include grassland as a proxy for the effect of CRP land on water quality.

Data for precipitation and temperature were calculated from Oregon State's PRISM dataset. PRISM provides the daily precipitation and temperature for 30km by 30km grid cells that cover the continental U.S. To find the annual precipitation for an individual HUC12 I sum the daily data for each PRISM grid cell across the watershed, and then sum the daily values over the year. To find the average annual temperature for each HUC12 I average daily temperatures across PRISM grid cells, and then average the daily values over the year.

Table (2.1) provides a description and summary statistics for the variables included in the analysis.

[4] Although CDL data includes other crops such as wheat, they are more difficult to accurately identify. Data on CDL accuracy can be found at http://www.nass.usda.gov/research/Cropland/sarsfaqs2.htm

Table 2.1: Summary Statistics

Variable	Mean	SD	10th Percentile	90th Percentile
Water Quality Model:				
CTSI (0-100)	60.81	9.89	47.00	73.00
Crops (thousand acres)	12.77	7.13	5.06	22.20
Grass (thousand acres)	5.58	3.72	1.55	10.52
Precipitation (cm)	39.26	16.60	21.21	58.96
Temperature (celsius)	9.57	2.37	6.48	12.95
N = 1,739				
Cropland Model:				
Crops (thousand acres)	13.37	7.40	4.29	23.75
Expected Price (index)	1.70	0.55	1.00	2.51
Fertilizer (index)	71.97	26.25	37.40	101.40
Fuel (index)	68.73	23.45	33.40	99.30
Precipitation (cm)	38.77	16.33	22.20	55.75
Temperature (celsius)	9.59	2.25	6.71	12.64
N = 27,472				

Notes: Number of observations, mean, standard deviation, 10th, and 90th percentile of variables used in the main regression analysis. These data cover the years 2001-2016 in Iowa, excluding 2008 for the water quality model. CTSI stands for Carlson's Trophic State Index.

Water Quality Model

I use the following dynamic panel data model to estimate the effects of land use on water quality:

$$Q_{i,j,t} = \mu_i + \beta_1 Q_{i,t-1} + \beta_2 C_{j,t} + \beta_3 G_{j,t} + \beta_4 W_{j,t} + \beta_5 T_{j,t} + \lambda_t + \epsilon_{i,t}$$

The dependent variable, $Q_{i,j,t}$, is a measure of water quality for lake i in HUC12 j at time t. For the measure of water quality, I use Carlson's Trophic State Index (CTSI). CTSI is an index of water quality from 0 to 100, where an increase indicates a degradation in water quality. I include the lag in Q to account for the stock effects of lake water; the coefficient is expected to be positive as a certain amount of nutrients in a lake carry over across years. Including a lag of Q implies both short and long term impacts from the other right hand side variables on water quality. The short term, i.e. contemporaneous impacts are the estimated coefficients on the variables, while the long term impacts are these coefficients multiplied by the dynamic multiplier (Greene, 2000, p.416):

$$\sigma = \frac{1}{1-\beta_1} \tag{2.1}$$

The variables C and G represent year t, HUC12 j's acres of cropland and grassland, respectively. The main coefficient of interest is β_2, which measures the short term marginal change in water quality due to an increase in cropland. Since an increase in CTSI represents lower water quality, β_2 is expected to be positive due to nutrient runoff. Of secondary interest is the marginal increase in water quality due to an increase in grassland, β_3, which could be considered a proxy to the effect of CRP on water quality. A negative sign on β_3 would indicate beneficial qualities of increased grassland near lake water. Since the CRP program requires active cropland to be retired, the total short term effect of CRP on lake water quality is ($-\beta_2$ + β_3).

I control for the effect of weather on water quality by including annual measures of precipitation, W, and temperature T. Although the focus of the paper is on the effect of crops on water quality, the effect of weather on water quality is an important and complicated topic in itself. For example, it is not clear a priori what the sign of these weather effects will be; increased rainfall, for example, can dilute existing nutrient levels, but can also increase nutrient runoff from nearby farms. Several papers have also highlighted the importance of studying the effects of weather on water quality, given the predicted increased variation in weather due to climate change (Delpla et al., 2009). The coefficient estimates on precipitation and temperature help shed light on these issues.

I control for time invariant, unobservable variables through lake level fixed effects, μ_i, thus the coefficients are identified by the variation of the data within a lake. The unobservable variables could be, for example, geographic features that are fixed over time, such as soil quality, slope, or surface area, and it can also include permanent man-made structures that can alter the flow of water to lakes, such as tile drains. Year dummy variables control for unobserved trends over time.

Each HUC12 watershed is contained within larger watersheds which share drainage properties. To control for correlation between HUC12's within the same drainage area, I cluster standard errors at the HUC8 level.

Cropland Response Model

To estimate the response of cropland to crop prices I estimate the following Nerlovian partial adjustment model (Nerlove, 1956):

$$C_{j,t} = \eta_j + \beta_1 C_{j,t-1} + \beta_2 \mathrm{E}[P_{j,t}] + \beta_3 T_{j,t-1} + \beta_4 W_{j,t-1} + \beta_5 F_{t-1} + \beta_6 G_{t-1} + \beta_7 t + \epsilon_{j,t}$$

This model assumes that a representative farmer in watershed j make spring acreage decisions based on last year's acreage, climate and operating costs, as well as the expected crop prices during fall harvest. Operating costs consist of fertilizer, F, and fuel, G.

The variable of interest is β_2, which represents the marginal change in cropland due to an increase in expected prices. For the expected price I construct a Laspeyres price index using futures prices on corn and soybeans, along with observed soybean and corn acreages (Huang et al., 2010; Evans and Potts, 2015):

$$E[p_{j,t}] = \frac{\sum_c (E[P_{j,c,t}] * C_{j,c,t_0})}{\sum_c (E[P_{j,c,t_0}] * C_{j,c,t_0})}, \tag{2.2}$$

where $c \in \{\text{corn, soybeans}\}$. Figure (??) shows the crop price index in Iowa from 2001-2016, averaged over HUC12 watersheds.

Estimation

Dynamic panel data models with fixed effects suffer from the well known "Nickell bias", which results from the *within* transformation that subtracts the time mean from each group μ_i in order to remove the fixed-effects (Nickell, 1981). In a dynamic model, this will cause the lagged, transformed dependent variable to be correlated with the error term, violating the assumed orthogonality condition. One solution is to use the Arellano-Bond model (henceforth

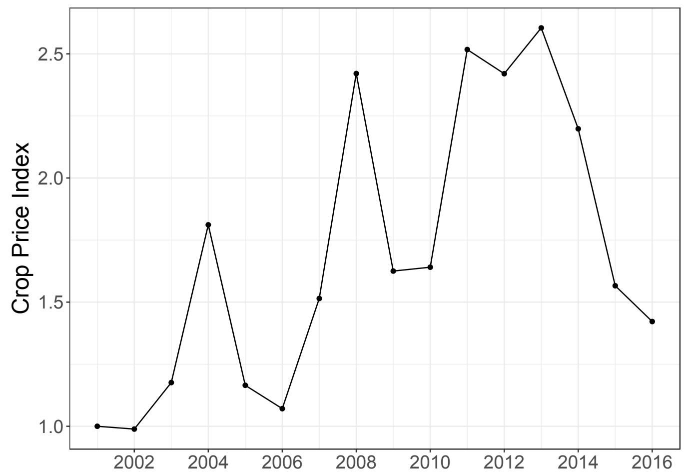


Figure 2.3: Crop Price Index, 2001-2016

Notes: Price index for corn and soybeans in Iowa, averaged over HUC12 watersheds from 2001-2016.

abbreviated as AB), also known as the“difference GMM” (Arellano and Bond, 1991), which constructs instruments for the lagged dependent variable using transformations of the data.

The model's nickname comes from using first differences of the data to remove fixed effects. However, when there are gaps in the data, as is the case with the CTSI measurements, they can result in a significant loss of observations. For example, all lakes in the data set are missing the year 2008, so neither $\Delta CTSI_{i,2008}$ or $\Delta CTSI_{i,2009}$ can be included in the estimation. Instead of the first difference transformation, I employ the forward orthogonal deviations (FOD) transformation, where the mean of all future observations of a variable is subtracted from the current observation for each year. This purges the fixed effects and allows for more observations than the first differences in an unbalanced panel.

AB estimates are typically estimated in both one and two-step variants. The two-step model uses a weighting matrix that is the inverse of an estimate of $Var(z'e)$, where z is the vector of instruments. This is the optimal weighting matrix in the sense that it is asymptotically efficient. However, in finite samples the two-step estimates have been shown to be biased downward. To fix this, I employ the finite sample bias correction described in Windmeijer (2005).

Results

Table (2.2) displays the results for the water quality model. All variables are estimated in log form, so the coefficients are interpreted as the elasticity of water quality with respect to each variable. Column (1) shows the estimated coefficients using the Arellano-Bond model, which instruments for the endogenous lagged dependent variable. The coefficient on the lag of CTSI is statistically significant and positive, which provides evidence that water quality conditions persist over time. The coefficient on *Crops* is positive and statistically significant, indicating that an increase in cropland increases CTSI and therefore lowers water quality. The elasticity is 0.0538 in the short run and, using equation (2.1), 0.0727 in the long run. I do not find a statistically significant effect of grassland on lake water quality.

The coefficient on precipitation is positive and statistically significant. This indicates that the overall effect of precipitation on water quality is detrimental. I do not find evidence of an effect of temperature on lake water quality.

Columns (2) and (3) of table (2.2) display the results of OLS and fixed effects (FE) estimation for comparison. The OLS estimates may suffer from omitted variable bias since it does not control for time invariant fixed effects; this could explain the larger magnitude of the coefficient on the lag of CTSI. As with the Arellano-Bond model, the effect of precipitation is positive and significant. The fixed effects model, which does not instrument for the endogenous lagged variable, obtains a similar estimate for the coefficient on the lagged variable and a marginally significant, positive effect for precipitation. Neither the OLS or fixed effects models find a significant effect of cropland or grassland on water quality.

Table (2.3) displays the results for the cropland response model. Again, the variables are logged so that the estimates can be interpreted as the elasticity of cropland to a specific variable. All variables show a statistically significant effect on cropland. The variable of interest, price, shows the expected positive relationship with cropland. The magnitude of the elasticity of cropland to prices, 0.066, is small but comparable to other estimates from the literature (Evans and Potts, 2015; Barr et al., 2011). Using the dynamic multiplier, the long run elasticity is 0.104.

Table 2.2: Regression Results: Water Quality Model

	(1) AB	(2) OLS	(3) FE
L.CTSI	0.260***	0.743***	0.300***
	(0.0410)	(0.0261)	(0.0529)
Crops	0.0538***	0.00757*	0.0251
	(0.0179)	(0.00438)	(0.0252)
Grass	0.00321	-0.00846	-0.00444
	(0.00832)	(0.00530)	(0.00916)
Prec.	0.0460***	0.0312***	0.0406*
	(0.0139)	(0.00928)	(0.0201)
Temp.	0.0324	0.0307*	0.0185
	(0.0261)	(0.0157)	(0.0304)
N	1,484	1,607	1,607

Notes: Arellano-Bond (AB), OLS, and fixed effects (FE) coefficient estimates and standard errors. Each observation is a water quality measurement from a specific lake in Iowa. Data includes the years 2001-2016, excluding 2008. All estimates include year fixed effects. CTSI stands for Carlson Trophic Secchi Index. Crops is equal to the sum of corn and soybean land. Standard errors are clustered by HUC8 watershed.

The results for the control variables are mostly intuitive. The weather variable coefficients indicate that cropland decreases in response to increases in the previous year's temperature, while it increases in response to the previous year's precipitation. The magnitude of these responses are small and roughly equal. Increases in last year's fuel costs have a negative effect on this year's cropland. On the other hand, increases in the cost of last year's fertilizer have a positive effect on this year's cropland. Although this result is not intuitive, it has been found in other research (Evans and Potts, 2015; Huang et al., 2010).

Columns (2) and (3) show the OLS and fixed effects results of the cropland response model. The signs and magnitudes of the coefficients are fairly similar across all three models, with the

Table 2.3: Regression Results: Cropland Model

	(1) AB	(2) OLS	(3) FE
L.Crops	0.304***	0.951***	0.670***
	(0.0798)	(0.00697)	(0.0278)
Price	0.0525***	0.0466**	0.0479**
	(0.0158)	(0.0222)	(0.0200)
L.lnPrec	0.00166	0.0235***	0.0160
	(0.0139)	(0.00844)	(0.0127)
L.lnTemp	-0.167***	-0.0306***	-0.0798***
	(0.0263)	(0.00889)	(0.0151)
L.Fuel	-0.106***	-0.106***	-0.101***
	(0.0224)	(0.0290)	(0.0273)
L.Fert.	0.0446***	0.0828***	0.0686***
	(0.0112)	(0.0186)	(0.0147)
N	24,033	25,750	25,750

Note: Arellano-Bond (AB), OLS, and fixed effects (FE) coefficient estimates and standard errors. Each observation is a HUC12 watershed in Iowa. Data includes the years 2001-2016. Crops is equal to the sum of corn and soybean land. Standard errors are clustered by HUC8 watershed.

exception of the coefficient on the lag of crops, where the FE and OLS estimates are larger than the Arellano-Bond estimate.

Application: The Renewable Fuel Standard

This section uses the previous elasticity estimates to measure the impact of the Renewable Fuel Standard (RFS) on lake water quality in Iowa. Since the RFS was enacted in 2005, I focus on the effects over 2006 to 2016, which is the most recent year with available data. The RFS mandated a large increase in ethanol, which is equivalent to a large increase in demand for corn since it is the primary feedstock. I follow the approach of Evans and Potts (2015) and use the price effects of this shock in demand to corn to connect the RFS to lake water quality. Specifically, I calculate the percent change in water quality using the following formula:

$$\%\Delta WQ = \%\Delta\text{Price} * \epsilon_{wq,c} * \epsilon_{c,p}, \tag{2.3}$$

where $\epsilon_{wq,c}$ represents the elasticity of water quality with respect to cropland, and $\epsilon_{c,p}$ represents the elasticity of cropland to prices. I calculate (2.3) using both the short and long run elasticities, which can be considered lower and upper bounds.

For the change in price, I use estimates from Hausman et al. (2012). Using a structural vector autoregression (SVAR) model, the authors estimate that an increase in the demand of corn acreage for ethanol of one million acres increases the price of corn and soybeans by 0.08 and 0.04 cents per bushel, respectively. According to data from USDA ERS, 1.603 billion bushels of corn were used to produce ethanol in 2005/2006, compared to 5.206 billion in 2015/2016, an increase of 3.603 billion bushels. Over the same period, the national average corn yield was 156 bushels per acres. Thus the shock in demand is equivalent to an approximate 23 million acre increase in demand for corn acreage. The above estimates from Hausman et al. (2012) imply this increase in demand would increase the price of corn by \$1.84 per bushel and the price of soybeans by \$0.92 per bushel. I use these changes to calculate the counterfactual price index for each HUC12 in 2016. The average HUC12 experienced an approximate 58% increase in the price index as a result of the RFS.

Finally, I calculate the percentage change in lake water quality using both the short and long term elasticities estimated from the previous analysis. Figure (2.4) The average lake experienced a 0.13% increase in CTSI in the short run, and a 0.27% increase in the long run. Thus, due to the very inelastic responses of lake water quality to cropland, and from cropland to crop prices, I find a negligible effect of the RFS on lake water quality.

Conclusion

U.S. agricultural and energy policies often have direct and indirect effects on the environment. In particular, policies which affect agricultural land use can alter lake water quality

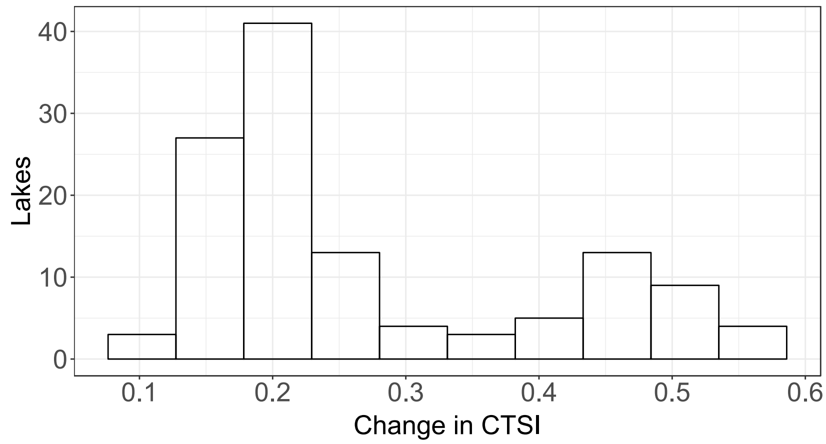


Figure 2.4: Estimated Change in CTSI due to the RFS

Notes: Distribution of the estimated change in CTSI due to the Renewable Fuel Standard (RFS) across 123 lakes in Iowa. Increases in CTSI represent a decrease in lake water quality, as represented by the clarity of the water.

through increased nutrient runoff. It is important to estimate these impacts in order to undertake thorough cost-benefit analyses of these policies.

This study focuses on estimating the effect of land use change on lake water quality in Iowa. High quality lake water measurements over 15 years are combined with satellite imagery and PRISM weather data to create a unique panel data set. Using a dynamic panel data model, I estimate the elasticity of CTSI to cropland to be 0.05% in the short run, and 0.07% in the long run, indicating that increases in cropland decrease lake water quality by a small amount. I also find a positive and significant coefficient on the lag of the dependent variable, which is evidence of a stock effect of lake water quality over time.

A second model estimated the elasticity of cropland to crop prices to be 0.066. Using these two elasticities, I estimate that the Renewable Fuel Standard decreased water quality by between 0.13 and 0.27%. The estimates may represent a lower bound since the paper only studies land use change along the extensive margin. Rather than expand cropland, farmers may alter crop rotations in favor of corn as a result of the RFS. Since corn requires a relatively high amount of fertilizer, the actual impact on water quality may be higher.

CHAPTER 3. OPTIMAL DEER MANAGEMENT IN THE MIDWEST

Abstract

Midwest deer management is a complicated balancing act. Hunters gain benefits from larger populations, but a larger population imposes increasing costs on farmers, through crop depredation, and on drivers through deer vehicle collisions. This problem is complicated by the growth dynamics and random growth shocks to deer populations. This paper formulates a dynamic model of deer management which explicitly incorporates costs, growth, and uncertainty into the problem. We [1] calibrate our model's parameters to reflect Iowa's conditions and simulate optimal deer management over 50 years. Our results show that the optimal deer herd size is approximately 189,000. This result is 8% higher than the state's current goals.

JEL codes: Q28, Q26, Q57

Keywords: optimal harvests, recreational licenses, renewable resources

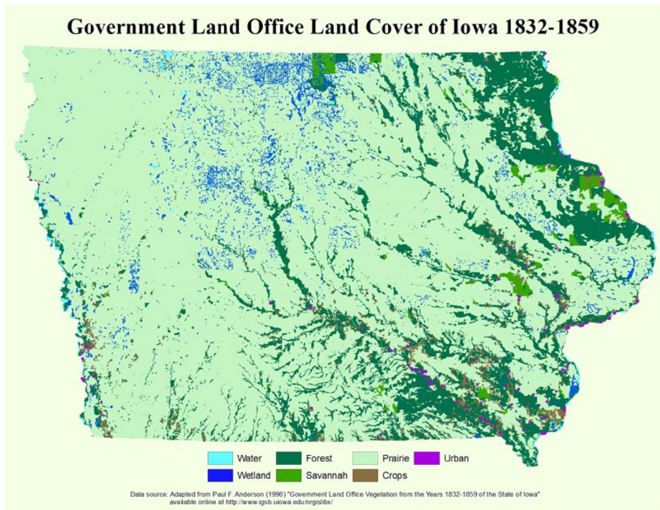
[1]Co-Author: Dr. Quinn Weninger, Iowa State University

Introduction

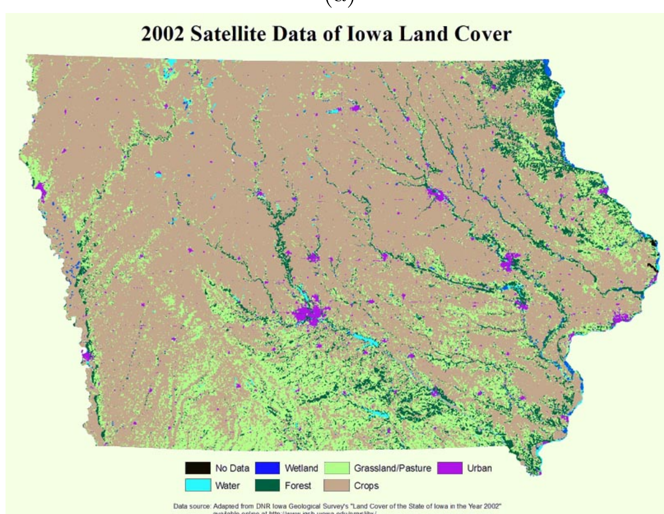
The U.S. Midwest faces a complex tradeoff between agricultural land use and the conservation of existing ecosystem services. Much of the natural landscape in some states has already been drastically altered to increase agricultural production. Iowa, for example, was composed of prairies, forests, and wetlands prior to 1860 (figure 3.1a), but is now almost entirely cropland and grassland/pasture (figure 3.1b). These changes have been beneficial to the whitetail deer, who have found themselves with no natural predators and an almost endless supply of food in the form of row crops (Paddock and Yabsley, 2007). With these favorable conditions, effective deer management is critical to maintaining a healthy balance with the human population. Left unchecked, the deer population in Iowa could double in as few as three years (IDNR, 2016).

The Iowa Department of Natural Resources (IDNR) is responsible for managing the white-tail deer population in the state of Iowa. According to IDNR- "Deer management in Iowa may be characterized as trying to maintain a balance between the public's demand for hunting and viewing opportunities with a need to keep deer numbers compatible with agricultural interests, highway safety, and habitat limitation" (IDNR, 2009). Thus an essential component of setting population targets is balancing both the benefits and costs of deer. Most of the benefits and costs can be classified as non-marketed factors, i.e., goods that are not traded in standard markets. This implies that market signals do not determine socially optimal deer herd sizes and complicates the task.

Following several public meetings with relevant stakeholders, the Iowa DNR began an effort in 2007 to reduce the size of the deer herd to levels observed in the mid to late 1990's. The overall consensus was that this time period represented an optimal deer population when all interests were considered. The main regulatory instrument used to affect population size is the issuance of deer hunting licences. Each year, following public forums as well as analysis of harvest and population trends, the IDNR adjusts the quota for buck and doe licenses by county to achieve its population target. In 2016, the IDNR stated that their mid to late 1990's target had been achieved on a statewide basis (IDNR, 2016). A crucial question is whether the stated population target maximizes the economic welfare of Iowa's citizens.



(a)



(b)

Figure 3.2: Iowa Land Uses Over Time

Notes: (a) Land uses in Iowa, 1832-1859. (b) Land uses in Iowa, 2002. Used with permission of the University of Iowa Office of the State Archeologist.

This paper develops a dynamic bio-economic model of the deer management problem facing the IDNR. We characterize the static and dynamic trade-offs involved in different deer population sizes. The dynamic model allows us to not only characterize the optimal levels of does and bucks, but also the optimal path of license prices that can maximize welfare over time. Our goal is to inform the management debate that has surfaced recently between Iowa hunters, who have voiced concerns about dwindling deer populations and the DNR policy to reduce herd size to mid to late 1990's levels (IDNR, 2009). To our knowledge, this is the first paper that uses a stochastic dynamic model to balance the benefits of hunting and viewing deer with the costs of crop depredation and vehicle collision. Attempting to balance these interests closely aligns with the complex decisions the DNR faces in many Midwest states.

Although the bio-economics literature is large, relatively few models have been applied to deer management. Our model most closely resembles Cooper (1993), who uses a dynamic bio-economic model to determine the optimal number of deer licenses in California. By assuming different values for three different ages of bucks, his model shows that the optimal harvesting strategy is periodic as the harvest peaks every three years before falling to allow bucks to grow. His model only considers the benefits of deer, which reflects the stark environmental, and perhaps cultural differences between California and Iowa. In fact, the main problem in California is that deer population numbers have dwindled as their natural habitat gets converted into developed land (CDFW, 2016). In Iowa, an additional problem is containing the deer population to avoid high costs from crop depredation and deer vehicle collisions, while retaining enough deer to satisfy hunters. Therefore, incorporating costs into the social welfare function is critical. We also differentiate our model by incorporating uncertainty in the growth function, conceptualized as annual weather shocks which can drastically affect deer survival rates (Loison and Langvatn, 1998). Introducing uncertainty captures a key component of the inherent difficulty of managing wildlife resources, and allows us to show the range of policy values available to the regulator along the optimal dynamic path.

Keith and Lyon (1985) use a bio-economic model to estimate the value of deer, but do not differentiate between bucks or does or analyze the dynamics of herd management. Schwabe et al. (2002) use a bio-economic model to compare different policies meant to reduce deer vehicle collisions, but focus on the transition equations and do not include a welfare function. Hussain and Tschirhart (2010) study the optimal allocation of deer licenses with uncertainty from hunting and lottery success, but only study the steady state and do not include costs from crop depredation or deer vehicle collisions. Rondeau and Bulte (2007) use a general equilibrium bio-economic model to analyze different compensation schemes for crop damage from wildlife. Rondeau and Conrad (2003) use a dynamic bio-economic model to analyze optimal culling of urban deer. However many states, including Iowa, rely solely on hunting to control the deer population. In addition, although they take into account both benefits and costs, their model does not distinguish between males and females. This is an important consideration with hunting, as hunters typically place a higher value on bucks than does due to their antlers.[2]

The following section describes the dynamic model we will use to study Midwest deer management. This is followed with a numerical application where the parameters are calibrated to Iowa data. The paper ends with a discussion on the policy implications.

Model

Our model will focus on the size of a Midwest deer herd across discrete time periods, which we take to be a single year. To simplify the model, each year is divided into two sub-periods. The model will more naturally conform to hunting and deer reproduction by setting the beginning year date to early Fall, say September 1st. We assume that all hunting, hunting mortality, and crop depredation take place during the first sub-period. Herd natural mortality, car collision mortality, and reproduction occur during the second sub-period. These assumptions and timing conventions simplify the model presentation while capturing the key forces that impact deer herd size over time.

We use $d_{b,t}$ to denote the number of male deer (bucks) and $d_{d,t}$ to denote the number of female deer (does) in the herd at the beginning of each period. Per-period net benefits are an inclusive and exhaustive measure of the total benefits less total costs associated with the herd. The net benefit function is described in detail below.

[2] This statement is based on anecdotal evidence as we are not aware of any peer reviewed studies which attempt to measure both the value of a buck and a doe.

The manager indirectly controls herd size by issuing hunting licenses for both male and female deer. We use $k_{i,t}$ to denote the number of male and female deer that are harvested in year t. Deer hunting is a random process. Our model applies at the aggregate level and therefore randomness in hunting success will average out across large numbers of hunters. We therefore assume a deterministic relationship between the number of tags sold and corresponding hunting mortality. Let $q_{i,t}$ represent the amount of doe $(i = d)$ and buck $i = b$ hunting licenses in year t. Deer harvest is calculated as:

$$k_{i,t} = \pi q_{i,t},$$

where π is the exogenous hunter success rate, i.e. kills per license.[3] We use $e_{i,t}$ to denote period t escapement for $i \in \{\text{d,b}\}$:

$$e_{i,t} = s_{i,t} - k_{i,t}$$

The escapement breeds and produces fawns during the second sub-period. Let α represent annual new fawns per doe, and let k represent the carrying capacity for deer in Iowa. We assume the following growth function for new fawns:

$$f_{i,t} = \alpha e_{d,t} \left(1 - \frac{e_{d,t} + e_{b,t}}{k}\right) \tag{3.1}$$

Equation (3.1) specifies a quadratic growth function where the potential number of new fawns per year, αe_d, is adjusted according to the ratio of herd size to carrying capacity. If herd size, $e_d + e_b$, is far below carrying capacity k, the herd will grow rapidly. If herd size exceeds carrying capacity, the herd will actually lose deer instead of gaining new fawns. This captures the idea that at certain herd sizes the existing resources, e.g. land and food, are not adequate to support the population.

Next, we incorporate random shocks z_t and deer vehicle collisions into the model. By random shocks, we are thinking of annual fluctuations in weather, which impact herd size but exhibit no serial correlation. Let $z_t \in \{\overline{z}, \underline{z}\}$ represent this random, i.i.d shock. Finally, let τ

[3]Endogenous hunter success, i.e. where α depends on the deer and hunter population, may be incorporated into future extensions of this paper.

represent the proportion of the deer herd killed in car collisions before the next period begins. The sex specific transition equations are:

$$d_{i,t+1} = z_t(1-\tau)\left(e_{i,t} + \frac{f_{i,t}}{2}\right). \tag{3.2}$$

We next determine hunting mortality, which is a function of the regulator's choice of the number of hunting licenses as well as the demand for hunting.

Regulation and License Purchase

The deer herd manager regulates hunting activity by issuing species-specific hunting licenses. It is common to structure licensing fees as a two-part tariff. Hunters must purchase a general hunting license- effectively an entry fee- plus species-specific tags. Regulations require tags be purchased in advance of the hunt. The tag purchase is non-refundable; the price is paid regardless of whether an animal is actually harvested. Harvesting success is by no means assured and therefore tags regularly go unfulfilled.

Let a_t denote the price of a general hunting license, and let $p_{b,t}$ and $p_{d,t}$ denote the price per buck and doe license, respectively. The demand for licenses is given as:

$$q_{i,t}(a_t, p_{i,t}) \tag{3.3}$$

We assume that q_i is non-increasing in a_t and strictly decreasing in tag price.

Deer Herd Benefits and Costs

We assume that the total benefits that accrue to hunters from deer hunting is captured by the sum of the consumer surplus from doe and buck hunting licenses:

$$B(a_t, q_b, q_d) = \sum_{i=d,b}\left(\int_0^{q_{i,t}} q_{i,t}^{-1}(a_t, s)ds\right) \tag{3.4}$$

The term $q_{i,t}^{-1}$ in equation (3.4) is the inverse demand function and s is the variable of integration. assumes that hunters enjoy greater benefits from healthy deer populations, which is strongly related to the number of does and bucks harvested per year and is a function of deer hunting license. In addition, the hunting benefits function are separable by does and bucks, which captures the generally accepted notion that hunters tend to place a greater value on harvesting bucks.

To further simplify the model, we assume the regulator chooses a_t in a way that does not affect marginal tag purchases. In other words, in each period the regulator sets a_t sufficiently low so that the marginal hunter is willing to enter the recreational market and purchase the same number of tags they would have purchased had a_t been set to zero. Hereafter we suppress a_t and write the hunting benefit function as $B(q_b, q_d)$.

The two major costs associated with deer are the cost of automobile collisions, C^c and crop depredation, C^d. We assume that both functions are non-decreasing, differential, and concave in the total deer escapement. To simplify notation, let $e = e(q_b, q_d) = e_d(q_d) + e_b(q_b)$. Then the equation for total cost function is:

$$C(q_b, q_d) = C^c(e) + C^d(e). \tag{3.5}$$

[4]

Management goal

For the purposes of the analysis that will follow, we will assume that a social planner will seek to manage the deer herd with the goal of maximizing the present value of the stream of net benefits, NB,

$$NB(q_{b,t}, q_{d,t}) = \max_{q_{b,t}, q_{d,t}} \sum_{t=0}^{\infty} B(q_{b,t}, q_{d,t}) - C(q_{b,t}, q_{d,t}) \tag{3.6}$$

The optimization problem is subject to the stock transitions equations (3.2).

Equation (3.6) can be reformulated as a Bellman equation. Hereafter, we adopt standard notation from the dynamic programming literature and use a prime to denote the value of a

[4]Notice that we have not included total payments for tags in the above expressions. The reason is that hunting license fees, given as a_t times the number of participating hunters, and tag revenue, $\sum_i p_{i,t} q_{i,t}$, are simply transfers from the hunter population to the regulator and therefore do not factor into changes in overall social welfare.

variable one period ahead; absent a prime will denote current period values. The distinction between sub periods is not required. The social planner's Bellman equation is

$$V(d_b, d_d, z) = \max_{q_b, q_d} \left\{ B(q_{b,t}, q_{d,t}) - C(q_b, q_d) + \beta E_{z'} \left[v(d_b', d_d'; z') \right] \right\} \tag{3.7}$$

Equation (3.7) contains two continuous state variables, d_b and d_d, which represent the size of the buck and doe herds, respectively, at the current period. The additional state variable z is the uncertainty parameter. Solving this equation, subject to equations (3.2), will reveal the optimal choice functions, $q_b^*(d_b, d_d, z)$ and $q_d^*(d_b, d_d, z)$ and the optimal value function, $V^*(d_b, d_d, z)$

Functional Form

We assume a linear inverse demand curve for buck and doe hunting licenses,

$$p_i = \eta_i - \gamma_i q_i, \ i \in \{b, d\} \tag{3.8}$$

Integrating (3.8) with respect to i gives the consumer surplus functions:

$$S(q_i) = \eta_i q_i - \frac{\gamma_i}{2} q_i^2, \ i \in \{b, d\} \tag{3.9}$$

The benefit function is therefore the sum of $S(q_i)$ across i.

We also assume a linear cost function. Let C^c and C^d be the marginal cost of deer vehicle collisions and crop depredation, respectively. Recall that ψ is the proportion of the deer herd who are involved in a deer vehicle collision. The cost function is then

$$C(e) = (\tau C^c + C^d)e \tag{3.10}$$

To gain intuition into the social planner's problem, we examine the first order conditions of equation (3.7) using these specified functional forms. Substituting equations (3.2) into equation (3.7) and taking derivatives with respect to q_b and q_d gives the following equations:

$$\eta_b - \gamma_b q_b^* + (\tau C^c + C^d)\pi = \beta E[V_b] z(1-\tau)\pi(-1 + \frac{\alpha}{2k} e_d) \tag{3.11}$$

$$\eta_b - \gamma_d q_d^* + (\tau C^c + C^d)\pi = \beta E[V_d] z(1-\tau)\pi(-1 - \frac{\alpha}{2} + \frac{\alpha}{k} e_d + \frac{\alpha}{2k} e_b) \tag{3.12}$$

Equation (3.11) shows how the social planner balances the marginal costs and benefits of bucks over time. The first two terms on the left hand side represent the increase in consumer surplus due to a unit increase in buck licenses. The second term represents the decrease in current period costs due to an unit increase in licenses; since π is the success rate of hunters, an increase of one buck license decreases the herd by π, which is multiplied by the marginal cost of deer in parenthesis. The right hand side of equation (3.11) is the discounted expected value of the deer herd tomorrow, given a unit increase in buck licenses today. The discounted, marginal expected value of a buck tomorrow is corrected by a series of terms which take into account the shock to the deer population, z, deer vehicle collisions, $(1-\tau)$, hunter success, π, and deer growth. Deer growth is composed of two terms, the first of which is negative due to the direct effect of decreasing the buck population. The second term, $\frac{\alpha}{2k}e_d$, represents the marginal increase in fawns due to less pressure on the carrying capacity from a decrease in bucks.

Equation (3.12) largely mirrors equation (3.11), except for two additional terms, $-\frac{\alpha}{2}$ and $\frac{\alpha}{k}e_d$, in the expression for deer growth. The first term represents the indirect loss of fawns due to an increase in doe licenses; in other words, after hunting season there are less does to reproduce. The second term is the increase in fawns due to less pressure on the carrying capacity from the indirect decrease in fawns.

Given the complexity of the model, an analytical solution is intractable. In the following section we present a numerical example using data from Iowa.

Application to Iowa

To parameterize the consumer surplus functions we draw on previous estimates from the literature. A 2006 U.S. Fish and Wildlife report estimated the economic value of deer hunting in each U.S. state. This value represents the area under an estimated demand curve for deer

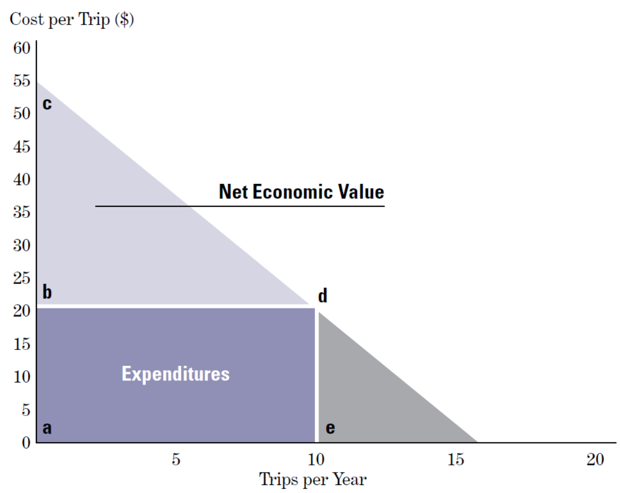


Figure 3.3: The economic value of deer hunting. Reprinted from Aiken (2006).

hunting trips, less expenditures (see figure 3.3). In other words, it is the net consumer surplus from deer hunting trips. For Iowa, a hunter in Iowa received an average of \$543 per year of economic value from deer hunting (Aiken, 2006). With 280,000 deer licenses sold (IDNR, 2009), the total economic value of Iowa deer hunting in 2006 was therefore \$155.6 million dollars. We therefore choose the demand parameters η_1, η_2, γ_1, and γ_2 to match this value. In addition, we choose to calibrate the intercept parameter for bucks as higher than does, in order to capture the greater value hunters place on bucks. For this exercise we leave the slopes for bucks and does equal. Our baseline values of $\eta_1 = .9$, $\eta_2 = .65$, $\gamma_1 = .1$, and $\gamma_2 = .1$ produce a total consumer surplus of \$155.4 million dollars, which is reasonably close to the above estimate. This calibration is necessarily ad hoc due to a lack of detailed information on the demand for hunting both bucks and does in Iowa; improving the accuracy of these parameters is left for future research.

Cost related to deer accrue to Iowa residents through deer vehicle collisions and through crop depredation. There were about 15,000 deer killed in 2006 through vehicle collisions, representing about 0.3% of the population (IDNR, 2009), thus we set $\psi = 0.03$. From 2002 to 2007, the Grinnell Mutual insurance company reported that the average claim for a deer related accident was \$2,135, thus C_c=\$2,135. Using these numbers, deer caused approximately \$34,160,000 of vehicular damage in 2006.

Table 3.1: Iowa Parameters

Definition	*Value*
Initial bucks	$d_0 = 160,000$
Initial does	$b_0 = 240,000$
Intrinsic growth rate	$\alpha = 1.85$
Demand function intercepts (bucks,does)	$\eta = (.9, .65)$
Demand function slopes (bucks,does)	$\gamma = (.1, .1)$
Exogenous growth shock	$z \in [0.67, 1.33]$
Deer vehicle collision rate	$\tau = 0.03$
Cost per collision	$\$2,135$
Corn consumed per deer (bushels/year)	$a_c = 3.4$
Soybeans consumed per deer (bushels/year)	$a_s = 4$
Price of corn (per bushel)	$p_c = \$4$
Price of soybeans (per bushel)	$p_s = \$6$
Carrying capacity	$k = 800,000$
Discount factor	$\beta = 0.95$
Hunter success rate	$\pi = .45$

Note: Baseline parameter values for numerical approximation. All dollar values are in 2006 U.S. dollars.

Data on crop depredation due to deer in Iowa is limited. Hunters can obtain a limited number of depredation licenses to hunt on a producer's property, and properties are eligible for additional licenses if estimated damages from deer exceed \$1,000 (IDNR, 2009). This means that official depredation complaints underestimate the total damages as many will go unreported. Instead of using depredation reports, we compute our own estimates based on the average deer diet. The average white tailed deer consumes about 4 pounds of food per day. According to Iowa DNR, about 78% of an Iowan deer's diet comes from crops [5], so each deer consumes an estimated 3.12 pounds of crops per day. Farmers in Iowa typically plant crops in April and harvest in October. Assuming the first crops don't sprout until May and harvest lasts through the end of October, crops are available for deer consumption for about 152 days. Thus each deer consumes about 474 pounds of crops per year. Assuming half is corn and half is soybeans, a deer consumes 3.4 bushels of corn and 4.0 bushels of soybeans per year [6]. We use the average 2006 prices of corn and soybean, which are \$3.03 and \$6.58, respectively [7]. Thus

[5] "White-tailed deer" pdf link at http://www.iowadnr.gov/portals/idnr/uploads/Hunting/

[6] One bushel of corn ears is 70 pounds; one bushel of soybeans is 60 pounds.

[7] https://www.nass.usda.gov/

one deer consumed approximately \$37 worth of crops in 2006, and the herd of 400,000 deer caused approximately \$14,800,000 worth of crop depredation.

The natural growth of the deer population is determined by the intrinsic growth rate, γ, the carrying capacity, K, and the stochastic shock z. Does typically give birth to between one and two fawns per year. For the DNR's population projection model, the growth rate was set to 1.85, which we also use in this paper. Finally, Iowa DNR note that Iowa land can support 1-3 healthy deer per square mile, but this number could be as high as 100 depending on the environment (IDNR, 2009). Given the highly uncertain nature of this parameter, we set an initial carrying capacity of 800,000 but test the sensitivity of the results to this parameter.

To solve the model we use the collocation method; that is, we approximate the value function using a Chebychev polynomial. We make an initial guess for the collocation coefficients and solve equation (3.7), then iterate on this guess until the coefficients converge.[8] We use a 15td degree polynomial approximate; higher degrees of approximation did not yield a significant increase in accuracy. We tested the result using different sets of initial guesses and the model converged to the same coefficients. A plot of the residuals of the value function show that they stay with a range of plus or minus 10e-2, with less accurate results when the buck population is near zero. These results indicate a reasonably accurate approximation.

Numerical Results

Figure (3.4a) shows a three dimensional image of the approximated value function, while figure (3.4b) shows a contour map. The value function has a concave shape, owing to the increasing costs of larger deer herds. The value function sharply decreases after 400,000 does, as the fawn they produce impose increasing costs upon agriculture through crop depredation, and drivers through deer vehicle collision. The value function displays less curvature as bucks increase as they are not as critical a factor in deer herd growth, and thus also growth in costs.

Figures (3.6a) and (3.6b) show the optimal policy function for doe and buck hunting licenses, respectively. Both curves display a linear shape. The optimal doe policy, however, also display

[8]Code was written in Matlab using the compEcon package; see Miranda and Fackler (2004)for more details on the collocation method and compEcon.

a flat portion near zero, since hunting does to very low numbers decreases the future herd size, implying decreasing future hunter benefits as well.

Simulation

This section uses the numerical approximations to the buck and doe policy functions to simulate the social planner's optimal choices over time. We choose as a start date 2006, which is the year Iowa DNR began a policy to reduce deer numbers to approximate levels in the 1990's. Thus we can examine how close the model's solution is to the actual path taken by DNR. We assume that the deer herd size in 2006 was composed of 160,000 bucks and 240,000 does, numbers that reflect DNR's own estimates (IDNR, 2009).

Figures (3.8a) through (3.8d) show the results of the simulation. The optimal policies are depicted as prices through the simple linear transformation in equation (3.8). Red dashed lines represent the 95% confidence intervals, showing the uncertainty in the system due to random shocks to the deer population. The system reaches a steady state after approximately 15 years. Doe license prices start at just over \$40 and eventually rise to almost \$56. Buck license prices start at about \$73, experience a slight drop, then rise to almost \$81.

Both the doe and buck herds experience a decrease over time. The doe herd starts from 240,000 and gradually decreases to about 97,000. The buck herd size starts at 160,000 and decreases to about 92,000. The herd size gradually decreases in order to maximize welfare over time. Rather than instantly arriving at the optimal amount, the social planner allows hunters to harvest a larger amount of deer in earlier periods in order to extract the benefits. The herd then grows after the hunting season, and the hunters harvest a slightly lower amount, eventually reaching the steady state over time. The curvature is due to the quadratic growth function. This can be seen by comparing the results to Cooper (1993), whose model allows the deer herd to instantly adjust to carrying capacity every period. As a result there is no adjustment period while the herd reaches the optimal size.

The total deer herd size in the steady state is about 189,000. Iowa DNR's stated goal is to return to the approximate deer herd size of the mid to late 90's, a goal of about 175,000 according to a 2009 report (IDNR, 2009). Thus the steady state of our model is 8% higher than DNR's recommendation. These number, however, should be viewed with caution as more work needs to be done to calibrate the demand functions accurately.

It is somewhat surprising that the model produced an optimal deer herd size only slightly higher than the DNR's goal. Again, their goal was decided upon through public meetings of all stakeholders, which came to the agreement that deer herd sizes in the mid to late 90's were optimal. The model comes to a roughly similar number by weighing the current and future benefits of hunters, agricultural producers, and drivers. The big difference is how the goal is met. Currently, DNR determines both the quantity and price of licenses. The price of licenses was historically set to a level roughly similar to neighboring states, with periodic rises due to inflation. In 2017, the price of both buck and doe hunting licenses was $28.50. Our model, on the other hand, ties prices to quantities of licenses so that the value of a deer is reflected in the price. The steady state of doe prices is roughly twice the current price, while the buck license prices are almost three times the current price.

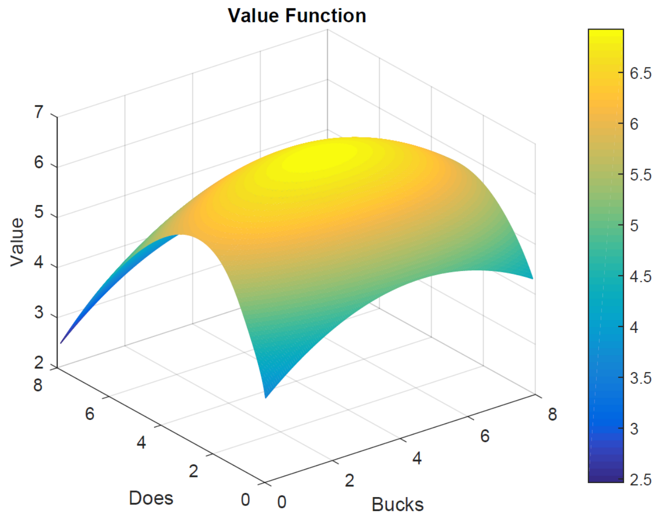
Conclusion

This paper studied the problem of how a social planner could effectively manage a Midwest deer herd. We formulated a dynamic model which required the social planner to consider the tradeoffs between current and future deer benefits. An important consideration for deer herd management in the Midwest, as opposed to previous studied regions such as California, are the costs that deer impose on society through crop depredation and deer vehicle collisions. Given the highly favorable conditions in farming states, a deer herd left unchecked could impose huge economic costs on society. The social planner must balance these costs with hunting benefits in order to optimize social welfare.

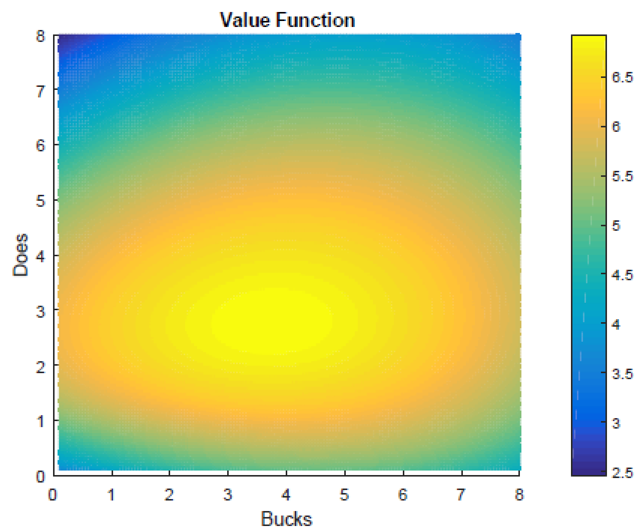
Our results indicate steady state doe and buck license prices of $56 and $81, respectively, along with a total herd size of 189,000. Had our optimal policies been implemented in 2006, these numbers would have been reached by approximately 2021. Our total deer herd size in the steady state was only slightly higher than the current goal of the Iowa DNR.

The uncertainty behind some of our parameters implies more research is needed to gain a more accurate picture of the optimal management of deer. In particular, our study would benefit from new estimates of the demand for deer which differentiates between does and bucks. In addition, we found little consensus for what the total deer capacity of Iowa is, a key parameter to the deer growth function. A precisely measured parameter would help shed more light on the dynamic path of does and bucks.

.



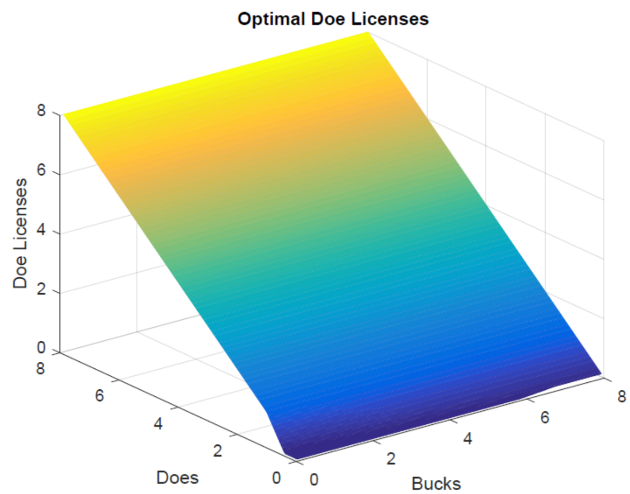
(a)



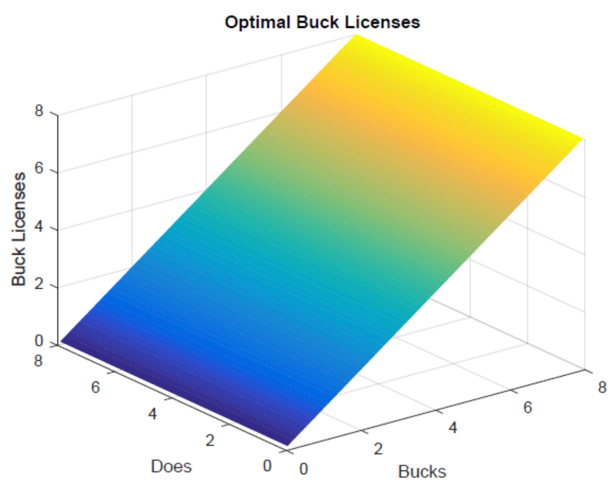
(b)

Figure 3.5: Numerical Approximation of the Value Function

Note: (a) Numerical approximation to the value function. Units for x and y axes ("does" and "bucks") is hundreds of thousands. Units for Z-axis values ("value") is hundreds of millions. (b) Contour map of the value function.



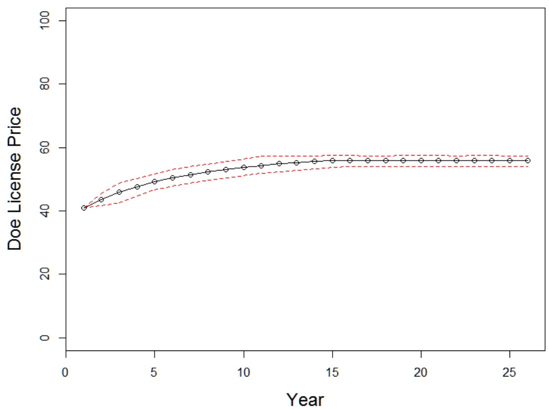
(a)



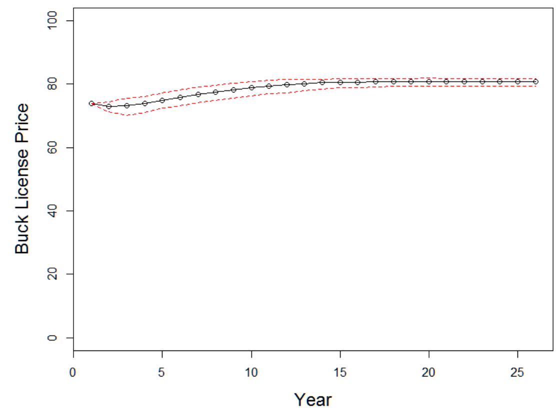
(b)

Figure 3.7: Numerical Approximation of the Policy Functions

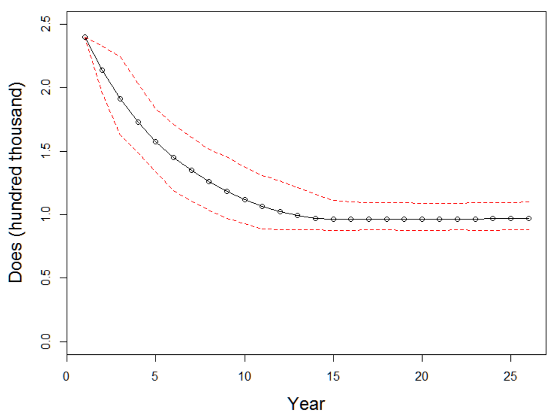
Notes: (a) Numerical approximation to the policy functions for doe hunting licenses. Units for x,y,and z axes are hundreds of thousands. (b) Numerical approximation to the policy functions for buck hunting licenses. Units for x,y,and z axes are hundreds of thousands.



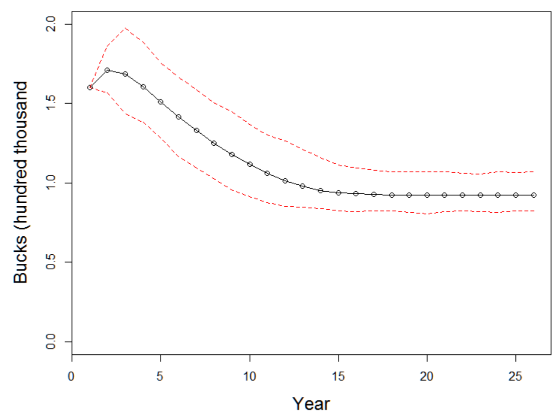
(a) Doe license prices.



(b) Buck license prices.



(c) Doe herd size.



(d) Buck herd size.

Figure 3.9: Simulation results

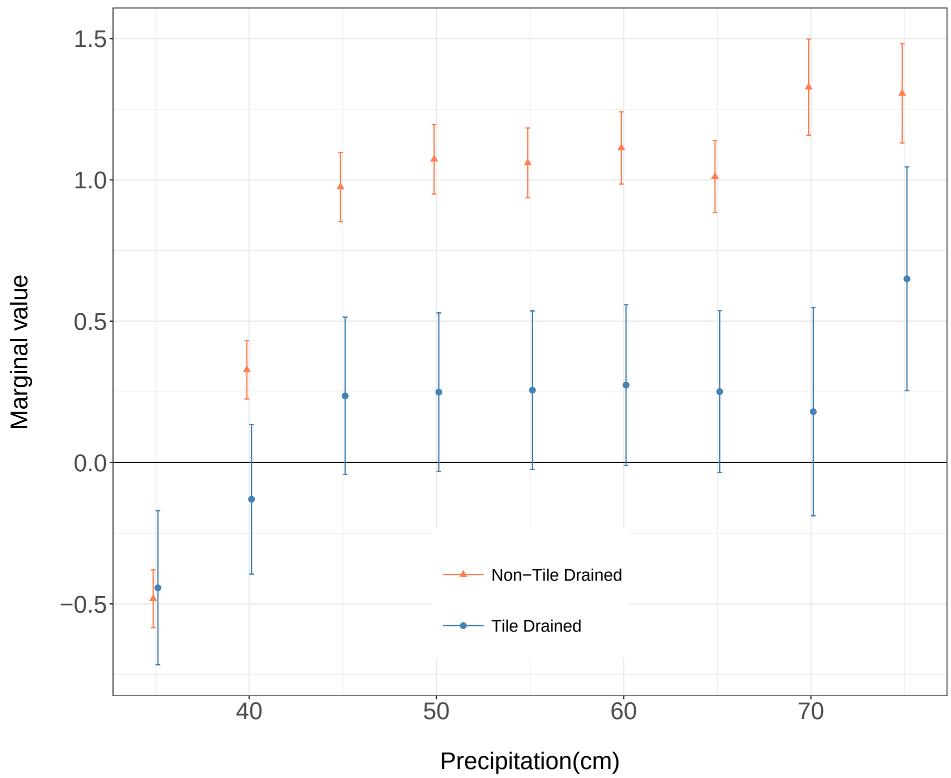
Notes: Simulation results over 25 years using a starting doe and buck population of 240,000 and 160,000, respectively, corresponding to Iowa deer herd size in 2006. For herd sizes, units along the y-axis is hundreds of thousands. Red dashed lines indicate 95% confidence intervals.

APPENDIX A. CHAPTER ONE APPENDIX

Functional Form

The functions $f(prec)$ and $g(GDD)$ of the climate variables are unknown, but past research has strongly indicated that it will be nonlinear. Our baseline model contains a quadratic form for annual precipitation and temperature, which is a common specification. We test the sensitivity of the results to several more flexible forms, including a step-function and a linear spline with three knots. All models show important differences in the value of climate on tile drained versus non-tile drained farms. For example, figure (A.1) plots the coefficients of precipitation for the step-function model, which divides the range of precipitation into five centimeter bins, with an omitted reference category for farms with less than 35 centimeters of precipitation. The plot shows that, except for the lowest bin category, the value of precipitation is higher on non-tile drained farms. The 95% confidence intervals show no overlap, providing evidence that the values are statistically different.

There has also been debate over whether it is appropriate to aggregate climate variables over the growing season, since the effect of precipitation and temperature on crop yields depends on the stage of crop development, and also whether using growing degree days is necessary, as opposed to Celsius temperatures (Massetti et al., 2015). To address this, we estimate a seasonal model which divides the growing season into two time periods: April through June ("spring") and July through September ("summer"). For each period, we estimate the response of farmland values to average Celsius temperature, total precipitation, and their squares. We also exclude harmful degree days from this model, as its effects may already be captured by the quadratic temperature curve (Massetti et al., 2015). Table (A.3) shows the results of this model. The differences in precipitation appear to be in the Spring, where the value is three



Notes: Coefficients and 95% confidence intervals for coefficients for precipication from a regression of the log of farmland values on climate and control variables, using a step-function for precipitation and growing degree days. The omitted reference category is for observations with less than 35 cm of precipitation.

Figure A.1: Plot of Coefficients for Precipitation Using a Step-Function

times higher on non-tile drained land, while the value of precipitation in the Summer is nearly equal. Note, however, that the damage estimates from climate change, shown in table (1.4), are virtually the same whether using annual or seasonal climate variables.

Tile Drainage: Historical Estimates

From 1920 to 1978, estimates of tile drainage were produced through the decennial Census of Drainage and sporadically through the Census of Agriculture. Unfortunately, these estimates used inconsistent methodologies. In addition, tile drainage is inherently difficult to estimate since it is buried underground, sometimes with no records of installation. Combined, these problems helped create a wide variation in historical drainage estimates. The 1969, 1974, and 1978 estimates from the Census of Agriculture, for example, indicated that the total amount of acres with tile drainage were approximately 60, 43, and 108 million, respectively. Since tile drainage is rarely removed, these estimates are likely significantly biased and thus unsuitable for comparisons over time.

A separate set of county level estimates were produced in 1982, 1987, and 1992 by the National Resources Inventory (NRI). These datasets were based on statistical sampling techniques and produced more consistent estimates over time. However, they also had some notable drawbacks. For example, it could be difficult for a staff member to identify whether the sampling location contained tile drainage. Even if they did, the survey was limited to only listing three land practices, which means it is possible that tile drainage was omitted in some cases (Sugg, 2007, p.2). Another data set, produced by the World Resources Institute (WRI) in 2007, estimated tile drainage in ten mid-western states based on data on row crops and soil quality. That study assumes that if crops are being grown on poorly drained soil, than they are likely to contain tile drainage (Sugg, 2007, p.3). The resulting estimates were combined with 1992 NRI estimates from the remaining U.S. states.

STATSGO Variables

Table A.1: STATSGO variables

Variable	STATSGO label	Definition
Slope	slope_r	The difference in elevation between two points, expressed as a percentage of the distance between those points.
K-factor	kffact	An erodibility factor which quantifies the susceptibility of soil particles to detachment by water.
Percent Clay	claytotal_r	Mineral particles less than 0.002mm in equivalent diameter as a weight percentage of the less than 2.0mm fraction.
Soil Class	nirrcapcl	The percent composition of the map unit that has the capability class displayed in the Non-Irrigated Capability Class.
Permeability	ksat_r	The amount of water that would move vertically through a unit area of saturated soil in unit time under unit hydraulic gradient.
Water Capacity	awc_r	The amount of water that an increment of soil depth, inclusive of fragments, can store that is available to plants.
Drainage Class	drainagecl	Identifies the natural drainage conditions of the soil and refers to the frequency and duration of wet periods.

Notes: Definitions can be found in documentation available at www.nrcs.usda.gov/

Regression Results

Table A.2: Regressions Results: Quadratic Model

	Probit	OLS		
	(1)	(2)	(3)	(4)
		Pooled	Tile Drainage	No Tile Drainage
Prec.	0.00842***	0.0141*	0.0793***	0.172***
	(0.00205)	(0.00814)	(0.0193)	(0.0106)
Prec.^{2}	-0.0000836***	-0.0000753	-0.000698***	-0.00146***
	(0.0000188)	(0.0000726)	(0.000172)	(0.0000942)
GDD	0.317***	2.865***	6.700***	7.453***
	(0.0357)	(0.155)	(0.389)	(0.215)
GDD^2	-0.0908***	-0.592***	-1.760***	-1.858***
	(0.00835)	(0.0367)	(0.100)	(0.0561)
HDD	-0.0157***	-0.144***	-0.259***	-0.416***
	(0.00253)	(0.0124)	(0.0315)	(0.0165)
Prec.*Drainage	0.000264*			
	(0.000157)			
Drainage Class	-0.00680			
	(0.00887)			
Clay	0.000621***	-0.00292***	0.00159	0.00931***
	(0.000195)	(0.000876)	(0.00148)	(0.00106)
Permeability	-0.00114***	0.000570	-0.00652***	-0.0140***
	(0.0000983)	(0.000471)	(0.00124)	(0.000755)
K-Factor	0.0967***	0.0205	0.550***	1.421***
	(0.0154)	(0.0687)	(0.120)	(0.0889)
Soil Class	0.00155***	0.00120***	0.00989***	0.0237***
	(0.0000808)	(0.000339)	(0.00127)	(0.000915)
Slope	-0.00467***	-0.0104***	-0.0436***	-0.0836***
	(0.000277)	(0.00115)	(0.00460)	(0.00294)
Water Capacity	0.00120	0.0253***	0.0505***	0.0591***
	(0.000992)	(0.00415)	(0.00658)	(0.00458)
Latitude	-0.00368**	-0.0224***	-0.0419***	-0.0616***
	(0.00166)	(0.00721)	(0.0118)	(0.00766)

Table A.2: (Continued)

Regressions Results: Quadratic Model

Pop. Density	-0.00150***	0.0221***	0.0290***	-0.00656***
	(0.000181)	(0.00186)	(0.00494)	(0.00202)
Pop. Density^2	0.0000107***	-0.000317***	-0.000379***	-0.000110***
	(0.00000213)	(0.0000443)	(0.000141)	(0.0000396)
Per Capita Income	0.000165*	-0.00198***	-0.00249***	0.000629
	(0.0000944)	(0.000429)	(0.000887)	(0.000450)
Inverse Mills Ratio			1.206***	2.332***
			(0.151)	(0.0931)
N	812,304	812,304	128,716	683,588
R^2		0.09	0.06	0.10

Table A.3: Regressions Results: Seasonal Model

	Probit	OLS		
	(1)	(2) Pooled	(3) Tile Drainage	(4) No Tile Drainage
Prec.(Spring)	0.0140***	0.137***	0.113***	0.364***
	(0.00337)	(0.0140)	(0.0307)	(0.0167)
Prec.^{2}(Spring)	-0.000342***	-0.00224***	-0.00296***	-0.00761***
	(0.0000558)	(0.000225)	(0.000553)	(0.000311)
Temp.(Spring)	0.0811***	-0.295***	0.0578	0.993***
	(0.0204)	(0.0986)	(0.143)	(0.113)
Temp.^{2}(Spring)	-0.000727	0.00844***	0.00520	-0.00239
	(0.000527)	(0.00251)	(0.00406)	(0.00265)
Prec.(Summer)	0.00527***	-0.0155	0.0912***	0.0819***
	(0.00183)	(0.00951)	(0.0221)	(0.0103)
Prec.^{2}(Summer)	-0.000000511	0.000553***	-0.000633	0.000236
	(0.0000319)	(0.000168)	(0.000400)	(0.000164)
Temp.(Summer)	0.133***	1.365***	3.453***	3.137***
	(0.0326)	(0.148)	(0.274)	(0.175)
Temp.^{2}(Summer)	-0.00467***	-0.0307***	-0.0892***	-0.0971***
	(0.000698)	(0.00316)	(0.00714)	(0.00429)
Prec.*Drainage	-0.0000426			
	(0.000131)			
Drainage Class	0.00999			
	(0.00748)			
Clay	0.000210	-0.00234***	-0.00210	0.00489***
	(0.000185)	(0.000845)	(0.00130)	(0.000930)
Permeability	-0.00127***	0.000333	-0.00713***	-0.0171***
	(0.0000949)	(0.000468)	(0.00131)	(0.000824)
K-Factor	0.107***	-0.0169	0.657***	1.502***
	(0.0144)	(0.0661)	(0.124)	(0.0882)
Soil Class	0.00143***	0.00168***	0.00958***	0.0227***
	(0.0000748)	(0.000348)	(0.00115)	(0.000878)

Table A.3: (Continued)

Regressions Results: Seasonal Model

Slope	-0.00464***	-0.00825***	-0.0405***	-0.0813***
	(0.000256)	(0.00117)	(0.00442)	(0.00297)
Water Capacity	0.00174*	0.0289***	0.0668***	0.0661***
	(0.000927)	(0.00416)	(0.00663)	(0.00466)
Latitude	0.00128	-0.0192**	-0.0553***	0.0137*
	(0.00184)	(0.00797)	(0.0136)	(0.00822)
Pop. Density	-0.00154***	0.0230***	0.0263***	-0.00577***
	(0.000176)	(0.00187)	(0.00464)	(0.00203)
Pop. Density^2	0.0000124***	-0.000326***	-0.000340***	-0.0000936**
	(0.00000210)	(0.0000444)	(0.000130)	(0.0000399)
Per Capita Income	0.000157*	-0.00215***	-0.00254***	0.000259
	(0.0000865)	(0.000422)	(0.000872)	(0.000441)
Inverse Mills Ratio			1.168***	2.251***
			(0.141)	(0.0921)
N	812,304	812,304	128,716	683,588
R^2		0.09	0.06	0.11

Table A.4: Regressions Results: Linear Spline Model

	Probit	OLS		
	(1)	(2)	(3)	(4)
		Pooled	Tile Drainage	No Tile Drainage
Prec. Spline 1	0.00850***	0.0267***	0.0685***	0.138***
	(0.00103)	(0.00382)	(0.0116)	(0.00623)
Prec. Spline 2	-0.00961***	-0.0186***	-0.0658***	-0.132***
	(0.00107)	(0.00459)	(0.0124)	(0.00696)
Prec. Spline 3	0.000992	-0.00959***	-0.00258	-0.00104
	(0.000665)	(0.00296)	(0.00501)	(0.00320)
Prec. Spline 4	-0.000158	0.0256***	0.0168*	0.0237***
	(0.000986)	(0.00520)	(0.0102)	(0.00523)
GDD Spline 1	0.349***	2.257***	3.332***	5.997***
	(0.0188)	(0.0715)	(0.268)	(0.163)
GDD Spline 2	-0.559***	-2.085***	-4.469***	-7.929***
	(0.0188)	(0.0792)	(0.380)	(0.244)
GDD Spline 3	0.0763***	-0.891***	-1.157***	-0.0456
	(0.0202)	(0.0926)	(0.177)	(0.0999)
GDD Spline 4	0.324***	1.017***	2.982***	4.869***
	(0.0284)	(0.132)	(0.420)	(0.211)
HDD	0.00564**	-0.0972***	-0.0463	-0.0520***
	(0.00261)	(0.0129)	(0.0283)	(0.0129)
Prec.*Drainage	0.000159			
	(0.000140)			
Drainage Class	0.00166			
	(0.00800)			
Clay	0.000374**	-0.00523***	-0.00135	0.00248**
	(0.000179)	(0.000866)	(0.00136)	(0.00100)
Permeability	-0.00137***	-0.000269	-0.00539***	-0.0141***
	(0.0000950)	(0.000447)	(0.00122)	(0.000720)
K-Factor	0.107***	0.155**	0.468***	1.340***
	(0.0142)	(0.0671)	(0.116)	(0.0832)
Soil Class	0.00126***	0.000237	0.00505***	0.0146***
	(0.0000748)	(0.000337)	(0.000940)	(0.000669)

Table A.4: (Continued)

Regressions Results: Linear Spline Model

Slope	-0.00358***	-0.0119***	-0.0261***	-0.0564***
	(0.000264)	(0.00115)	(0.00350)	(0.00208)
Water Capacity	0.000857	0.0164***	0.0418***	0.0422***
	(0.000910)	(0.00401)	(0.00643)	(0.00442)
Latitude	0.00564***	0.0189***	0.0272**	0.0925***
	(0.00155)	(0.00715)	(0.0120)	(0.00769)
Pop. Density	-0.00199***	0.0184***	0.0287***	-0.00970***
	(0.000181)	(0.00179)	(0.00479)	(0.00197)
Pop. Density^2	0.0000136***	-0.000291***	-0.000372***	-0.0000929**
	(0.00000232)	(0.0000423)	(0.000135)	(0.0000378)
Per Capita Income	0.000131	-0.00190***	-0.00295***	-0.000368
	(0.0000905)	(0.000428)	(0.000879)	(0.000442)
Inverse Mills Ratio			0.780***	1.720***
			(0.125)	(0.0742)
N	812,304	812,304	128,716	683,588
R^2		0.09	0.06	0.11

Table A.5: Regressions Results: Step-Function Model

	Probit	OLS		
	(1)	(2)	(3)	(4)
		Pooled	Tile Drainage	No Tile Drainage
Prec.(35-40cm)	-0.0366***	0.0838*	-0.443***	-0.482***
	(0.0113)	(0.0473)	(0.139)	(0.0522)
Prec.(40-45cm)	0.00423	0.211***	-0.130	0.328***
	(0.0112)	(0.0517)	(0.135)	(0.0526)
Prec.(45-50cm)	0.0420***	0.275***	0.236*	0.975***
	(0.0119)	(0.0528)	(0.142)	(0.0623)
Prec.(50-55cm)	0.0412***	0.353***	0.249*	1.073***
	(0.0129)	(0.0533)	(0.143)	(0.0627)
Prec.(55-60cm)	0.0409***	0.347***	0.256*	1.060***
	(0.0137)	(0.0535)	(0.143)	(0.0630)
Prec.(60-65cm)	0.0450***	0.333***	0.274*	1.113***
	(0.0147)	(0.0543)	(0.145)	(0.0652)
Prec.(65-70cm)	0.0353**	0.379***	0.251*	1.012***
	(0.0170)	(0.0577)	(0.146)	(0.0649)
Prec.(70-75cm)	0.0489**	0.483***	0.180	1.328***
	(0.0198)	(0.0784)	(0.188)	(0.0868)
Prec.(75cm+)	0.0407*	0.613***	0.650***	1.306***
	(0.0222)	(0.0881)	(0.202)	(0.0896)
GDD (1.5-1.75)	0.0734***	0.503***	0.873***	1.526***
	(0.00565)	(0.0193)	(0.0754)	(0.0453)
GDD (1.75-2)	0.0829***	0.652***	0.981***	1.799***
	(0.00710)	(0.0251)	(0.0836)	(0.0512)
GDD (2-2.25)	0.0415***	0.667***	0.673***	1.299***
	(0.00845)	(0.0323)	(0.0655)	(0.0420)
GDD (2.5-2.75)	-0.0181*	0.662***	0.203***	0.419***
	(0.0105)	(0.0392)	(0.0669)	(0.0423)
GDD (2.75-3)	-0.0301**	0.523***	-0.318***	0.164***
	(0.0119)	(0.0470)	(0.0901)	(0.0511)
GDD (3-3.25)	-0.0168	0.354***	-0.513***	0.272***
	(0.0144)	(0.0582)	(0.117)	(0.0608)

Table A.5: (Continued)

Regressions Results: Step-Function Model

GDD (3.25-3.5)	-0.00258	0.534***	-0.333**	0.703***
	(0.0167)	(0.0727)	(0.141)	(0.0749)
GDD (3.5+)	0.0101	0.332***	-0.659***	0.761***
	(0.0195)	(0.0875)	(0.179)	(0.0935)
HDD	-0.00148	-0.104***	-0.0994***	-0.147***
	(0.00233)	(0.0113)	(0.0238)	(0.0110)
Prec.*Drainage	-0.0000188			
	(0.000105)			
Drainage Class	0.00854			
	(0.00609)			
Clay	0.000351*	-0.00357***	-0.0001000	0.00574***
	(0.000180)	(0.000853)	(0.00135)	(0.000972)
Permeability	-0.00135***	-0.000415	-0.00690***	-0.0189***
	(0.0000946)	(0.000462)	(0.00142)	(0.000862)
K-Factor	0.0991***	0.103	0.473***	1.491***
	(0.0139)	(0.0658)	(0.123)	(0.0852)
Soil Class	0.00139***	0.000523	0.00741***	0.0212***
	(0.0000717)	(0.000336)	(0.00122)	(0.000865)
Slope	-0.00396***	-0.0106***	-0.0313***	-0.0738***
	(0.000249)	(0.00109)	(0.00432)	(0.00261)
Water Capacity	0.00181*	0.0202***	0.0564***	0.0620***
	(0.000935)	(0.00415)	(0.00673)	(0.00474)
Latitude	0.00126	-0.00179	0.000499	0.0142**
	(0.00132)	(0.00659)	(0.00997)	(0.00689)
Pop. Density	-0.00181***	0.0199***	0.0292***	-0.0126***
	(0.000175)	(0.00181)	(0.00501)	(0.00205)
Pop. Density^2	0.0000119***	-0.000299***	-0.000386***	-0.0000753**
	(0.00000218)	(0.0000429)	(0.000141)	(0.0000384)
Per Capita Income	0.000138	-0.00181***	-0.00261***	0.000230
	(0.0000911)	(0.000419)	(0.000885)	(0.000433)
Inverse Mills Ratio			1.014***	2.296***
			(0.158)	(0.0941)
N	812,304	812,304	128,716	683,588
R^2		0.09	0.06	0.11

APPENDIX B. CHAPTER TWO APPENDIX

Econometrics Criteria

Kubis and Schneider (2012) advise the following criteria to evaluate the results of the Arellano-Bond model:

1. The Hansen J test does not reject the H_o of the valid instruments.

2. The number of instruments is smaller than then number of cross-sectional units.

3. The Arellano and Bond AR(2) test statistic for the first-differenced residuals is insignificant.

4. The magnitude of the coefficient on the lagged dependent variable lies in between the pooled-OLS and fixed-effects estimates.

Item (i) is the usual test statistic used to test the validity of instruments when the number of moment conditions is greater than the number of parameters. Roodman (2009) notes that this test statistic "is usually and reasonably thought of as a test of instrument validity. But it can also be viewed as a test of structural specification". He also notes that careful attention should be paid to the magnitude of the p-value and not just whether the test fails or not. For example, a p-value as high as 25% should still be viewed cautiously, as this means that the odds are 1 in 4 that one would observe a statistic that large. Conversely, an extremely high p-value can also be a cause for concern. For example, a test statistic at or near 1 is a classic sign that too many instruments have weakened the test.

Item (ii) also addresses the problems that can occur from overfitting. The total number of instruments available is quadratic in T- thus the number of instruments can grow quickly relative to the sample size. This can cause overfitting of the endogenous variables that are being instrumented for, as well as imprecise estimates of the optimal weighting matrix used in two-step GMM. For this reason, the robustness of the estimates to different numbers of instruments will be investigated (although in all cases they will be less than N). Two simple ways of reducing the number of instruments are to reduce the number of lags of the dependent variable used as instruments, and to collapse the instruments matrix. The latter can be explained by considering the instruments matrix $\boldsymbol{Z}$. Assume for clarity that the data set is balanced and transformed through first differences (the intuition using a FOD transformation and unbalanced data is the same). $\boldsymbol{Z}$ consists of a stack of blocks of the form:

$$\boldsymbol{Z} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ y_{i1} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & y_{i2} & y_{i1} & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & y_{i3} & y_{i2} & y_{i1} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots \end{bmatrix}$$

Each row corresponds to one period; the first row is all zeros because there are no lags of the dependent variable available to be used as instruments. This equation also clearly shows how the number of instruments differs in each period, since more lags are available as T increases. "Collapsing" Z simply means squeezing Z horizontally and adding together formally distinct columns (Roodman, 2009):

$$\tilde{\boldsymbol{Z}} = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ y_{i1} & 0 & 0 & \cdots \\ y_{i2} & y_{i1} & 0 & \cdots \\ y_{i3} & y_{i2} & y_{i1} & \cdots \\ \cdots & \cdots & \cdots & \ddots \end{bmatrix}$$

Collapsing has the effect of making the instrument count linear in T, while retaining information since no lags are actually dropped. The robustness check in this paper will include a combination of limited lags and a collapsed instrument matrix to investigate a range of possible instrument sets, while still allowing for a Hansen test in eac h model to check instrument validity and structural specifications.

Item (*iii*) tests for an AR(2) process in the first difference of residuals, although the actual model is estimated using the FOD transformation. First differences are used because the error terms in a FOD transformation are all interrelated, which makes them unhelpful in testing for auto-correlation. An AR(2) test is used because AR(1) is expected of the first differences [1], since $\Delta\epsilon_{i,t}$ and $\Delta\epsilon_{i,t-1}$ both share the common term $\epsilon_{i,t-1}$. If an AR(2) process is detected in the first difference of residuals, it is a sign that lags of the dependent variable are endogenous, making them poor instruments. Thus, with a null hypothesis of an AR(2) process in the data, a model that passes this test will have a p statistic higher than the conventional testing threshold of 5%.

Finally, Item (*iv*) takes advantage of the opposing theoretical biases from pooled OLS and fixed-effects estimates in a dynamic setting. Specifically, while the pooled OLS estimate of the coefficient on the lagged dependent variable is biased *upwards*, fixed effects estimates on the same coefficient are biased *downwards*. Therefore, one test of the validity of the GMM results is for the estimate on this parameter to be in between the pooled OLS and fixed-effects estimates (Blundell et al., 2001).

[1] All specifications in the results section found AR(1), with p-values near zero in every case

BIBLIOGRAPHY

Aiken, R. (2006). Net Economic Values of Wildflife-Related Recreation in 2006. Technical report, U.S. Fish and Wildlife Service.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Atasoy, M., Palmquist, R. B., and Phaneuf, D. J. (2006). Estimating the effects of urban residential development on water quality using microdata. *Journal of environmental management*, 79(4):399–408.

Auffhammer, M., Hsiang, S. M., Schlenker, W., and Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, page ret016.

Auffhammer, M. and Schlenker, W. (2014). Empirical studies on agricultural impacts and adaptation. *Energy Economics*, 46:555–561.

Balkcom, K. S., Blackmer, A. M., Hansen, D. J., Morris, T. F., and Mallarino, A. P. (2003). Testing soils and cornstalks to evaluate nitrogen management on the watershed scale. *Journal of environmental quality*, 32(3):1015–1024.

Barr, K. J., Babcock, B. A., Carriquiry, M. A., Nassar, A. M., and Harfuch, L. (2011). Agricultural land elasticities in the united states and brazil. *Applied Economic Perspectives and Policy*, 33(3):449–462.

Blundell, R., Bond, S., and Windmeijer, F. (2001). *Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator*, volume 15. Emerald Group Publishing Limited.

Burke, M., Craxton, M., Kolstad, C., Onda, C., Allcott, H., Baker, E., Barrage, L., Carson, R., Gillingham, K., Graff-Zivin, J., et al. (2016). Opportunities for advances in climate change economics. *Science*, 352(6283):292–293.

Burke, M. and Emerick, K. (2016). Adaptation to climate change: Evidence from us agriculture. *American Economic Journal- Economic Policy*, 8(3):106–140.

CDFW (2016). California Deer Conservation and Management Plan. Technical report, California Department of Fish and Wildlife.

Chatzopoulos, T. and Lippert, C. (2015). Endogenous farm-type selection, endogenous irrigation, and spatial effects in ricardian models of climate change. *European Review of Agricultural Economics*, page jbv014.

Cooper, J. C. (1993). A bioeconomic model for estimating the optimal level of deer and tag sales. *Environmental and Resource Economics*, 3(6):563–579.

Dauer, D. M., Ranasinghe, J. A., and Weisberg, S. B. (2000). Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in chesapeake bay. *Estuaries*, 23(1):80–96.

Delpla, I., Jung, A.-V., Baures, E., Clement, M., and Thomas, O. (2009). Impacts of climate change on surface water quality in relation to drinking water production. *Environment International*, 35(8):1225–1233.

Deschenes, O. and Greenstone, M. (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *The American Economic Review*, pages 354–385.

Downing, J. A., Antoniou, G., Boatwright, L., Conrad, S., Kendall, D., Li, J., and Kaleita, A. (2006). Iowa lakes survey: Summer 2005 scientific initiatives report. Technical report, Iowa Department of Natural Resources.

EPA, U. (2002). National water quality inventory 2000 report.

Evans, S. G. and Potts, M. D. (2015). Effect of agricultural commodity prices on species abundance of us grassland birds. *Environmental and Resource Economics*, 62(3):549–565.

Feick, S., Siebert, S., and Döll, P. (2005). A digital global map of artificially drained agricultural areas.

Fezzi, C. and Bateman, I. (2007). The impact of climate change on agriculture: nonlinear effects and aggregation bias in ricardian models of farmland value. *Journal of the Association of Environmental and Resource Economice*, 2(1):57–92.

Fisher, A. C., Hanemann, W. M., Roberts, M. J., and Schlenker, W. (2012). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather: comment. *The American Economic Review*, 102(7):3749–3760.

Greene, W. H. (2000). Econometric analysis 5th edition. *International edition, New Jersey: Prentice Hall.*

Hausman, C., Auffhammer, M., and Berck, P. (2012). Farm acreage shocks and crop prices: an svar approach to understanding the impacts of biofuels. *Environmental and Resource Economics*, 53(1):117–136.

Heckman, J. J. (1977). Sample selection bias as a specification error (with an application to the estimation of labor supply functions).

Hendricks, N. (2016). The economic benefits from innovations to reduce heat and water stress in agriculture. Working paper.

Huang, H., Khanna, M., et al. (2010). An econometric analysis of us crop yield and cropland acreage: implications for the impact of climate change.

Hussain, A. M. T. and Tschirhart, J. (2010). Optimal harvest licensing when harvest success is uncertain. *American Journal of Agricultural Economics*, page aap019.

IDNR (2009). A Review of Iowa's Deer Management Program. Technical report, Iowa Department of Natural Resources.

IDNR (2016). Trends in Iowa Wildlife Populations and Harvest. Technical report, Iowa Department of Natural Resources.

Jaynes, D. and James, D. (2007). The extent of farm drainage in the united states. *US Department of Agriculture*.

Keith, J. E. and Lyon, K. S. (1985). Valuing wildlife management: a utah deer herd. *Western Journal of Agricultural Economics*, pages 216–222.

Klippart, J. H. (1861). *The principles and practice of land drainage*. R. Clarke & Company.

Kubis, A. and Schneider, L. (2012). Human capital mobility and convergence: a spatial dynamic panel model of the german regions. Technical report, IAB Discussion Paper.

Kurukulasuriya, P., Kala, N., and Mendelsohn, R. (2011). Adaptation and climate change impacts: a structural ricardian model of irrigation and farm income in africa. *Climate Change Economics*, 2(02):149–174.

Kurukulasuriya, P., Mendelsohn, R., et al. (2008). A ricardian analysis of the impact of climate change on african cropland. *African Journal of Agricultural and Resource Economics*, 2(1):1–23.

Loison, A. and Langvatn, R. (1998). Short-and long-term effects of winter and spring weather on growth and survival of red deer in norway. *Oecologia*, 116(4):489–500.

Lougheed, V. L., Crosbie, B., and Chow-Fraser, P. (2001). Primary determinants of macrophyte community structure in 62 marshes across the great lakes basin: latitude, land use, and water quality effects. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(8):1603–1612.

Mankin, K. R., Wang, S.-H., Koelliker, J. K., Huggins, D. G., and de Noyelles, F. (2003). Watershed-lake water quality modeling: Verification and application. *Journal of Soil and Water Conservation*, 58(4):188–197.

Massetti, E. and Mendelsohn, R. (2011). Estimating ricardian models with panel data. *Climate Change Economics*, 2(04):301–319.

Massetti, E., Mendelsohn, R., and Chonabayashi, S. (2015). How well do degree days over the growing season capture the effect of climate on farmland values? *Working Paper*.

Meador, M. R. and Goldstein, R. M. (2003). Assessing water quality at large geographic scales: relations among land use, water physicochemistry, riparian condition, and fish community structure. *Environmental Management*, 31(4):0504–0517.

Mendelsohn, R., Nordhaus, W. D., and Shaw, D. (1994). The impact of global warming on agriculture: a ricardian analysis. *The American economic review*, pages 753–771.

Miranda, M. J. and Fackler, P. L. (2004). *Applied computational economics and finance*. MIT press.

Montgomery, R. H. and Reckhow, K. H. (1984). Techniques for detecting trends in lake water quality1.

Nerlove, M. (1956). Estimates of the elasticities of supply of selected agricultural commodities. *Journal of Farm Economics*, 38(2):496–509.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Oregon State University (2016). PRISM climate group. Online; accessed March 16th, 2016.

Paddock, C. D. and Yabsley, M. J. (2007). Ecological havoc, the rise of white-tailed deer, and the emergence of amblyomma americanum-associated zoonoses in the united states. In *Wildlife and emerging zoonotic diseases: the biology, circumstances and consequences of cross-species transmission*, pages 289–324. Springer.

Pavelis, G. A. (1987). *Farm drainage in the United States: History, status, and prospects.* Number 1455. US Department of Agriculture, Economic Research Service.

Rabotyagov, S., Kling, C., Gassman, P., Rabalais, N., and Turner, R. (2014). The economics of dead zones: Causes, impacts, policy challenges, and a model of the gulf of mexico hypoxic zone. *Review of Environmental Economics and Policy*, 8(1):58–79.

Reinsborough, M. J. (2003). A ricardian model of climate change in canada. *Canadian Journal of Economics/Revue canadienne d'économique*, 36(1):21–40.

Rondeau, D. and Bulte, E. (2007). Wildlife damage and agriculture: a dynamic analysis of compensation schemes. *American Journal of Agricultural Economics*, 89(2):490–507.

Rondeau, D. and Conrad, J. M. (2003). Managing urban deer. *American Journal of Agricultural Economics*, 85(1):266–281.

Roodman, D. (2009). A note on the theme of too many instruments*. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.

Schilling, K. E., Gassman, P. W., Kling, C. L., Campbell, T., Jha, M. K., Wolter, C. F., and Arnold, J. G. (2014). The potential for agricultural land use change to reduce flood risk in a large watershed. *Hydrological Processes*, 28(8):3314–3325.

Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2005). Will us agriculture really benefit from global warming? accounting for irrigation in the hedonic approach. *American Economic Review*, pages 395–406.

Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2006). The impact of global warming on us agriculture: an econometric analysis of optimal growing conditions. *Review of Economics and Statistics*, 88(1):113–125.

Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2007). Water availability, degree days, and the potential impact of climate change on irrigated agriculture in california. *Climatic Change*, 81(1):19–38.

Schlenker, W. and Roberts, M. J. (2006). Nonlinear effects of weather on corn yields. *Applied Economic Perspectives and Policy*, 28(3):391–398.

Schwabe, K. A., Schuhmann, P. W., and Tonkovich, M. (2002). A dynamic exercise in reducing deer-vehicle collisions: Management through vehicle mitigation techniques and hunting. *Journal of Agricultural and Resource Economics*, pages 261–280.

Secchi, S., Gassman, P. W., Williams, J. R., and Babcock, B. A. (2009). Corn-based ethanol production and environmental quality: A case of iowa and the conservation reserve program. *Environmental Management*, 44(4):732–744.

Seo, S. N. and Mendelsohn, R. (2008). Measuring impacts and adaptations to climate change: a structural ricardian model of african livestock management1. *Agricultural economics*, 38(2):151–165.

Sheffield, J., Barrett, A. P., Colle, B., Nelun Fernando, D., Fu, R., Geil, K. L., Hu, Q., Kinter, J., Kumar, S., Langenbrunner, B., et al. (2013). North american climate in cmip5 experiments. part i: evaluation of historical simulations of continental and regional climatology*. *Journal of Climate*, 26(23):9209–9245.

Shi, Y. J., Phipps, T. T., and Colyer, D. (1997). Agricultural land values under urbanizing influences. *Land economics*, pages 90–100.

Simpson, T. W., Sharpley, A. N., Howarth, R. W., Paerl, H. W., and Mankin, K. R. (2008). The new gold rush: fueling ethanol production while protecting water quality. *Journal of Environmental Quality*, 37(2):318–324.

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

Sugg, Z. (2007). Assessing u.s. farm drainage: Can gis lead to better estimates of subsurface drainage extent? Technical report, World Resources Institute.

Tong, S. T. and Chen, W. (2002). Modeling the relationship between land use and surface water quality. *Journal of environmental management*, 66(4):377–393.

Tu, J. (2011). Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. *Applied Geography*, 31(1):376–392.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of econometrics*, 126(1):25–51.

Zhang, W. and Nickerson, C. J. (2015). Housing market bust and farmland values: Identifying the changing influence of proximity to urban centers. *Land Economics*, 91(4):605–626.