ESSAYS ON THE BENEFITS OF SOIL AND WATER CONSERVATION

A Dissertation

by

MICHAEL ALEXANDER BLACK

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Richard T. Woodward
Committee Members,	Anastasia Shcherbakova
	Ximing Wu
	Cristine Morgan
	William McIntosh
Head of Department,	Mark L. Waller

August 2021

Major Subject: Agricultural Economics

ABSTRACT

In this work, I estimate the value of soil and water quality improvements using revealed and stated preferences of anglers and farmers, respectively. I also identify the causal effect of a groundwater management program in India. I find that the standard approach to valuing changes in environmental quality may often underestimate the true marginal benefits of potential improvements. I find that farmers are willing to pay for some soil quality improvements separate from any benefits realized in changes in crop revenues or production costs, which suggests that the current approach to estimating the benefits of soil conservation may not be capturing the full set of benefits for farmers. Finally, I find that a large groundwater management program in India had - at best - moderate success in conserving groundwater.

In the first essay, I explore the value of short- and long-term changes in water quality. The economic value of water quality improvements is often assessed using short-term changes in quality measures. Meanwhile, water quality regulators seek long-term, permanent changes. I use a pooled cross-section of angler surveys administered in Texas from 2001 - 2015 and a panel of water quality data to measure the effect of short-term and long-term variation in water quality on anticipated angler utility. Using a two-stage estimation process, I find that anglers are willing to pay substantially more for long-term changes in water quality than for short-term changes. Approaches that capture only short-term variation in water quality may thus result in significant undervaluation.

In the second essay, I explore the value of soil quality improvements. The on-farm value of soil quality improvements has historically been captured by changes in production costs or crop revenues. I design a discrete choice experiment that captures the willingness-to-pay from farmers to improve three manageable soil quality characteristics, providing for the first time an estimate of the value of soil quality improvements without involving the potentially confounding effects of conservation practices. I find that, on average, farmers are willing to pay more for improvements in water infiltration than organic matter which realistically

occur with the adoption of a no-till regime. When I incorporate preference heterogeneity through the use of mixed logit models, the 95% confidence intervals for all willingness-to-pay measures include zero, even within distinct sub-groups. This suggests that while some farmers are willing to pay more than others for soil quality improvement, there is a lot of heterogeneity and many farmers may not be willing to pay anything to improve soil quality.

In the third essay, I identify the effect of a participatory groundwater management program on groundwater levels in India. The Andhra Pradesh Farmer-Managed Groundwater System (APFAMGS) was a multi-year program that provided farmers with tools to manage groundwater collectively through constant monitoring of local aquifers. Using a two-way fixed effects difference-in-differences model, I find the program had some moderate success in the intra-year groundwater levels throughout a full agricultural year. The success of the program, however, did not extend to average groundwater levels nor was the success consistent across years.

DEDICATION

Dedicated to my father, who
always
protects the embers of curiosity
from the stiff winds of life

ACKNOWLEDGMENTS

First and foremost, I thank my chair Dr. Richard T. Woodward. From him I learned how to solve dynamic optimization problems, how to adjust the rear derailleur of my bike, and how to cook a good onion soup. I thank him for his patience and guidance.

For the first essay, I thank the Human Dimensions of Natural Resources Lab at Texas A&M University for providing angler survey data; Dr. Kirk Winemiller for providing information on fish preferences and behavior; general feedback from presentations at Resources for the Future, Camp Resources XXVI, and Heartland Environmental and Resource Economics Workshop. Special thanks to Dr. David Anderson, Dr. Yung-Ping Tseng, and Dr. Adam Landon for additional survey data. The essay would not be possible without their diligent data management.

For the second essay, I thank Dr. Cristine Morgan and Dr. Dianna Bagnall for their assistance with all soil-related questions. I would like to specifically thank Dr. Morgan's undergraduate class on soil, where she gave us holes to crawl into, poetry to read, and clumps of dirt to turn over in our hand. Her love of soil is infectious.

For the third essay, I credit Nishita Sinha for finding the APFAMGS program and finding data to make this work possible. Her keen eye for empirical economics and insistence on quality make her an ideal collaborator. Special thanks to the Bharati Integrated Rural Development Society (BIRDS) for providing important data on the precise location of treatment areas.

For her support through the most stressful nights and depressing regression results, I thank my partner Amy. She lifted my spirit with adventure, and consistently reminded me that economists don't have all the answers. My happy place is on a highway with her, speeding towards the promise of new food and new air.

Lastly, I thank my dog Momo. On our walks, he has heard me rehearse all of my talks, and has been quite the sounding board for new ideas. Good boy, Mo.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Dr. Richard T. Woodward (advisor), Dr. Anastasia Shcherbakova, Dr. Ximing Wu, Dr. William McIntosh of the Department of Sociology, and Dr. Cristine Morgan of the Soil Health Institute.

The survey instrument used in the second essay was developed with help from Dr. Richard T. Woodward, Dr. Cristine Morgan, Dr. Dianna Bagnall, Dr. William McIntosh, Dr. Marissa Cisneros, Dr. Erin Kiella, Dr. Srinivasulu Ale, and Sayantan Samanta. The instrument was deployed by the United States Department of Agriculture, National Agricultural Statistics Service (USDA NASS).

The third essay was co-authored with Dr. Nishita Sinha.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate research was funded from the USDA National Institute of Food and Agriculture Grant # 2018-67019-27975, and from support from the Department of Agricultural Economics at Texas A&M University.

NOMENCLATURE

AIC	Akaike information criterion
APFAMGS	Andhra Pradesh Farmer-Managed Groundwater Systems
APWELL	Andhra Pradesh Groundwater Borewell Irrigation Schemes
ASC	Alternative specific constant
BIRDS	Bharati Intergrated Rural Development Society
DCE	Discrete choice experiment
FAO	Food and Agricultural Organization of the United Nations
LCA	Latent class analysis
MWTP	Mean willingness-to-pay
NASS	National Agricultural Statistics Service
NRCS	Natural Resources Conservation Service
OLS	Ordinary least squares
RPL	Random parameters logit
RUM	Random utility model
SWQMIS	Surface Water Quality Monitoring Information System
TCEQ	Texas Commission on Environmental Quality
TPWD	Texas Parks and Wildlife Department
USD	United States Dollar
USDA	United States Department of Agriculture
USFWS	United States Fish and Wildlife Service
VOTT	Value of travel time
WTP	Willingness-to-pay

TABLE OF CONTENTS

Page

ABSTRACT ... ii

DEDICATION ... iv

ACKNOWLEDGMENTS ... v

CONTRIBUTORS AND FUNDING SOURCES ... vi

NOMENCLATURE ... vii

TABLE OF CONTENTS ... viii

LIST OF FIGURES ... x

LIST OF TABLES ... xi

1. INTRODUCTION ... 1

2. THE VALUE OF SHORT AND LONG TERM CHANGES IN WATER QUALITY 4

  2.1 Introduction ... 4
  2.2 The Model ... 8
  2.3 Data ... 10
    2.3.1 Angler data ... 10
    2.3.2 Water quality data ... 15
  2.4 Estimation Results ... 19
  2.5 Policy Simulation and Welfare ... 23
  2.6 Conclusion and Limitations ... 28

3. AN EMPIRICAL ESTIMATE OF THE VALUE OF SOIL QUALITY IMPROVEMENTS ... 30

  3.1 Introduction ... 30
  3.2 Survey Development ... 32
    3.2.1 Study Area ... 32
    3.2.2 Attribute Selection ... 33
    3.2.3 Experimental Design ... 38
  3.3 Survey Results ... 39
  3.4 Models ... 40
    3.4.1 Conditional Logit Analysis ... 42

3.4.2 Latent Class Analysis ... 47
3.4.3 Mixed Logit Analysis ... 50
3.5 Willingness-To-Pay Simulations ... 55
3.6 Policy Implications ... 59
3.7 Discussion and Limitations ... 61

4. THE EFFECT OF A PARTICIPATORY GROUNDWATER MANAGEMENT PROGRAM ON GROUNDWATER LEVELS ... 65

4.1 Introduction ... 65
4.2 Program Background ... 68
4.3 Program Evaluation ... 70
4.3.1 Empirical Model ... 75
4.3.2 Data ... 77
4.3.3 Spatial Interpolation ... 78
4.4 Results ... 83
4.5 Discussion ... 86

5. SUMMARY AND CONCLUSION ... 90

REFERENCES ... 92

APPENDIX A. APPENDIX FOR CHAPTER 1 ... 101

APPENDIX B. APPENDIX FOR CHAPTER 2 ... 105

B.1 Full Willingness-to-Pay Estimates ... 105
B.2 Survey Refinement ... 107
B.3 Survey Instrument ... 110

APPENDIX C. APPENDIX FOR CHAPTER 3 ... 126

LIST OF FIGURES

FIGURE Page

2.1 *Typical-Trip* And *Most-Often* Survey Questions 11

2.2 Water Quality Over Time For Most-Visited Sites 17

2.3 Long-term Water Quality Across Texas 18

2.4 Map Of The Brazos River Watershed 24

3.1 Cash Rental Rates In Texas (Average, 2008 - 2020) 37

3.2 Selected Summary Statistics, Farmer Survey 41

3.3 Class-Conditional Response Probabilities, Latent Class Analysis 51

3.4 Marginal Willingness-To-Pay By Model Specification: Conditional Logit 58

3.5 Marginal Willingness-To-Pay By Model Specification: Mixed Logit 60

4.1 Map Of The APFAMGS Program Area 69

4.2 Seasonal Rainfall And Groundwater Levels In Andhra Pradesh 73

4.3 Groundwater Levels In Program Area 74

4.4 Depth To Water Table, Observed Versus Krigged Data 79

4.5 Depth To Water Table, Krigging Measurement Error 81

LIST OF TABLES

TABLE Page

2.1 Summary Of Angler Survey: Demographic Information 14

2.2 First And Second Stage Estimation Results 20

2.3 Per-Trip WTP For Simulated Transparency Improvement In A Watershed 27

2.4 Bootstrapped WTP For Simulated Transparency Improvement In A Watershed 27

3.1 Discrete Choice Experiment Attributes and Levels 36

3.2 Selected Summary Statistics, Farmer Survey 40

3.3 Conditional Logit Estimation Results 46

3.4 Latent Class Analysis Manifest Variables 48

3.5 Conditional Logit Estimation Results By Latent Class 52

3.6 Mixed Logit Estimation Results 54

3.7 Mixed Logit Estimation Results: Latent Class Analysis 55

3.8 Median Soil Quality Indicators For Study Area 56

4.1 Kriging Cross-Validation Summary Statistics 82

4.2 Effect Of APFAMGS On Annual Depth To Groundwater 84

4.3 Effect Of APFAMGS On Intra-Year Depth To Groundwater 85

4.4 Intra-Year Effects, Dynamic, Raw Data 87

A.1 Alternate Specifications: First Stage Using “Most-Often” 101

A.2 Alternate Specifications: First Stage Using “Typical-Trip” 102

A.3 Alternate Specifications: Second Stage Using “Typical-Trip” 103

A.4 Alternate Specifications: Second Stage Using “Most-Often” 104

B.1 Marginal Willingness-To-Pay By Model/Group: Conditional Logit 105

B.2 Marginal Willingness-To-Pay By Model/Group: Mixed Logit ... 106
B.3 Pilot Survey Summary Statistics ... 107
B.4 Pilot Survey Conditional Logit Estimation Results ... 108
B.5 Required Sample Sizes To Identify Effects ... 110
C.1 Two Way Fixed Effects: Intra-Year Effects, Dynamic, Krigged ... 126

1. INTRODUCTION

Estimating the costs and benefits of policies is a core component of modern economics. In the United States, federal regulation requires that "regulatory action shall not be undertaken unless the potential benefits to society from the regulation outweigh the potential costs to society".[1] Although not always enshrined in law, benefit-cost analysis is performed worldwide prior to regulatory action or new policy.

It is, however, exceedingly difficult to estimate all the costs and benefits of polices. Society has lots of people in it, and those people may be impacted differently. There are also some costs and benefits that don't have inherent price-tags. For example, the construction of a new dam may generate electricity but displace people. It may create a beautiful lake but demolish the habitat of a river snail. The calculus of identifying and accounting for the various costs and benefits has been particularly developed in the field of resource and environmental economics and has been appropriately named non-market valuation.

In the first essay I show some evidence that a current approach to valuing the benefits of environmental improvement is severely underestimating the true benefits. In the second essay I show that farmers value their soil beyond what it produces. In the third essay I provide evidence that a large-scale project in India provided - at best - modest benefits to groundwater levels.

Conditional logit (and more recently mixed logit) models are a popular class of models that can model an entire demand system by observing individual choices. The breakthrough work of McFadden (1974) has been applied numerous times to estimate recreational, agricultural, and environmental choices to ultimately estimate the value of changes to the environment, which is often an important part of the benefit-cost calculus for regulations. It is increasingly common to observe choices in a panel format, and in panel settings, the inclusion of fixed effects is standard to control for a myriad of characteristics that we don't observe

[1] Executive Order 12291

in demand systems. The larger class of fixed effects models are praised for their ability to control for a wide swath of unobserved effects, but the well-known curse is that characteristics that don't vary over time (or any other singular dimension) are not identifiable: they are often differenced-out and forgotten about.

Water quality clearly changes over time. Anyone who has driven over multiple bridges in a day knows, however, that water quality also changes over space. The dominant approach for estimating the benefits of water quality improvement uses water quality changes over time to identify conditional logit choice models. They use fixed effects, and forget about the change over space. Existing methods in the non-market valuation (Murdock, 2006) and political science (Plümper and Troeger, 2007) literature show that it is entirely possible to use a two-stage regression to recover both the *within* (time-varying) and *between* (space-varying) effects of a model. I apply these methods and show that changes across space matter much more to individuals than changes across time. In other words, permanent changes matter more than temporary changes. This finding is miraculously novel, and means that the current approach to estimating the benefits of environmental improvement may significantly underestimate the true benefits.

Upstream from the anglers experiencing changes in water quality, farmers manage the soil that may cause the changes. Farmers rely on their soil for their agricultural production, and this channel has dominated the way we calculate the benefit of soil quality for farmers since the middle of last century (Ciriacy-Wantrup, 1947). Clearly, if you improve your soil and that improves your yield, that is a convenient way to put a price-tag on the benefits of whatever you did to improve your soil. It is nonetheless only a partial picture of the total benefits. To get a more complete picture of the benefits, I talk to farmers and essentially ask them about how they value their soil.

Using observations of their behavior is difficult or impossible because of the lack of data. Asking them directly is problematic because of hypothetical bias (Murphy et al., 2005). A team of scientists and I instead create a choice experiment where we take great care to create a hypothetical but realistic scenario where we can observe choices dealing with soil quality. By observing the trade-offs farmers make with their soil - something impossible to do with existing data - I estimate the value of soil quality improvements removed from the confounding effects of conservation practices and production decisions. My estimate of the benefits of soil quality improvement stand alone in its separation from agricultural production, and more work needs to be done to both validate my estimates and to estimate similar measures for other areas around the country and world.

The implications for my estimates of the benefits of water and soil quality improvements are potentially large. Any regulation compares potential costs and benefits, and my estimates have never been accounted for in regulation decisions. No one has estimated the value of changes in long-term conditions, and no one has worked with farmers to value soil apart from production. These new values tip the scale in favor conservation, and could alter the ultimate decision to enact some conservation policy.

It is also important to understand the costs and benefits of previous policies when regulating for the future. Indeed, the modern empirical economics field is dominated by the evaluation of previous policies. Recently, a large-scale groundwater management program in India sought to conserve groundwater resources through better monitoring. There is currently another proposal for a massive duplication of that project, since there has been some evidence that the program was successful. I apply causal econometric methods to formally analyze the project, and find that the program was largely unsuccessful. There is some evidence of weak success, but the benefits to groundwater levels are just that: weak.

This work focuses on the benefits of conserving agricultural resources. Not all conservation efforts are successful, and I explore a particular partial failure. I also provide an estimate of benefits that does not rely on conservation success to be true. When conservation *is* successful, benefits accrue to downstream individuals, and I show an easy way to properly capture a fuller picture of those benefits.

2. THE VALUE OF SHORT AND LONG TERM CHANGES IN WATER QUALITY

2.1 Introduction

The current standard approach in valuing changes in water quality using recreation demand data is to focus on contemporaneous water quality conditions and use fixed-effects estimators to control for all time-invariant heterogeneity across sites (Moeltner and Von Haefen, 2011; Dundas, von Haefen, and Mansfield, 2018; Melstrom et al., 2015; Von Haefen and Phaneuf, 2008). This is powerful but costly because the effects of any observable time-invariant characteristics can not be simultaneously identified alongside site fixed-effects. The value of any time-invariant effect in panel recreational demand models is swept under the rug of the alternative-specific constants (ASCs). In this paper, we show that the value of long-term water quality conditions (time-invariant by construction) can be estimated, and ignoring these values potentially results in significant undervaluation of policies that would lead to improvements in water quality.

There is good reason to suspect that variation in long-term conditions may have a larger effect on a recreational fishermans choices than short-term variation. First, in most settings the choice set available to anglers will include sites that have consistently different fishing conditions. When considering where to fish, an angler may have a better understanding of these long-term typical conditions than the current conditions of all sites in the choice set. Moreover, recent work suggests that researchers may encounter substantial measurement error in using current water quality conditions (Keiser, 2019). Anglers, too, could perceive current conditions with substantial error, as the costs of acquiring the knowledge of current conditions for all sites could be large.[1]

Nonetheless, most modern recreational demand models use contemporary water quality conditions at the time of the choice occasion to model angler behavior (Pendleton and

[1] The difference between the value of an attribute and the value of *information* on that attribute is an important distinction that we do not address in this paper.

Mendelsohn, 1998; Parsons et al., 2009; Whitehead et al., 2018). As we will show, however, the standard inclusion of ASCs does not prevent the identification of the effect of long-term water quality variation on angler utility. Both Murdock (2006) and Ji, Keiser, and Kling (2020) have shown that the ASCs can be decomposed in a second stage regression to identify effects that only vary across sites. As we define it, short-term (or acute/contemporary) water quality varies across time and space, while long-term water quality only varies across space. In this paper we show that standard methodologies in the recreational demand literature can be adapted to simultaneously identify the effect of variation in short- and long-term water quality on angler utility, while still accounting for unobserved alternative heterogeneity. We also provide evidence that changes in long-term conditions affect angler utility much more than changes in short-term conditions, which has significant implications for welfare measurement.

To understand our basic model, first consider a simple form of indirect angler utility:

$$V_{ijt} = X_{jt}\beta + C_{ijt}\gamma + \alpha_j, \tag{2.1}$$

where the anticipated indirect utility of fishing for angler i choosing site j on choice occasion t is a function of the water quality characteristics (X_{jt}) of site j at time t, and the costs for angler i to get to site j on occasion t (C_{ijt}). The inclusion of site-fixed effects (α_j) has become standard, since there may be unobserved site-specific characteristics that, when not controlled for, result in a biased estimate of γ and thus biased welfare results (Murdock, 2006). The inclusion of α_j controls for *all* time-invariant site-specific features of the choice set for any angler. However, researchers may be interested in some time-invariant site-specific characteristics that affect anticipated angler utility.

We hypothesize that long-term water quality also affects an angler's anticipated indirect utility, i.e.

$$V_{ijt} = X_{jt}\beta^S + \bar{X}_j\beta^L + C_{ijt}\gamma + \alpha_j, \tag{2.2}$$

where $\bar{X}_j$ is the long-term water quality specific to site j that does not change over the period of study, and β^S and β^L capture the effect of variation in short- and long-term water quality on anticipated indirect utility, respectively. If $\beta^S = \beta^L$, then short- and long-term changes in water quality affect indirect utility by the same magnitude, and the exercise of separately identifying the effects becomes moot. There is no reason, however, to assume changes in short- and long-term conditions affect indirect utility by the same magnitude. For example, long term water quality conditions may be widely known and play a major role in where an angler goes to fish, but current conditions (as they deviate from the long-term average) may be less certain and less widely known. Even though X_{jt} and $\bar{X}_j$ are measured on different time-scales, their effect on indirect utility in (2.2) is on the same scale: a single choice occasion. If (2.2) is the true form of indirect utility, there is a problem: we cannot simultaneously identify α_j and β^L. One option would be to assume that α_j in (2.2) is a random effect, thus relegating it to the error term. As mentioned above, this may lead to biased estimation when time-invariant characteristics are correlated with other observed characteristics.

The classic fixed effects estimator is a *within*-estimator, which uses only within-unit variation for identification (Wooldridge, 2010). Thus, if we could estimate (2.2) directly using OLS, β^S would be the marginal effect of $(X_{jt} - \bar{X}_j)$ on anticipated utility, and both α_j and β^L would be differenced out.[2] There are several approaches to recover β^L. The first is to use a between estimator, which is a cross-section OLS regression on the de-meaned values of regressors and regressands. The second is to use a hybrid model that would allow direct estimation of (2.2), but this requires the assumption that after controlling for $\bar{X}_j$, the remaining site-specific fixed effects must be random and zero in expectation (Allison, 2009). A third approach is fixed effect vector decomposition (Plümper and Troeger, 2007), which

[2]The within-estimator in a fixed effects framework subtracts the time-averaged value from each regressor, so any effect that is constant over time is "differenced out" and is not used in identification.

proceeds in two stages. The first stage would use conditional logistic regression to estimate

$$U_{ijt} = X_{jt}\beta^S + C_{ijt}\gamma + \alpha_j + \varepsilon_{ijt}, \tag{2.3}$$

and the second stage uses the vector of estimated α_j to estimate

$$\hat{\alpha}_j = \alpha + \bar{X}_j\beta^L + u_j. \tag{2.4}$$

Any of the three approaches may work if utility were directly observable, but none of the utility models above can be directly estimated.

In the recreational demand literature, random utility models (RUMs) are used to estimate the key features of recreationalists' utility. Researchers frequently estimate models that assume that utility takes the form of equation 2.1, and we use this approach to estimate the short-term effect, β^S. However, it is still possible to recover the long-term effect, β^L, in equation 2.2. We use a two-stage approach first suggested by Murdock (2006) and recently used by Ji, Keiser, and Kling (2020) that is analogous to the fixed-effects vector decomposition (Plümper and Troeger, 2007) in a choice model framework.

While the methods employed here would be ideally suited to panel data, such data sets for recreation demand are extremely rare (e.g., Yi and Herriges (2017)). We use a more common type of data set, a pooled-cross section of angler surveys. The surveys we use were administered in the state of Texas from 2001 - 2015 and we combine these data with a panel of water quality data from sites across the state. In the first stage we estimate a random utility model (RUM) to recover the marginal effect of travel costs and short-term water quality conditions on anticipated angler utility, while controlling for any unobserved site-specific heterogeneity. Then in the second stage we regress the estimated vector of site fixed effects on time-invariant water quality, recovering the long-term effects on angler utility. We apply the results of our model to a simulation that improves water quality for a watershed in central Texas, and find that per-trip benefits for long-term improvements are larger than per-trip benefits for short-term improvements. Our results suggest that valuing water quality using only short-term variation is a significant undervaluation of the willingness-to-pay for permanent changes in water quality.

2.2 The Model

We hypothesize that an anglers choice of where to fish is a function of the current water quality conditions and the long-term average conditions at each site on a given choice occasion. Our data are in a repeated cross-section format, and we use a conditional logistic regression to model only the site-choice decision conditional on the participation decision already having been made. We assume a standard random utility model (McFadden, 1974), in which angler i chooses sites based on his or her anticipated utility from site j on choice occasion t:

$$U_{ijt} = X_{jt}\beta^S + \bar{X}_j\beta^L + C_{ijt}\gamma + \alpha_j + \varepsilon_{ijt}, \tag{2.5}$$

where α_j is an alternative-specific constant (ASC) to control for unobserved heterogeneity across sites, X_{jt} is a vector of water quality characteristics unique to site j in year t and $\bar{X}_j$ is a vector of long-term water quality measurements. Finally, C_{ijt} is a measure of the travel cost for angler i to get to site j with costs unique to year t, and ε_{ijt} is an i.i.d. random term with a Type I extreme value distribution.

Following the standard random utility model, we assume that angler i chooses site j at time t when the expected utility from choosing site j is the maximum of all possible sites k in the choice set J, i.e.:

$$U_{ijt} \geq U_{ikt}, \quad \forall k \in J. \tag{2.6}$$

Ideally, we would model the probability that angler i chooses site j at time t as:

$$P_i(j,t) = \frac{e^{V_{ijt}}}{\sum_{k \in J} e^{V_{ikt}}}, \tag{2.7}$$

where V_{ijt} is the deterministic portion of equation 2.5. However, α_j and β^L cannot be simultaneously identified, since α_j and $\bar{X}_j$ both vary only over j. We therefore proceed in a two-stage estimation process. In the first stage, we use conditional logistic regression to identify the following indirect utility function:

$$V_{ijt} = X_{jt}\beta^S + C_{ijt}\gamma + \hat{\alpha}_j. \tag{2.8}$$

In the second stage, we use simple OLS and regress the estimated ASCs ($\hat{\alpha}_j$) on long-term water quality measures:

$$\hat{\alpha}_j = \alpha + \bar{X}_j\beta^L + u_j. \tag{2.9}$$

The estimated slope coefficient, β^L, captures the effect of long-term water quality conditions at site j on angler is anticipated indirect utility. The intercept plus residual in this second stage regression, $\alpha + u_j$, captures the remaining unobserved heterogeneity across sites, i.e. $\alpha_j = \alpha + u_j$ from equation 2.5. The two stage estimation process is analogous to Murdock (2006), with two differences. First, Murdock uses a random parameters logit (RPL) (McFadden and Train, 2000) which allows each individual to have a separate preference for site characteristics. Second, the site characteristics in Murdocks model do not vary over time. In our work, we emphasize that when a site characteristic varies over time, its average over time is itself a time-invariant characteristic and one that has potentially important economic significance.

We use a conditional logit in the first stage, which as Murdock explains is a simplification of the RPL.[3] Two well-known threats to inference in recreational demand models come from unobservable and endogenous site characteristics (Moeltner and Von Haefen, 2011). Our first stage of estimation controls for unobservable site-specific characteristics and we are able to identify the ASCs directly, in part because the number of sites greatly exceeds the number of water quality characteristics in our model (Von Haefen and Phaneuf, 2008). Identification

[3]The coefficient on site characteristics in the conditional logit is the mean taste parameter from the RPL, and the standard deviation draws from the RPL are omitted in the conditional logit.

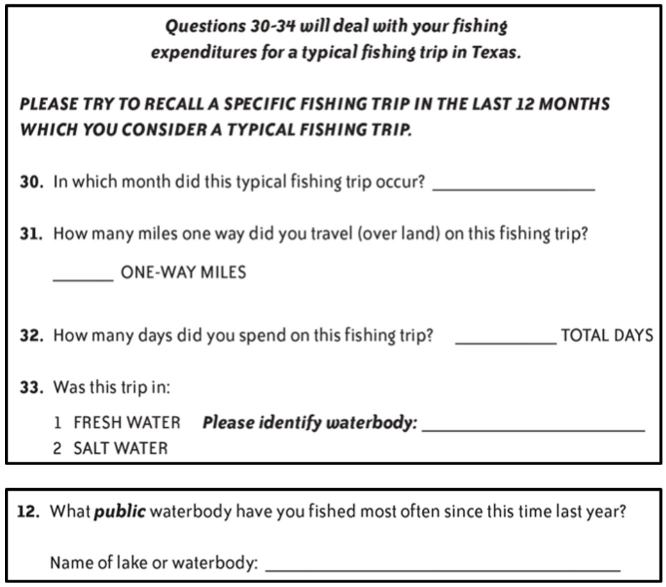
of our second stage would be biased if there is any remaining endogeneity in equation 2.9, e.g. if angler choices affect average water quality, in which case unbiased identification of the second stage would require an instrumental variables approach (Moeltner and Von Haefen, 2011; Murdock, 2006). In our context, however, it is unlikely that endogeneity will bias our estimates. Unlike catch rates or congestion (two common culprits of endogeneity), we do not believe the measurements used to capture water quality in our analysis are affected by angler choices. Hence, we assume anglers make the decision on where to fish based in part on the exogenously determined water quality measurements.

2.3 Data

Our site-choice and water quality data are obtained from two sources. Recreational fishing data were drawn from the periodically administered Texas Parks and Wildlife Department (TPWD) *Survey of Texas Anglers*. We use 5 surveys administered between 2001 - 2015, which elicit information on a single choice of fishing location, allowing us to estimate a site-choice model. Water quality data are obtained through the Texas Commission on Environmental Quality (TCEQ) *Surface Water Quality Monitoring Information System*. Information on dissolved oxygen, specific conductance, transparency, and pH for each fishing site in Texas is averaged for each site for a given year to arrive at a single annual measure of water quality. By spanning a 15-year period with a consistent measure of angler behavior and water quality, these data offer a unique opportunity to explore how water quality affects angler choices.

2.3.1 Angler data

Beginning in 2001, the TPWD included two questions in the *Survey of Texas Anglers* about the fishing sites that anglers choose. One question, hereafter referred to as the *typical-trip* question, asks each respondent to "…recall a specific fishing trip in the last 12 months which you consider a typical fishing trip," and subsequently asks respondents to identify the type of waterbody, and if it is freshwater, to name the waterbody. Respondents are asked to provide information on the timing of the trip (month of the year), one-way travel distance,


Questions 30-34 will deal with your fishing expenditures for a typical fishing trip in Texas.

PLEASE TRY TO RECALL A SPECIFIC FISHING TRIP IN THE LAST 12 MONTHS WHICH YOU CONSIDER A TYPICAL FISHING TRIP.

30. In which month did this typical fishing trip occur? ________________

31. How many miles one way did you travel (over land) on this fishing trip?

______ ONE-WAY MILES

32. How many days did you spend on this fishing trip? __________ TOTAL DAYS

33. Was this trip in:

1 FRESH WATER **_Please identify waterbody:_** ______________________

2 SALT WATER

12. What **_public_** waterbody have you fished most often since this time last year?

Name of lake or waterbody: ___________________________________

Figure 2.1: *Typical-Trip* And *Most-Often* Survey Questions

days spent on the trip, and itemized estimates of the cost of the trip. Another question, hereafter referred to as the *most-often* question, asks each respondent: "What public waterbody have you fished most often since last year?" Despite their apparent similarities, these questions do not always elicit the same response from anglers, so we estimate our models separately using both questions. As described below, we prefer using the *typical-trip* question because of the additional information embedded in the question and better framing. Images of the *typical-trip* and *most-often* questions can be seen in Figure 2.1.

In the survey data we observe the 5-digit zip-code of the home address,[4] and for each Texas zip-code we find the geographic centroid using U.S. Census Bureau data. For each reported trip, we observe the name of the fishing site, but not the access point. For lakes, we find the geographic centroid of the waterbody using Google Maps.[5] For rivers, however,

[4]The data for the 2001, 2004, and 2009 surveys initially included only 3-digit zip-codes; the additional digits had been lost with the death of the principle investigator (Robert Ditton). Thanks to diligent data management, we were able to contact other researchers associated with the surveys to recover the lost information. Special thanks to Dr. David Anderson for the 2001 information, Dr. Yung-Ping Tseng for the 2004 information, Dr. Adam Landon for the 2009 information, and generally the Human Dimensions of Natural Resources Lab at Texas A&M University.

[5]Using the centroid of a lake could be problematic if the only nearby road substantially lengthens the

there are many potential access points that might be hundreds of miles apart. There are two common approaches of setting an access point when the access point is not observed. The dominant approach uses the nearest access point to the angler's home zip-code, and another approach uses the midpoint of the river. Recent evidence suggests the second approach can lead to significant bias in the estimated travel cost parameter (Ji, Herriges, and Kling, 2016), so we opt for the first approach. We manually tag access points for each river from the survey by virtually floating downstream and marking points where roads intersect or end at the river. To obtain a specific site on a river for each angler, we first calculate the distance from the anglers home zip code to every access point along the river. The nearest location is used as the site location on that river in the anglers choice set. The location on the river therefore varies depending on the anglers home zip-codes.

Travel distances are calculated using the Stata package *osrmtime*,[6] which uses the *Open Source Routing Machine*[7] to generate optimal routes by car. Self-reported distances are available for the *typical-trip* question and, although the elicited distance may suffer from recall bias (Mazurkiewicz et al., 1996), the self-reported and distances and those calculated using *osrmtime* share a correlation coefficient of 0.32.

We calculate the cost angler i would incur by traveling to site j in year t as:

$$C_{ijt} = \rho\left(\frac{income_i}{2000} \cdot duration_{ij}\right) + \eta_t(dist_{ij}). \tag{2.10}$$

Information on annual income is elicited in the survey.[8] Dividing by 2000 approximates the

route to get as close to the centroid by road as possible. The routing software addresses this by allowing for "jumps"; when a route is close to the centroid, the route can leave the road and jump to the centroid or take a different road to get a little closer to the centroid, potentially lengthening the travel distance. The routing software will make the jump rather than lengthen the route, and reports the length of the jump for every route. The user can determine ex-post if some jumps are too large and selectively re-route to get a more realistic estimate.

[6]https://www.uni-regensburg.de/wirtschaftswissenschaften/vwl-moeller/medien/huber/osrm_paper_online.pdf

[7]http://project-osrm.org/

[8]Income is reported as a categorical variable, where each survey contains several total income bins. For any angler that reports income within a given bin, we take the median value of the bin and consider that the angler's income. For example, if an angler reported income of $25,000 - $35,000, we assume that angler's income in $30,000. The lowest/highest income bin is typically "below/above $15,000/$100,000", and in this

salary as an hourly wage. The hourly wage is multiplied by the estimated round-trip travel time ($duration_{ij}$: the time in hours for angler i to travel to site j and back - calculated in *osrmtime*) to capture the value of travel time (VOTT). Next, the VOTT is weighted by a scalar that discounts the VOTT, ρ , assumed to be 0.33.[9] The two-way travel distance estimated in *osrmtime* is multiplied by η_t: the per-mile cost of travel for an angler in year t.[10]

Each survey iteration was administered independently, drawing a random sample from the set of licensed anglers each year. Demographic information on age, race, gender, zip-code, and income is not sufficient to uniquely identify individuals through time, so traditional panel estimation or repeated choice models are unavailable. Instead, we estimate a pooled cross-sectional model.

Demographic data on Texas anglers from usable responses are summarized in Table 2.1. Across all survey years, the average age is approximately 47 - well below typical retirement age.[11] Average income across all years is approximately $83,800 in 2015 U.S. dollars, though we observe wide variance. In particular, respondents from the 2012 survey have a noticeably higher income than other years. We compare the demographics of the survey respondents to the United States Fish and Wildlife Service (USFWS) *National Survey of Fishing, Hunting, and Wildlife-Associated Recreation* results for Texas for 2001, 2006, and 2011 (the only years of the USFWS survey), and find similar observable characteristics between our data and the USFWS surveys. Average one-way distance exhibits similarly wide variance, which ranges from nearly 194 miles in 2001 to just 60 miles in 2015.

---

case we assume the angler's income is half the width of the other income bins away from the lowest/highest amount. For example, if the lowest bin was "below $15,000" and the average bin width was $10,000, we assume the anglers with the lowest income earn $10,000 per year. On the other hand, if the highest bin was "above $100,000" and the average bin width was $10,000, we assume the anglers with the highest income earn $105,000 per year. We drop anglers with missing income observations.

[9]We test a range of weights from 0.1 to 0.9, and the weight has little impact on the results.

[10]We use per-mile cost estimates from the American Automobile Association (AAA), assuming the angler is using a 4WD Sport-Utility Vehicle (SUV) and drives 10,000 miles per year.

[11]If the sample had a greater proportion of individuals over the age of 66 (Full Retirement Age, e.g. eligible for 100% of Social Security benefits), using foregone wages in formulating travel costs could be problematic, since retirees have no wage to forgo.

Table 2.1: Summary Of Angler Survey: Demographic Information

	Texas Parks and Wildlife Department Survey					USFWS National Survey: Texas		
Year	**2001**	**2004**	**2009**	**2012**	**2015**	2001	2006	2011
Age	46	44	46	48	49	35-44	35-44	45-54
	(11)	(12)	(13)	(15)	(11)			
Income (2015 USD)	86,719	77,443	73,603	98,984	84,211	66,735 - 100,102	58,928 -88,390	79,591 - 106,122
	(39,771)	(37,230)	(35,574)	(48,792)	(48,536)			
Reported Distance (One Way Miles)	66	74	66	94	60			
	(81)	(81)	(74)	(136)	(82)			
Calculated Distance (One Way Miles)	194	131	81	97	73			
	(121)	(139)	(85)	(105)	(93)			
Calculated Travel Costs (2015 USD)	452	341	258	328	247			
	(217)	(259)	(181)	(230)	(193)			
Percent White	90	93	91	96	93	97	90	75
Percent Male	86	85	83	91	80	71	73	76

Notes: Summary statistics for the Texas Parks and Wildlife Department Survey of Texas Anglers and United States Fish and Wildlife Service National Survey results for the state of Texas. Standard deviations in parentheses. The USFWS National Survey does not provide averages for age and income of respondents. Instead, the median category is reported in the table. The median income category has been inflated to reflect 2015 USD.

2.3.2 Water quality data

Our water quality data are obtained from TCEQ's *Surface Water Quality Monitoring Information System* (SWQMIS). This database contains information on 5,980 parameters[12] at waterbodies across the state. Unfortunately, the data are neither uniform nor comprehensive. Instead, a handful of parameters are measured per site per visit and monitoring visits occur at the discretion of volunteer monitoring organizations, so the frequency of measurement per site is not consistent.

We choose pH, specific conductance, dissolved oxygen, and transparency as our water quality parameters for several reasons. First, these four parameters are taken directly in the field as opposed to a lab sample that occurs sometime after a water sample is taken. Second, these measurements are the most often measured parameter group, and often taken as a bundle together. Third, we focus on water quality parameters that can be affected by upstream management, which is the policy-relevant scale for water-quality improvements across the state. Soil conservation, for example, may result in decreased levels of sediment and nutrient runoff, thus affecting downstream pH, specific conductance, dissolved oxygen, and transparency. Recent evidence suggests nutrient pollution is a growing problem in Texas Kuwayama et al. (2020).

Finally, we choose parameters that affect the quality of fishing for anglers. All four measurements can affect the health and abundance of fish species (Alabaster and Lloyd, 2013; Miller, Bradford, and Peters, 1988). Dissolved oxygen is a measure of how much oxygen - a vital resource for fish - is in the water. While different fish species have different low-oxygen tolerance levels, we expect the marginal effect of dissolved oxygen on angler utility to be positive. Specific conductance is a measurement of water's ability to pass electric flow. While pure distilled water is a poor conductor of electricity, conductance increases as water contains more metallic ions from dissolved salts or inorganic material as electricity can pass from ion to ion. Specific conductance (electrical conductance corrected for temperature) is

[12] https://www.tceq.texas.gov/assets/public/waterquality/crp/data/ParameterInventory_12202018.txt

relatively stable across time, but will change during high rainfall events when runoff picks up soil sediment and distributes it downstream.[13] We anticipate the marginal effect of specific conductance on angler utility to be negative, since higher levels of specific conductance indicate more dissolved salts/ions from soil runoff since higher levels of dissolved salts are detrimental to most freshwater fish species, which are the focus of this work. The pH of water is a measure of acidity, measured on a unit-less scale ranging from 0 - 14, with 7 being a neutral value. As water becomes either too acidic or too basic, fish populations may be detrimentally affected. The marginal effect of pH on angler utility may likely take the shape of an inverted-U, with high and low values of pH being undesirable. In our data, however, we do not observe extreme values of pH and averages are above 7 so that for the most part, larger values are further from neutral. Hence we assume a linear relationship over the relevant range of values. Finally, we expect increasing transparency of water to positively affect an angler's utility: the ability to see for both the angler and the sought fish should be a desirable attribute.

In a more data-rich setting, the month of the trip from the survey could be matched to the same month's water quality information. However, both our site choice and water quality data are too sparse. On average, only 11% of stations have water quality observations for every month of the year. Hence, we are forced to use a single annual choice and a single annual measure of water quality. To arrive at the single annual measure of water quality, we take several rounds of averages for each quality parameter. First we average across depths, then across time (days, months, year) and finally, if a site has multiple stations, we take a average across the nearby stations since we do not observe the access point of the lake used by the angler and access points may change across anglers. We only consider measurements that are taken between the hours of 7am and 8pm, since measurements taken at night may not be representative of the fishing conditions experienced by anglers. Finally, because water quality changes may be slow and there is a gap of three or more years between each survey,

[13]https://www.waterboards.ca.gov/water_issues/programs/swamp/docs/cwt/guidance/3130en.pdf

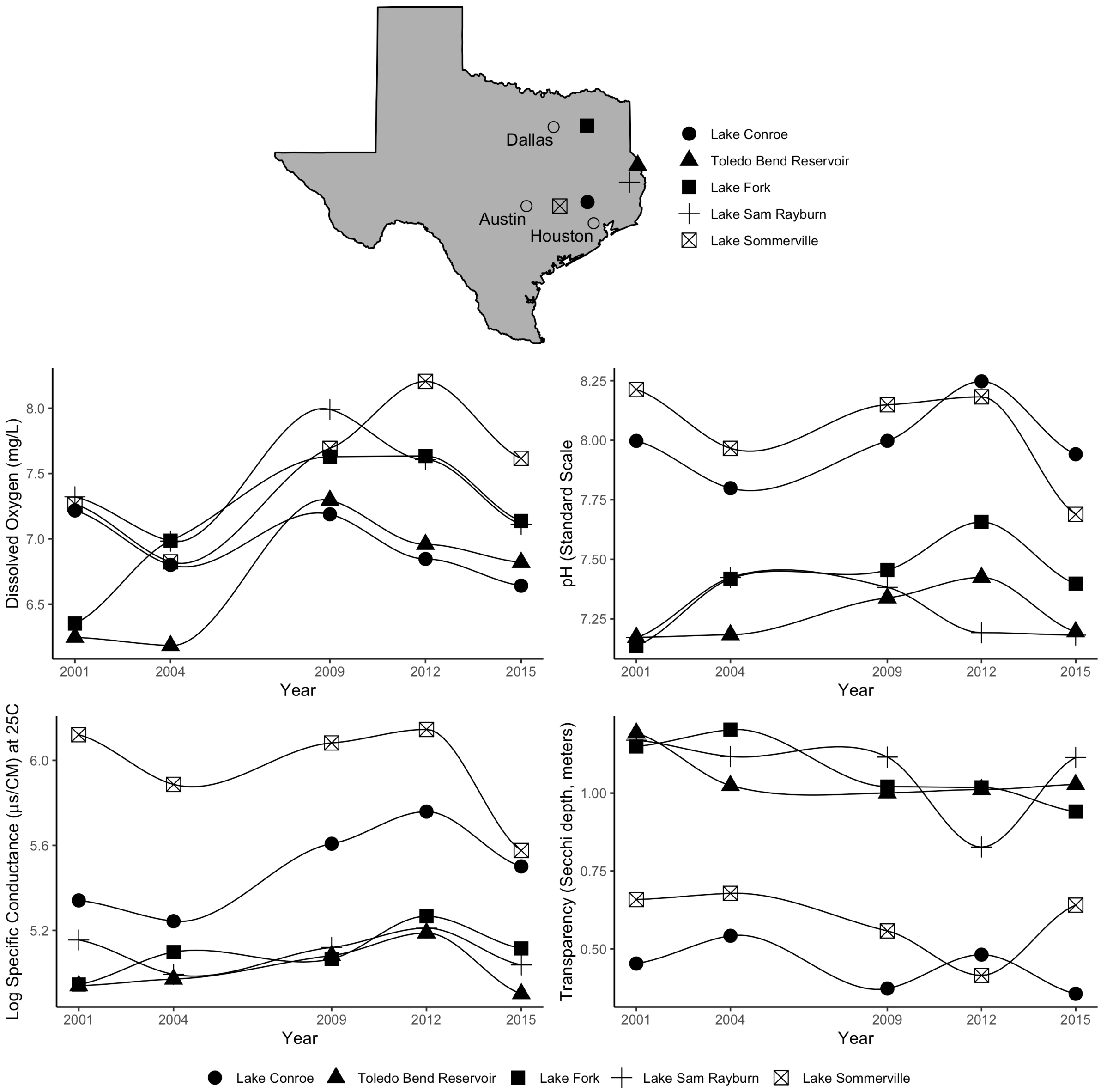


Figure 2.2: Water Quality Over Time For Most-Visited Sites

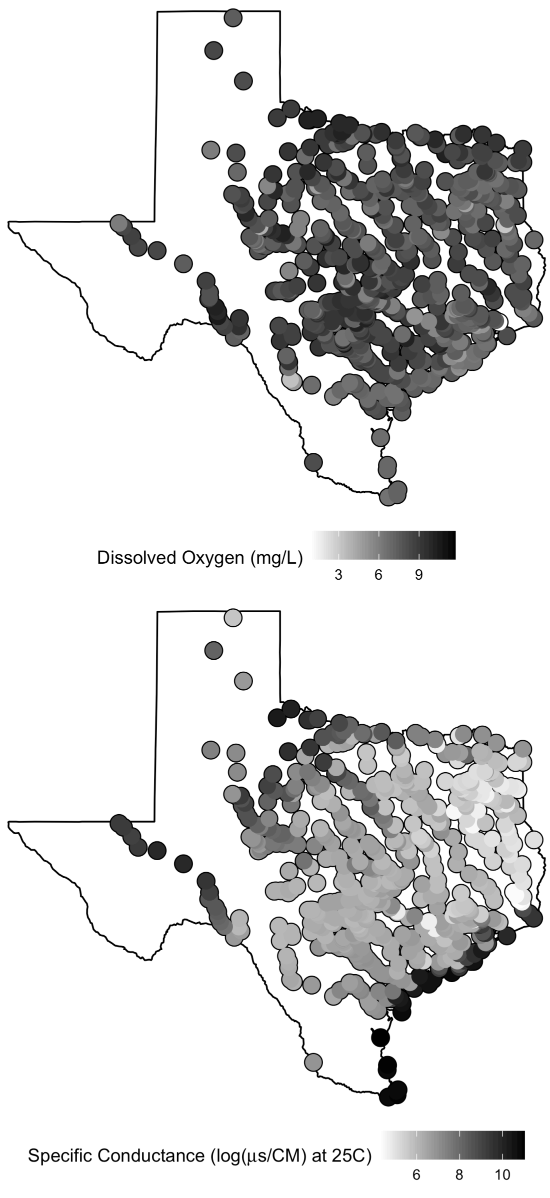


Figure 2.3: Long-term Water Quality Across Texas

we include the leading 10 months after the survey. We also include the lagging 10 months, since anglers may have experienced water quality at some sites during the previous fishing season. Including the leading and lagging months also increases our sample size of water quality monitoring stations: on average, the number of stations with at least one observation increases by 8% when we include the extra months.

Water quality along our selected dimensions varies over time and space. For example, Figure 2.2 illustrates how water quality changes over time for five popular sites. It is apparent that some sites are chronically different from other sites. Lake Sommerville and Lake Conroe, for example, have consistently higher specific conductance and lower transparency than the other popular sites.

Figure 2.3 shows the spatial distribution of long-term water quality conditions. Each point on the map represents a single water quality monitoring station, and the color indicates the long-term average from 2000 - 2016. Areas near the Gulf Coast or in the western part of the state have higher specific conductance, which is likely the result of saline water or agricultural runoff. The eastern part of the state has higher pH, likely from underlying parent material of local soil, or from the higher incidence of pines (whose needles, when dropped, are acidic). The central part of the state appears to have clearer water, while there is no clear spatial pattern of dissolved oxygen.

2.4 Estimation Results

Table 2.2 presents the results of the estimation of both the short-term (first-stage) and long-term (second-stage) models. The alternative-specific conditional logit (site-choice RUM) and the OLS linear model are estimated using Stata 14.2. The first-stage is estimated using the *asclogit* command with clustered standard errors.[14] The second-stage is estimating using the standard *reg* command, with heteroskedastic-robust standard errors. After dropping missing information for water quality and/or travel cost, we estimate the models using observations for 1,687 trips taken across the 5 survey years to a set of 198 sites across Texas.

[14]Similar to Dundas, von Haefen, and Mansfield (2018), we cluster at the individual level (i.e. angler).

Table 2.2: First And Second Stage Estimation Results

	Short-Term	Long-Term
Travel Cost	-0.00609***	
	(0.00)	
Dissolved Oxygen	0.102*	-0.0486
	(0.0586)	(0.551)
pH	0.537***	4.473**
	(0.204)	(1.908)
Specific Conductance	-0.311***	-1.788**
	(0.107)	(0.692)
Transparency	0.316***	3.042***
	(0.112)	(1.166)
Constant		-33.08**
		(13.02)
Observations	262,103	198
R-squared		0.089
Adjusted R-squared		0.070
AIC	11,978.2	
Log-Likelihood	-5,781.1	

Standard errors in parentheses
**p<0.1, **p<0.05, ***p<0.01*

The short-term model contains travel cost, the four water quality measures, and a suite of alternative-specific constants (ASCs - not presented in Table 2.2) as explanatory variables. The presence of the ASCs controls for unobserved heterogeneity at the site level (Murdock, 2006). There may be similar unobserved heterogeneity across anglers, which we could potentially address by allowing the preference parameters on travel cost or water quality vary by individual. However, there is evidence that specifying a random-parameters logit (RPL) with ASCs leads to poor in-sample prediction (Klaiber and von Haefen, 2011). The repeated cross-sectional nature of our survey data also prevents the use of an RPL. The long-term model contains the four long-term measures of water quality as explanatory variables.

Table 2.2 presents the main estimation results. First, we estimate the short-term model by assuming anglers have linear preferences over water quality (Column 1). Next, we estimate the long-term model using the estimated ASCs from the short-term model (Column 2).[15]

In the short-term model, we observe a significant negative coefficient on travel cost, as expected; anglers anticipate lower utility from sites that are more costly to travel to, *ceteris paribus*. Dissolved oxygen has a significant positive effect on anticipated utility; anglers anticipate higher utility for sites with higher levels of dissolved oxygen. While the sign of the coefficient is as expected, its relatively low statistical significance is surprising since dissolved oxygen is one of the most important factors influencing fresh water fisheries. The coefficient on pH is positive, suggesting for the Texas sites considered here, angler anticipated utility tends to increase as water becomes more basic. The coefficient on specific conductance is negative, suggesting that angler anticipated utility decreases with increased conductance. In other words, as a site is loaded with more salt and metallic ions, the specific conductance increases and anticipated utility of fishing at that site decreases. The coefficient on transparency is positive, suggesting that angler anticipated utility increases as water becomes clearer.

The estimated coefficients in the long-term model (Column 2, Table 2.2) are similarly signed to the short-term model, with the exception of dissolved oxygen. This suggests that the direction of the marginal effect for changes in water quality in either the short- or long-term will be similar. However, the *magnitude* of the coefficients between the short- and long-term models is very different: the long-term coefficients are approximately seven times larger than their short-term counterparts. As we discuss below, the stark difference in the relative size of the coefficients between the short- and long-term models is the main empirical finding of this study.

Before proceeding to a welfare analysis, we explore the robustness of our main results by

[15] There are $J-1$ ASCs identified in the short-term model. The omitted site is set to zero, and is included in the long-term model so there are J observations in the OLS regression.

estimating a range of alternate specifications with both the typical-trip question (Table A.2) and the most-often question (Table A.1). First, we present the base model identical to the short-term model in Table 2.2. Second, we estimate the base model using an alternate form of travel cost.[16] Third, we restrict the choice set for each angler to only include those sites that are 150 miles or less to each anglers home zip-code centroid. Fourth, we only include anglers who reported at least one trip over the prior 12 months, excluding the small group of anglers (less than 2%) who reported a typical-trip while simultaneously indicating they took no trips. Fifth, we estimate the base model but only include lakes. Sixth, we estimate the base model using rivers only. Seventh, we estimate the model dropping the 2001 survey because, as discussed above, respondents in the 2001 survey reported much longer trips than the other survey years. Finally, we estimate two additional specifications including only day trips (Column 8, Table A.2) and where the difference between the reported and calculated distance is less than 100 miles (Column 9, Table A.2), but only for typical trip question since the necessary complementary data were not available for the most often question.

In Tables A.3 and A.4, we estimate the long-term model using the "typical-trip" and "most-often" survey question, respectively, for the preceding short-term models. The alternate specifications are the same as in Table A.2, though for each specification we estimate the model with and without regional fixed effects.[17]

Across all the estimated specifications for the long and short-term models, the estimated coefficients are generally similar in sign, though the magnitudes and significance varies. In Table A.2, we observe that changes in short-term dissolved oxygen, pH, and transparency affect anticipated utility for lake-based anglers, while specific conductance does not. Conversely, changes in short-term specific conductance and dissolved oxygen are important for river-based anglers, and pH and transparency are not. Curiously, we do not observe the same pattern in the difference in preferences between lake and river anglers when using the

[16]We specify the alternate form of travel cost as: $C_{ijt} = \tau\rho\left(\frac{income_i}{250}\right) + \eta_t(dist_{ij})$, which uses daily income and we assume that the duration of the trip, τ is one day. In other words, we use the value of a day of fishing instead of the value of travel time.

[17]Regional fixed effects are used for each of the state's 24 major river basins.

"most-often" question.

The difference in estimated coefficients across model specifications suggests our point estimates are not robust. The lack of robust estimates is likely a function of sparse water quality data coupled with an imperfect survey instrument.

The coefficient from the short-term model, which is often used in recreation demand studies to estimate the value of a improvement in a site characteristic, captures the total value of improvements in those characteristics only if the coefficient in the short-term and long-term models are equal. Our data strongly rejects that equivalence. We find compelling evidence that anglers pay much more attention to long-term site characteristics when making their recreation choices. Hence, using the coefficients from the first-stage site-choice model, which only capture the impact of period-to-period variation on site choices, would result in a significant underestimate of the value of water quality to recreational anglers in Texas.

Furthermore, environmental regulation is crafted to improve the long-term conditions of sites. Given the reality of regulation and the results of our estimation, we show in the next section that when we only consider the short-term improvements of water quality (as has been done in previous recreational demand studies[18]), we substantially underestimate the benefits of water quality improvements.

2.5 Policy Simulation and Welfare

To consider the value of short- and long-term changes in water quality in a policy context, we use the results from our two-stage econometric model with the "typical-trip" question (Table 2.2) and simulate a hypothetical policy scenario. Specifically, we assume an improvement in water quality conditions for all sites in the watershed of the Brazos River (hereafter referred to as *the watershed*). The Brazos River is the third longest river in Texas, and its watershed is the second largest by area.[19] We select this particular watershed for three reasons: 1) It is centrally located, and thus we assume it has sites that are feasible for most anglers

[18]For example: Pendleton and Mendelsohn (1998); Parsons et al. (2009); Whitehead et al. (2018)

[19]From the Texas Water Development Board: https://www.twdb.texas.gov/surfacewater/rivers/river_basins/brazos/index.asp

to choose; 2) It contains 46 sites, making up 23% of the entire choice set for anglers across our survey years; and 3) Agricultural operations dominate the landscape, and we assume that the policy scenario we investigate could be physically attainable by altering agricultural land management. A map of the watershed and the sites affected by our scenario is seen in Figure 2.4.[20]

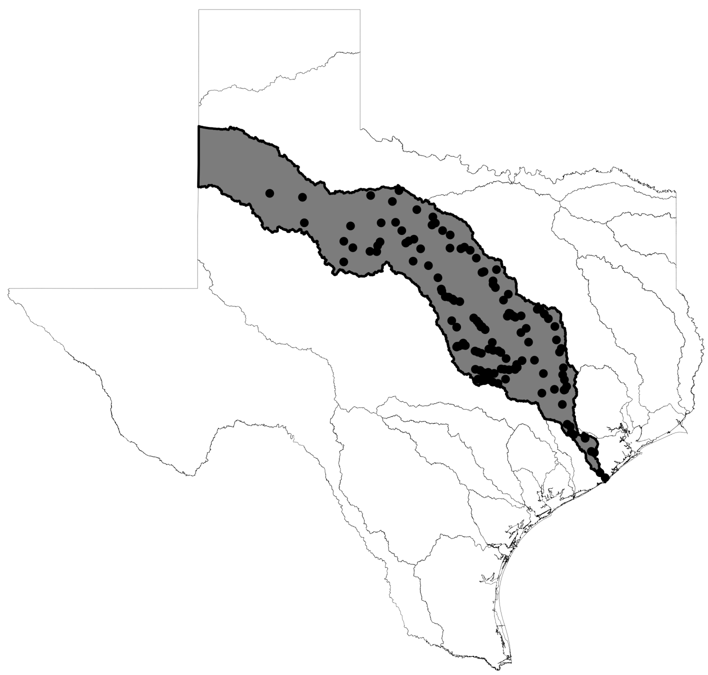


Figure 2.4: Map Of The Brazos River Watershed

In our simulation, we assume a hypothetical policy wherein conservation practices are widely adopted by farmers in the watershed, with the primary downstream impact being a reduction in sediment loads. Nutrient pollution - often associated with sediment runoff - is a growing and realistic problem for the state of Texas (Kuwayama et al., 2020). The effect of conservation practices on downstream water quality is often complicated, dependent on

[20] In Figure 2.4, there appear to be more than 46 sites. As explained previously, a river (like the Brazos) contains multiple fishing locations considered to be independent fishing sites. When we alter water quality in our hypothetical scenario, we do so for every site along each river as well as at every lake.

inherent geomorphological characteristics of watershed, and is often measured by nitrogen and phosphorus loads (Moriasi et al., 2020). A 16-year study on downstream impacts of conservation practices for a portion of the Brazos river watershed, for example, only measured downstream nitrogen and phosphorus loads and found mixed results (Smith, Harmel, and Haney, 2020). We do not have strong evidence how upstream agricultural conservation practices will impact downstream dissolved oxygen, pH, specific conductance, and transparency.

We simplify the hypothetical policy by focusing only on an improvement in transparency. We speculate that region-wide changes in agricultural practices (like reduced tillage, cover crops, terraces, buffer zones, etc.) will result in a reduction of sediment loads, leading to improvements in transparency.

Next, we decide on the level of transparency improvement. We calculate the standard deviation of transparency over time for all sites with more than three observations over the time horizon. The standard deviation of transparency for most sites in the watershed is less than 0.2 meters and the average standard deviation is 0.18 meters (approximately 7 inches). We make the simplifying assumption that the conservation policy scenario improves transparency at each site in the watershed by 0.18 meters. We apply the improvement in transparency to the short- and long-term conditions separately. That is, we first evaluate the welfare impacts of an improvement in transparency for all sites in the watershed by 0.18 meters with the assumption that this will be a one-period shock, holding long-term transparency constant at its status quo level. Then we carry out the same exercise but assume that the change leads to a permanent increase in transparency.

After creating a new vector of water quality conditions, we calculate the willingness-to-pay (WTP) per-trip for the change as:

$$WTP_i = \left[ln \left(\sum_{j \in J} exp(V^*_{ijt}) \right) - ln \left(\sum_{j \in J} exp(V_{ijt}) \right) \right] \cdot \gamma^{-1} \tag{2.11}$$

where V_{ijt} is the indirect utility of angler i choosing site j in year t, V^*_{ijt} is the same measure but with the new vector of water quality, and γ is the coefficient on the travel cost variable.

This standard equation (Haab and McConnell, 2002) yields the per-trip expected WTP for the change in water quality for angler i. It is calculated for each angler using parameters from all of the nine specifications and sample restrictions presented in Tables A.2 and A.3. Because the angler's choices are measured for a single choice occasion, both the short- and long-term impacts capture the effect on a representative angler's welfare at a single choice occasion; the duration of the welfare change is the same, the difference is whether the angler perceives the change as a short-term change or a permanent shift in water quality.

As seen in Table 2.3, we find slight differences in per-trip WTP across model specifications, and significant differences in per-trip WTP for short- and long-term transparency improvements. Mean WTP for a 0.18 meter increase in short-term transparency at all sites in the watershed ranges between -$0.27 and $2.56, depending on the model specification and sample restriction. Mean WTP for a 0.18 meter increase in long-term transparency ranges between $6.28 and $25.17. While the point estimate for mean WTP varies across models, there is a consistently significant difference between WTP for short-term and long-term changes; we find strong evidence that on a per-trip basis anglers are willing to pay more for long-term improvements in water quality than for the same change in the short-term. Since polices targeting water quality improvements generally seek long-term quality standards, analysts should be careful of how they are calculating the benefits of such policies. The current standard approach would only use short-term measures of water quality, which can significantly understate the true benefits.

To construct confidence intervals around the WTP point estimates for the base model, we use a bootstrapping approach. Using 200 independent random draws of 1,687 anglers, we calculate the WTP measurements for the base model for each iteration, keeping the mean WTP for each draw. The results are shown in Table 2.4. On average, the per-trip benefit to a Texas angler for short-term improvements to the watershed is $1.38, with a 95% confidence interval ranging from $0.67 to $1.96. The same measure per-season (calculated by multiplying the per-trip WTP by the number of reported trips per randomly selected

Table 2.3: Per-Trip WTP For Simulated Transparency Improvement In A Watershed

Model	Mean WTP: Short-Term	Mean WTP: Long-Term
Base Model	1.36	14.24
Alternate Travel Cost	1.09	11.45
Choice set restricted to 150 miles	-0.27	6.28
Positive trips only	1.17	15.37
Lakes only	1.34	8.08
Rivers only	2.56	25.17
Drop 2001 survey	0.32	7.90
Day trips only	0.60	7.89
Distance error of less than 100 miles	0.40	12.84

Notes: Mean WTP is presented in 2015 USD. Rows represent alternate model specifications.

angler) is \$36.46, with a 95% confidence interval from \$17.88 - \$52.57. If the water quality changes are perceived as long-term, the benefits are much larger. On average, the per-trip benefit for long-term improvements for transparency in the watershed is \$15.76 (95% confidence interval: \$0.50 - \$21.37), and the per-season benefit is \$417.57 (95% confidence interval: \$254.85 - \$561.83). As stated above, while we are not confident in the magnitude of the model coefficients, we are confident in the difference in the relative magnitude of the coefficients in the short- and long-term models.

Table 2.4: Bootstrapped WTP For Simulated Transparency Improvement In A Watershed

Statistic	Mean	95% CI: Lower	95% CI: Upper
Short-term, per trip	1.38	0.67	1.96
Short-term, per season	36.46	17.88	52.67
Long-term, per trip	15.76	9.50	21.37
Long-term, per season	417.57	254.85	561.93

Bootstrap results from 200 iterations of random angler selection from base model

2.6 Conclusion and Limitations

There is growing evidence that the recreational demand literature should consider important temporal aspects of water quality. Recent work shows that welfare estimates may not be stable over time (Ji, Keiser, and Kling, 2020), and using only site-specific short-term variations may result in significant measurement error (Keiser, 2019). Long-term water quality is a measure that is stable over time and plays a large role in the site-choice decision for anglers. We show that existing methods can be adapted to identify the effect of long-term variations on angler anticipated utility, while still identifying the traditional effect of short-term variations and controlling for unobserved site characteristics. Moreover, using standard fixed-effects estimate to value changes in water quality that are intended to be permanent will result in a significant underestimate of the true value. We are able to estimate the welfare impacts of improving short- and long-term water transparency and find significant differences between the two.

As far as our empirical estimates of the value of water quality to Texas anglers, there are several important limitations to our data that lead to imprecise value estimates. First, we do not observe the participation decision of the angler. For this reason, we are not able to capture the extensive margin of anglers, which may bias the welfare estimates. Second, we are missing substantial water quality data. The *SWQMIS* database has rich spatial data, but the frequency of measurements for many sites is quite low. With more measurements per site, we would have a better idea of the water quality conditions. Third, we do not observe the targeted fish species. If anglers are targeting specific types of fish, their preference for water quality conditions may change; some fish are more tolerant of muddy or briny water. For anglers targeting these species, low water transparency may be *preferred*. By failing to take into account angler preferences between clear- and muddy-water species, we are likely biasing the coefficients on water quality towards zero, which should result in more conservative welfare estimates.[21]

[21] Another approach could employ both ASCs and random parameters to account for angler heterogeneity.

Despite the data limitations, we believe our results strongly support the conclusion that anglers per trip willingness-to-pay is more for long-term improvements in water quality than for short-term improvements; we believe this is an important empirical contribution to the recreational demand literature. More generally, we show that standard coefficients estimated with site fixed effects capture only the value of short-term variation in water quality; the value of long-term water can be estimated in a second stage regulation and for many policy interventions this may be the more appropriate measure.

Some actions that might be undertaken to affect water quality have short-term consequences, while others have long-lasting impacts. Our work, therefore, has important implications for policy makers or environmental damage assessment, because we have shown that only accounting for the effect of short-term changes on angler welfare is not the complete story; the effect of long-term changes is potentially an important driver of site-choice, and thus an aspect of angler welfare that should not be ignored.

We opt for what Klaiber and von Haefen (2011) call a "second-best" strategy which prioritizes in-sample predictions, since the goal of this paper is to show the two-stage estimation process using ASCs, and since the repeated-cross-section nature of the data does not allow us to use a random parameters logit

3. AN EMPIRICAL ESTIMATE OF THE VALUE OF SOIL QUALITY IMPROVEMENTS

3.1 Introduction

During the past century, the United States has addressed soil conservation through a variety of conservation acts and services.[1] Today, the United States Department of Agriculture (USDA) spends approximately $6 billion each year on conservation programs.[2] Despite the evident importance of soil conservation, we have an incomplete picture of the value of soil quality improvements.

The benefits of soil quality improvements can accrue both on and off the farm, and recent literature has mostly focused on the off-farm benefits. It is widely recognized that soil provides a bevy of ecosystem services that are valuable to society, such as flood control, recreational benefits, and carbon storage (Dominati et al., 2014; Jonsson and Daviosdottir, 2016; Lal, 2014; Hansen and Hellerstein, 2007; Hansen and Ribaudo, 2008). Researchers have used a variety of non-market valuation tools to value the effect that soil conservation has on the provision of these ecosystem services.[3]

Meanwhile, estimation of the on-farm benefits of soil quality improvements is stuck in time. Early work on the on-farm economic value of soil quality focused on changes in crop revenue or return-on-investments associated with soil conservation efforts (Ciriacy-Wantrup, 1947). After the Soil and Water Resources Conservation Act of 1977, there was renewed interest in the value of soil conservation, though most work was anchored in theoretical optimal control models (Barbier, 1990; Barrett, 1991; McConnell, 1983; Burt, 1981). A seminal paper on the optimal control of soil resources from McConnell (1983) assumes that

[1]See, for example; Public Law 74-46 and the establishment of the Soil Conservation Service, Public Law 73-67 and the establishment of the Soil Erosion Service, the Flood Control Act of 1936, the Watershed Protection and Flood Control Act of 1954, and the Soil and Water Resources Conservation Act of 1977

[2]https://www.ers.usda.gov/topics/natural-resources-environment/conservation-programs/

[3]For an exhaustive summary of soil ecosystem service valuation studies, see Jonsson and Daviosdottir (2016).

a farmer seeks to maximize net revenues for her operation by considering the depth of soil as the only state variable. The economic literature on soil conservation since the 1980's has overwhelmingly focused on soil erosion Barbier (1990); Barrett (1991); Burt (1981); King and Sinden (1988); Lee (1980); Seitz and Swanson (1980), and benefits of soil conservation are typically captured in changes in revenues or input costs.

Estimating the on-farm benefits of soil quality improvements beyond erosion effects has received little attention. Nonetheless, farmers may value improvements in other characteristics like organic matter and water infiltration. In this paper, we directly estimate the willingness-to-pay to improve a set of soil quality characteristics without translating the changes to revenue. We use a direct stated preference approach since revealed preference approaches are not appropriate in this setting. Revealed preferences for soil quality characteristics elicited through choices of conservation practices are confounded by unobservable factors such as inherent (non-manageable, e.g. depth of horizons, clay content, and soil order) soil characteristics, effort of the farmer, and farmer-specific cost considerations. Revealed preferences could be elicited from observing land purchases, but there is no widely available data on manageable soil quality. We therefore build a hypothetical scenario to identify farmer preferences for manageable soil quality characteristics. Specifically, we utilize a discrete choice experiment (DCE) which is a popular approach to valuing specific attributes of a resource (Hanley et al., 1998). We follow a well-established framework for constructing a DCE (Holmes, Adamowicz, and Carlsson, 2017), opting for a multinomial choice-sequence version, where respondents face greater than two alternatives from which to choose (multinomial), and make a choice multiple times (sequence) (Carson and Louviere, 2011). We ask respondents to choose between two land parcels to rent for their operation, where each land parcel has different levels of manageable soil quality and rental rates, and a status-quo option where they may choose neither parcel.

The United States Natural Resources Conservation Service (NRCS) provides technical and financial assistance to farmers to adopt soil conservation practices. To effectively balance the costs and benefits of soil conservation, farmers and agents need to understand the monetary value of soil quality improvement. For example, the official benefit-cost template for no-till conservation lists potential on-farm benefits such as "[i]ncreased infiltration", "high soil organic carbon", and "reduce[d]...potential for soil compaction".[4] Our estimates of the willingness-to-pay for improvements in these characteristics is the first step in monetizing these benefits, which helps the USDA make more informed choices.

This work focuses only on estimating the perceived benefits of improved soil health. To be fully policy-relevant, our estimates should be compared to the costs of achieving soil quality improvements across different types of farmers to better understand which farmers need the most or least incentive to adopt soil conservation practices. Our work can ultimately lead to more efficient funding and targeting of conservation practices for the NRCS and similar agencies.

3.2 Survey Development

3.2.1 Study Area

The selected study area is the watershed of the Brazos River in central Texas. We are particularly interested in the middle portion of the basin, where farms share similar geological and climatic characteristics. The watershed consists of a few major land resource areas (MLRAs), as defined by the United States Department of Agriculture, Natural Resource Conservation Service (United States Department of Agriculture, Natural Resources Conservation Service, 2006). The major MLRAs in the region are the Texas Blackland Prairie and Texas Claypan areas. In both MLRAs, most soils are entisols, mollisols, and vertisols, and have an ustic soil moisture regime. Cropland and grassland are the dominant land covers and the major crops are cotton, corn, and sorghum. The average annual rainfall across both areas is 27 to 45 inches, and the average annual temperature is 64 to 70 degrees (United States Department of Agriculture, Natural Resources Conservation Service, 2006).

Ustic soil moisture regimes are soils that are intermittently dry and wet throughout the

[4] https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/econ/data/?cid=nrcseprd1298864

year (United States Department of Agriculture, Natural Resources Conservation Service, 2015). The soil moisture is generally adequate during the growing season but can be scarce during other times of the year. This differs from udic regimes, for example, which are consistently more moist throughout the year. Other MLRAs in Texas can be characterized by different soil orders and soil moisture regimes, for which management concerns (and therefore preferences for soil health characteristics) may vary significantly. We therefore target the Texas Blackland Prairie and Texas Claypan areas so that the difference in observed preferences is purely a function of manageable soil quality and not also a function of inherent geologic or climatic context for the farmer.

3.2.2 Attribute Selection

We adhere to standard best-practices laid out by Johnston et al. (2017) in design of the discrete choice experiment (DCE). For example, we use prior empirical results and pretesting to create a design that efficiently identifies key parameters[5]. When estimating preference parameters, previous estimates can be used to create more efficient experimental designs. A common metric to measure is a design's *D-efficiency* (described in Appendix), which is a measure of how efficiently an experimental design can identify coefficients of interest. Given the number of choice sets and attributes for an experiment, a researcher can maximize the D-efficiency of the design (Rose and Bliemer, 2009) by selecting combinations of attribute levels. This section describes the design process that ensures both the choice occasion and the alternatives are realistic to avoid a long-standing concern of hypothetical bias (Murphy et al., 2005) and consequentiality (Groothuis et al., 2017).

In both characterizing the decision problem and defining the attributes of the choice alternatives, we sought early input from farmers in the study area. In June of 2018, we met with two groups of farmers: those who had adopted soil conservation practices, and those who had not. The groups were recruited with the help of local extension service agents, who identified the adopting status of each farmer. Each group was asked a similar set of

[5]See recommendation 4 in Johnston et al. (2017)

questions. After reviewing the recorded transcripts of both group meetings, we found that water management, organic matter, yield, and biomass of crops were important measures of soil health (Bagnall et al., 2020).

After the focus groups, we constructed an initial DCE where farmers were asked to choose between parcels of land to rent, with different soil health attributes and price for each alternative parcel. The initial list of attributes was refined through conversations with soil scientists. For example, water retention is a measure of how much water is held in a given type of soil, but soil scientists cautioned that the water-holding capacity of a given soil is more likely a function of its inherent characteristics and is not manageable.[6] Other attributes, like previous yield, were dropped because of endogeneity concerns.

In designing a choice experiment, it is important to frame alternatives in a realistic way (Johnston et al., 2017). Soil health is measured in a variety of ways. Some popular soil science measures are bulk density (grams/cm^3), organic matter (% of soil sample), soil respiration (mg of CO_2/kg soil/time), and other technical measurements (Bünemann et al., 2018). However, simple non-technical measurements - like the number of earthworms in a clump of soil - can also be good indicators of soil health (Plaas et al., 2019). Walking the tightrope between meaningful technical measurements and measurements that have meaning to farmers led us to explore the concept of *linking indicators*, a term coined by Boyd et al. (2015). A linking indicator is one that indicates some important ecological measure, and also has meaning to the person interpreting it. For example, the hydraulic conductivity of soil (a measure of how fast water travels through soil) is sometimes referred to as K_{SAT}, and is measured in μm/seconds. If hydraulic conductivity is selected as an attribute in a choice experiment, should it be called "hydraulic conductivity", "K_{SAT}", "How fast water moves through the soil", or something else? Should it be measured in μm/seconds, or perhaps scaled to meters (or feet) per day?

The list of soil attributes were first identified using the focus groups and an initial review

[6]Personal communication with Dr. Dianna Bagnall and Dr. Cristine Morgan.

of the soil literature, and then refined after discussions with soil scientists and initial pre-tests of the survey with farmers. See Table 3.1 for the final list of attributes and levels. The four attributes used in the choice experiment are water infiltration, organic matter, compaction, and price. Water infiltration is a measure of how well water flows through the soil. The inherent characteristics of any soil (like sand and clay content) significantly influence water infiltration rates, but within a given soil class, variation in infiltration rates is likely due to soil management. We measure water infiltration as time (in hours) for an inch of standing water to absorb. Organic matter is the amount of organic material in a unit of soil, and we measure it in percentage terms. In pre-testing, farmers were familiar with measuring organic matter in percentage terms. Soil compaction is measure of how much resistance roots encounter when penetrating downwards. In certain conditions, pressure from heavy equipment can cause a compaction of the soil beneath the ground, which acts as a root barrier. The root barrier restricts the downward growth of the roots, causing plants to become less physically stable and produce less fruit as energy is re-routed to root exploration (Unger and Kaspar, 1994). The levels for all three soil attributes were set based on real in-field measurements taken at several field stations in the project area (Bagnall and Morgan, 2021).

The final attribute is the monetary vehicle. The choice occasion is the rental of a new field, so naturally the rental rate serves as the ideal monetary attribute.[7] Unfortunately, the average rental rates vary significantly over the study area. Figure 3.1 shows the average cash rental rate in USD from 2008 - 2020 for the State of Texas, with the Brazos River Watershed outlined in the first panel. There are significant differences across field types: irrigated land is generally more expensive than non-irrigated land, and pasture land is generally the cheapest of the three land types. Even within broad field types, there is substantial heterogeneity in typical cash rental rates across the study area. The large variance in land prices is problematic in determining the attribute levels. For example, a farmer operating on non-irrigated land near Austin (in the lower-middle region of the watershed) may be

[7] Some farmers in the focus group indicated their rental agreements are sometimes structured as crop-shares, but most were engaged in cash rentals.

Table 3.1: Discrete Choice Experiment Attributes and Levels

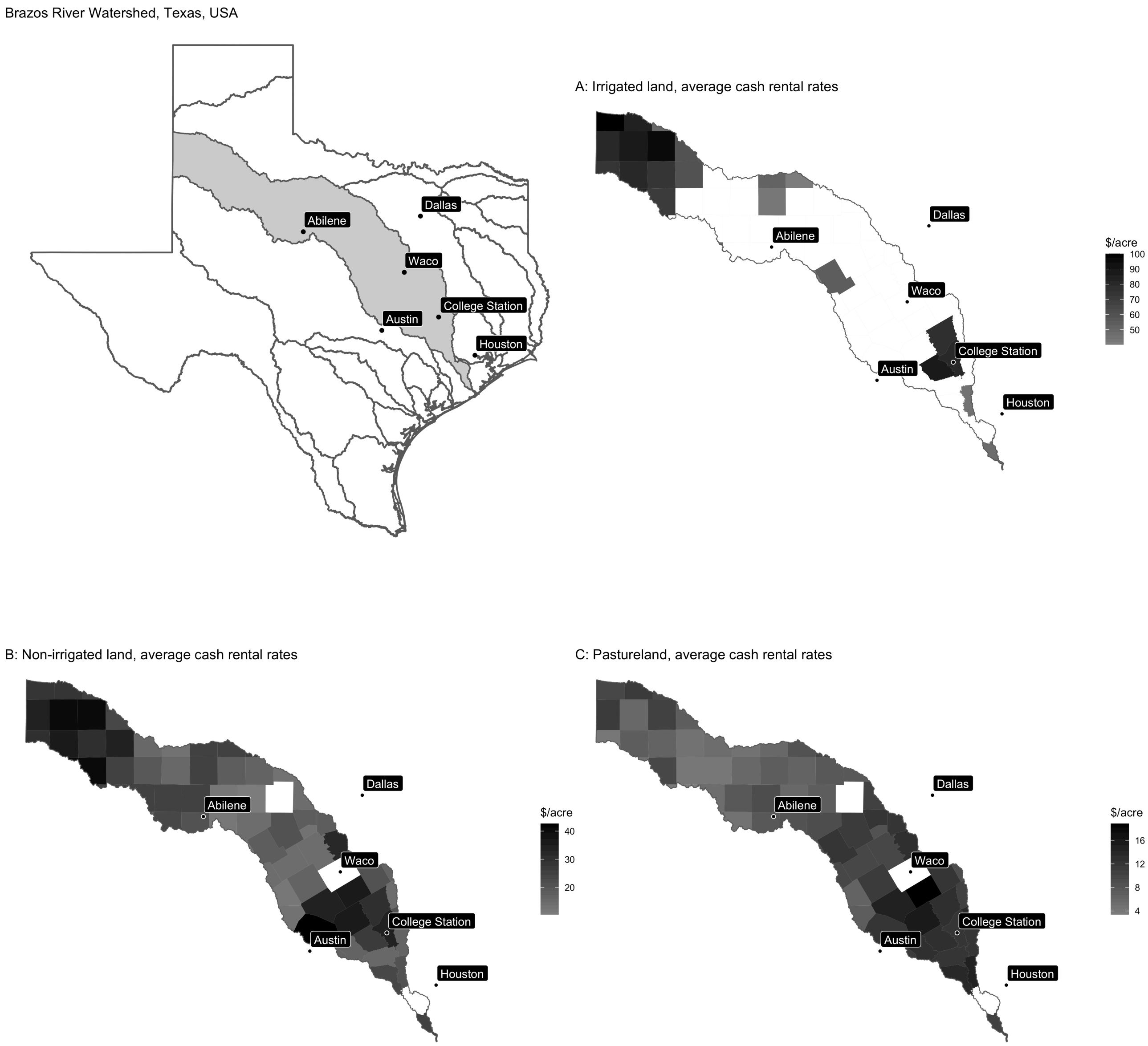
Attribute	Levels
Water infiltration	1 inch of standing water absorbed in 10 hours
	1 inch of standing water absorbed in 5 hours
	1 inch of standing water absorbed in 3 hours
Organic matter	0.5% by soil mass
	1% by soil mass
	2.5% by soil mass
Compaction	Does not restrict root growth
	Restricts root growth partially
	Restricts root growth substantially
Price	$10/acre less expensive that typical price
	Typical price
	$10/acre more expensive that typical price

Note: Soil attribute levels from Bagnall and Morgan (2021)

accustomed to rental rates of $40/acre. If we were to center the payment vehicle around $40 and present the survey to a farmer operating in an area where the average rental rates are closer to $10/acre, they may likely choose the opt-out alternative because the land is too expensive to rent.

To combat the large variance of the intended payment vehicle, we opt for a pivoted design, where the rental rates that each respondent is presented with are relative to their prior experience. Pivoted designs are relatively popular in the choice modeling literature (see Hensher and Rose (2007); Train and Wilson (2008); Hess and Rose (2009)). In a pivoted design, one attribute level is selected by the respondent based on some prior experience. The other attribute levels then *pivot* from the reference point, usually taking some symmetric deviation above and below the reference point. In our context, we ask the respondent to select his or her “base field.” The base field is some field that the farmer knows well. If a farmer were to choose any of his/her fields, there may be substantial selection bias where the base field is under no-till, but also happens to be in a favorable location with favorable inherent soil characteristics. To minimize the potential selection bias, we ask farmers to

Figure 3.1: Cash Rental Rates In Texas (Average, 2008 - 2020)



select roughly the same type of field, regardless of tillage practice. If the field is under no-till or strip-till, we ask the respondent to select a field where the tillage is marginally effective. If the field is under conventional tillage, we ask the respondent to select a field where no-till/strip-till could be advantageous. Once the base field is selected, we ask a series of several questions to learn about the base field's size, location, crop mix, tillage practices, ownership structure, topography, and perceptions on urban encroachment and soil health. Finally, we ask each respondent for the typical cash rental rate for their base field. The respondent's typical rental rate is the reference point, and we add two additional levels: \$10 above and \$10 below the reference point.

3.2.3 Experimental Design

To our knowledge, we are the first to estimate the value of manageable soil characteristics using a discrete-choice experimental approach. Without prior expectations for any of the preference parameters, we initially construct a traditional factorial design using the R package *support.CEs*(Aizaki, 2012).[8] We then deployed the choice experiment and used the preliminary results to estimate preference parameters. These initial parameters were then used as priors, and we re-designed the choice experiment using the Stata package *dcreate*, which maximizes the D-efficiency of the experimental design assuming a conditional logit choice model (Hole, 2015).

In the initial design phase, we faced a design challenge. We avoided the following conditions which make little sense in real fields:

Avoidance Condition 1. Field A has higher water infiltration, equal or lower organic matter, and equal or higher compaction than Field B, or vice versa.

Avoidance Condition 2. Field A has lower water infiltration, equal or more organic matter, and equal or lower compaction than Field B, or vice versa.

While there are established methods to avoid certain attribute level combinations for a given alternative, there is little consideration in the literature for alternatives that are prob-

[8]Specifically, we use a mix-and-match method described in Johnson et al. (2006): we find the orthogonal main-effects design (OMED), which is unique to our design of four attributes with three levels each. A new design is created by *rotating* the levels for each attribute up by one (and returning to the lowest for the highest attribute levels). The two designs represent two alternatives, with each row corresponding to a choice set. Next, each row is shuffled for both designs independently. Finally, a random draw is made from each design to construct the first choice set. The drawing continues with replacement until the number of drawable choice sets equals zero.

lematic only relative to otherwise unproblematic alternatives. There is no way to identify a problematic design until after the alternative combinations have been set. Our final design removes alternatives that *could* be problematic relative to others from the full-factorial combinations of all attributes and levels. For example, we removed alternatives that had the highest water infiltration rate, the lowest organic matter percentage, and the highest compaction level since this alternative would fail the first avoidance condition. We then iterate a full design until two conditions are met: 1) we avoid the identified problematic conditions and 2) we retain one choice occasion with a dominated alternative to check the monotonicity of preferences (Scarpa, Campbell, and Hutchinson, 2007).

The final DCE is embedded in a larger survey consisting of 12 sections that address the research questions of the interdisciplinary team of economists, sociologists, and soil scientists. The first section collects general demographic information, the second collects information specific to the farmer's base field, the third collects perception of soil quality and effectiveness of soil conservation practices, the fourth is the choice experiment, and the remaining 8 sections are sociological in nature and use the theory of planned behavior (TPB) to understand why farmers make the decision to adopt or not adopt a practice.

3.3 Survey Results

The final survey was administered through the USDA National Agricultural Statistics Service (NASS) to guarantee a representative and random sample in our study area. A total of 575 usable responses were collected in 2020 from a total of 2,833 mailed surveys, for an effective response rate of 20.3%. See Table 3.2 and Figure 3.2 for selected summary statistics. Note that because of NASS confidentiality concerns, we cannot present median values or any other individually identifying statistic. The average respondent is 68 years old and has been in operation for approximately 34 years. The average respondent rents more acres (681) than he owns (479), which is consistent with our initial focus group conversations about the prevalence of renting land in central Texas. The average base field operation is approximately 77 acres, and for those who have adopted no-till on the field, they have almost

9 years of experience with no-till on average. Approximately 15% of the average base field is prone to flooding. Figure 3.2 summarizes categorical responses across a set of key questions. Almost half of respondents plan on continuing to operate in five years. A vast majority of respondents did not use strip-till in 2018, and had no plans to change. The same is true for no-till, though a more sizable 15% of the sample used no-till in 2018. Most respondents did not use outside crop consultants, and of those who did, the vast majority relied on the consultant's advice "somewhat" or "very little". The selected base field for respondents is, on average, rented, under conventional-till, and is gently sloping. Most respondents consider their base-field to be neither floodplain nor hilly/upland.

Table 3.2: Selected Summary Statistics, Farmer Survey

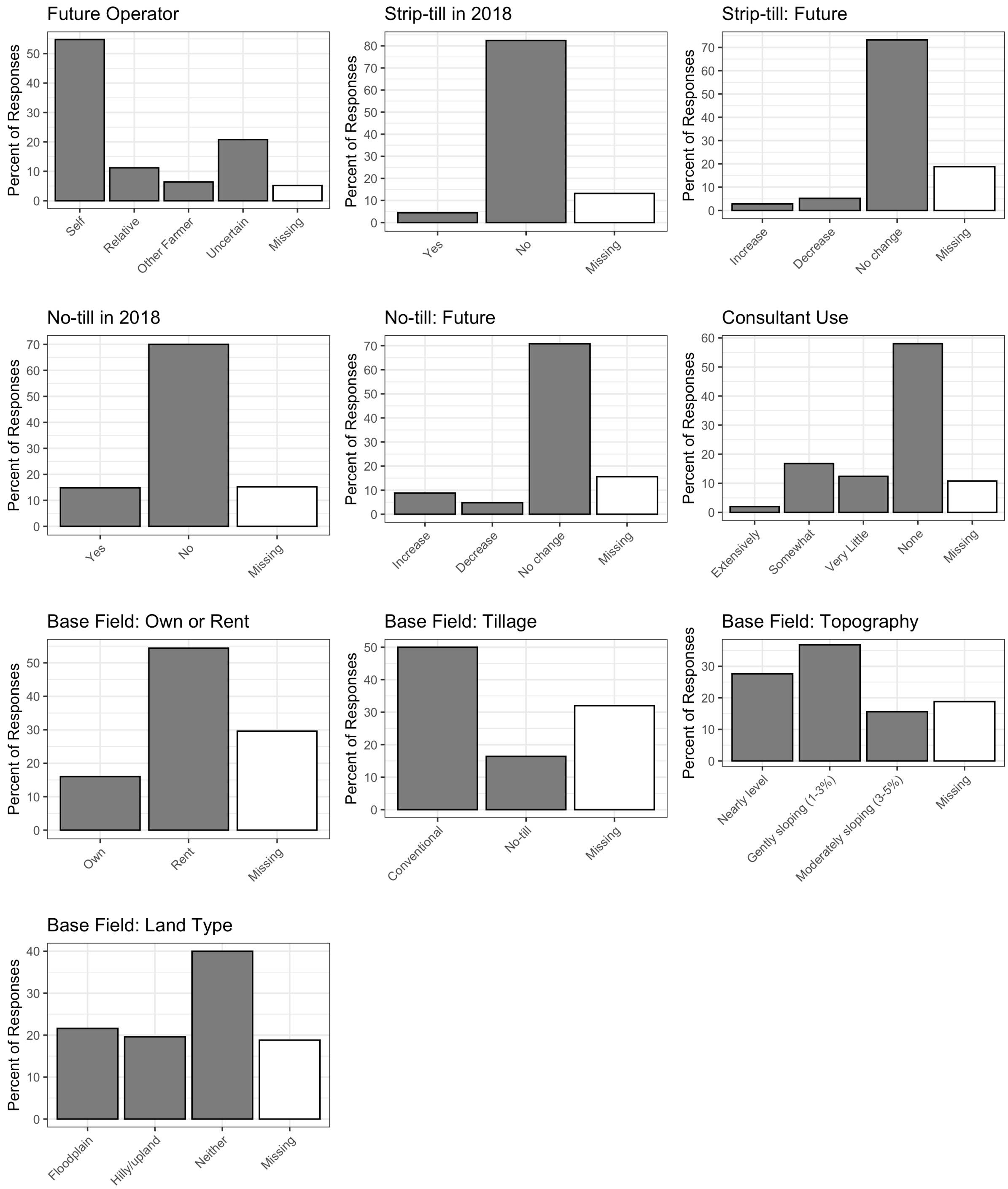
Survey Question	Mean	Std Deviation
Years in Operation	34.47	17.66
Age	68.30	12.02
Acres Owned	479.94	4,455.19
Acres Rented	681.25	2,700.69
Base Field Size (Acres)	77.03	174.08
Base Field Number of Years in No-till	8.91	35.77
Percent of Base Field Prone to Flooding	15.56	77.07

Note: Outliers cause large standard deviations. While outliers would normally be dropped, USDA NASS restrictions prevent the sharing of any data that could be used to identify a single operation. Outliers are therefore retained in our summary statistics.

3.4 Models

The statistical efficiency of the DCE design presupposes a conditional logit model of estimation (Hole, 2015). Our main modeling approach therefore characterizes respondent choice by using a random utility framework, estimated by a conditional logit (McFadden, 1974). The conditional logit setup ignores attribute preference heterogeneity across respondents. There is good reason to believe, however, that there may be substantial heterogeneity across

Figure 3.2: Selected Summary Statistics, Farmer Survey



respondent farmers. We try to capture preference heterogeneity in a variety of ways. First, we split the sample across base field rental rates, tillage practices, future operational plans, and base field topography. Second, we employ a latent class analysis to split the sample into distinct groups based on soil health attitudes. As we explain below, we identify four distinct groups of farmers based on these attitudinal questions. Third, we estimate several mixed logit regressions: one for every restricted sample in the conditional logit and latent class analyses.

In the next three subsections, we describe the model and estimation results for the standard conditional logit, latent class analysis, and mixed logit approaches.

3.4.1 Conditional Logit Analysis

Following the standard random utility model, we assume that respondent i chooses alternative j at choice occasion t when the expected utility from choosing alternative j is the maximum of all possible alternatives k in the choice set J, i.e.:

$$U_{ijt} \geq U_{ikt}, \quad \forall k \in J \tag{3.1}$$

where utility is comprised of two separable components: a deterministic component V_{ijt} and a structural error ε_{ijt} assumed to exhibit a Type I extreme value distribution, i.e. $U_{ijt} = V_{ijt} + \varepsilon_{ijt}$. In our case, $J = \{\text{Field A, Field B, Neither Field A nor Field B}\}$. We model the probability that respondent i chooses alternative j at choice occasion t as:

$$P_i(j,t) = \frac{e^{V_{ijt}}}{\sum_{k \in J} e^{V_{ikt}}}, \tag{3.2}$$

where V_{ijt} is the deterministic portion of utility, and takes the form:

$$V_{ijt} = \alpha_j + X_{jt}\beta + \gamma r_{jt}. \tag{3.3}$$

The α_j term is a set of two dummy variables that indicate if alternative j is Field A or Field

B. A negative α_j would suggest that respondents are less likely to choose one of the two main alternatives, and more likely to choose the opt-out alternative, *ceteris paribus*. The alternative-specific constants are not of particular interest in our study. The X_{jt} term is a matrix of values for the three selected soil attributes (water infiltration, organic matter, and compaction[9]) in the DCE. Finally, the r_{jt} term is the rental rate value for alternative j at occasion t. Recall that this is a pivoted variable, and thus takes one of three values: -1 if the rental rate is \$10 below the base rate, 0 if the rental rate is equal to the base rate, and 1 if the rental rate is \$10 above the base rate. The vector of estimated coefficients, β, and γ capture the marginal effect of a change in attribute levels on respondent utility.

Before estimating any model, we first filter out respondents that appear to not be making utility-maximizing decisions by eliminating respondents who choose a dominated alternative. We purposefully included a dominated alternative in choice occasion six to check for irrational choices. We also observe many respondents did not provide a base field rental rate, which is the center of the pivoted rental rate variable. While the pivoted design allows prior experience to enhance the realism in the choice experiment, the center pivot is not needed for estimation.[10] We retain respondents who did not provide the rental rate for their base field (approximately 19% of the sample) and assume the DCE was nonetheless realistic for this group.

See Table 3.3 for the conditional logit estimation results for each of the restricted samples described above. The base model which includes the entire sample is presented in column 1. All attributes significantly affect respondent utility, though there is a noticeable difference in magnitudes. We discuss relative magnitudes across attribute coefficients in the willingness-to-pay calculations below, where we scale the marginal change by a realistic change in soil management to more accurately compare differences in soil quality attribute preferences. The

[9]Compaction is a categorical variable that takes three values: high, medium, and low. We set the reference level to be "low" and omit this category from estimation

[10]The coefficient and variable for the rental rate in equation 3.3 can be represented as: $\gamma(r_i + r_{jt})$ where r_i is the rental rate of the respondent's base field, and r_{jt} is the pivoted value, which can be the same as the base rate, or \$10 above or below that rate. Note that r_i can be factored out and as in all conditional logits, drops out of the estimation process because it varies only over individuals.

coefficients in Table 3.3 are most helpful when interpreting coefficient signs or comparing across models. Again focusing on the base model, the coefficients on water infiltration and organic matter are positive, which means as water infiltration or organic matter increase for a field, respondent utility associated with choosing that field increases. On the other hand, the coefficients on high and medium compaction are negative, which means that a field with high or medium compaction yields lower respondent utility – compared to a field with low compaction. Finally, the coefficient on rental rate is negative, suggesting that as the rental rate of a field increases, the respondent utility associated with choosing that field decreases.

After estimating the base model, we split the sample in a variety of ways. First, we create three rental rate groups: a *low-price group* with base field rental rates of \$0 - \$20 per acre, a *medium-price group* with base field rental rates of \$21 - \$40 per acre, and a *high-price group* with a base field rental rate of more than \$40 per acre. Second, we create a *conventional tillage group* who use conventional tillage on their base fields, and a *no-till/strip-till group* who use either no-till or strip-till on their base fields.[11] Third, we create three groups based on the planned operator in five years: *self* if the respondent plans on continuing to operate himself, *other* if the respondent plans on having a relative or other person operating his fields, and *uncertain* if the respondent does not know who will be operating the farm in five years. Finally, we create two groups based on the topography of the base field: *level* if the base field is level (less than 1% slope), and *sloped* if the base field has any slope greater than 1%.

Columns 2 - 11 in Table 3.3 present the same conditional logit results for the restricted samples. Despite some noticeable differences in estimated coefficients across sample restrictions, all groups (sub-samples) yield somewhat similar results in coefficient magnitudes and identically-signed coefficients, all of which are significant at the 5% level. As we explore be-

[11] We tried creating three tillage groups: conventional, no-till, and strip-till, but the number of strip-till users was too small for subsequent analysis. We therefore absorbed the strip-till group into the no-till group, recognizing that there may be important differences in soil preferences between no-till and strip-till producers. Unfortunately, we lack the power to identify any differences between the unconventional tillage groups.

low, this leads to no significant differences in estimated willingness-to-pay across the groups. Note that all groups are not mutually exclusive; an individual in the medium rental rate group will also be in the level or sloped topography groups. However, within the group categories (rental rates, tillage practices, future plans, and topography), groups *are* mutually exclusive and a direct comparison of coefficient magnitudes can be made. For example, the high rental rate group is much less affected by changes in rental rates than the medium or low rental rate groups, as evidenced by the smaller coefficient on rental rate for the high rental rate group. Similarly, no-till/strip-till farmers appear to be much less sensitive to rental rates than conventional-till farmers in their choice to rent a field. At the same time, the no-till/strip-till farmers are more affected by all soil quality attributes than conventional-till farmers. Focusing on future plans, farmers who plan to continue to operate their own farms in five years are more affected by changes in water infiltration than farmers who are uncertain about their plans or are planning on selling their operation. Interestingly, there is little difference in the soil quality attribute coefficients for the two topography groups, though respondents with sloped base fields are more sensitive to rental rate changes than respondents on mostly level land.

Table 3.3: Conditional Logit Estimation Results

		Rental Rate Groups			Tillage Groups		Future Plan Groups			Topography Groups	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Base	Low Rental Rate	Medium Rental Rate	High Rental Rate	Conventional	No-till/Strip-till	Self	Other	Uncertain	Level	Slope
Rental Rate	-0.198^{***}	-0.259^{***}	-0.294^{***}	-0.134^{***}	-0.233^{***}	-0.121^{**}	-0.208^{***}	-0.240^{***}	-0.221^{***}	-0.128^{***}	-0.219^{***}
	(0.029)	(0.057)	(0.063)	(0.057)	(0.038)	0.061)	(0.039)	(0.063)	(0.072)	(0.053)	(0.036)
Water Infiltration	4.311^{***}	3.980^{***}	5.493^{***}	4.730^{***}	4.319^{***}	4.720^{***}	4.736^{***}	3.265^{***}	3.497^{***}	4.242^{***}	4.255^{***}
	(0.510)	(0.998)	(1.145)	(0.974)	(0.662)	(1.076)	(0.675)	(1.134)	(1.243)	(0.929)	(0.645)
Organic Matter	0.401^{***}	0.421^{***}	0.392^{***}	0.343^{***}	0.324^{***}	0.603^{***}	0.400^{***}	0.337^{***}	0.390^{***}	0.429^{***}	0.388^{***}
	(0.042)	(0.080)	(0.088)	(0.080)	(0.053)	(0.091)	(0.054)	(0.089)	(0.104)	(0.077)	(0.052)
High Compaction[a]	-1.745^{***}	-1.862^{***}	-1.613^{***}	-1.489^{***}	-1.637^{***}	-1.849^{***}	-1.566^{***}	-1.887^{***}	-2.070^{***}	-1.713^{***}	-1.744^{***}
	(0.109)	(0.217)	(0.219)	(0.206)	(0.134)	(0.257)	(0.142)	(0.230)	(0.272)	(0.204)	(0.135)
Medium Compaction[a]	-0.566^{***}	-0.483^{***}	-0.580^{***}	-0.437^{***}	-0.466^{***}	-0.647^{***}	-0.6105^{***}	-0.458^{***}	-0.573^{***}	-0.608^{***}	-0.512^{***}
	(0.078)	(0.149)	(0.168)	(0.152)	(0.099)	(0.172)	(0.104)	(0.165)	(0.186)	(0.146)	(0.097)
Field A Intercept	0.186^{*}	0.088	0.633^{***}	-0.178	0.374^{***}	-0.633^{***}	-0.057	0.783^{***}	0.570^{**}	-0.124	0.305^{**}
	(0.127)	(0.259)	(0.273)	(0.240)	(0.163)	(0.283)	(0.169)	(0.286)	(0.314)	(0.234)	(0.161)
Field B Intercept	0.073	0.305^{*}	0.452^{**}	-0.444^{**}	0.161	-0.328^{*}	-0.135	0.696^{***}	0.307	-0.219	0.251^{**}
	(0.111)	(0.216)	(0.239)	(0.215)	(0.143)	(0.237)	(0.148)	(0.242)	(0.267)	(0.205)	(0.139)
Observations	7,815	1,986	1,716	2,133	4,776	1,635	4,275	1,701	1,431	2,283	4,932
Log-Likelihood	−2,311.282	−5611.013	−471.865	−635.231	−1,399.470	−504.524	−1,310.564	−486.749	−401.685	−689.568	−1,453.835
AIC	4,636.564	1,236.026	957.730	1,284.462	2,812.940	1,023.047	2,635.127	987.498	817.370	1,393.137	2,921.67

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

[a] *Omitted base category: low compaction*

3.4.2 Latent Class Analysis

As seen in Table 3.3, restricting the sample of respondents based on a single observable variable (like base field rental rate or tillage practice) results in relatively minor differences in coefficient estimates across groups. Splitting the sample does not remove potentially confounding opinion or behavioral patterns that can persist within restricted samples. For example, the *medium-price group* may have distinctly different groups of respondents, and estimating a single coefficient on various preference parameters hides potentially important heterogeneity.

Latent class analysis uses a set of categorical responses (referred to as *manifest* variables) to group respondents into distinct groups, or classes. The assigned class is referred to as the latent class, and is assumed to remove potential confounding between the set of manifest variables. Latent class models have been used, for example, to group agricultural producers based on their production decisions, the size of their operations, and their risk preferences (Chinedu et al., 2018). Latent class analysis (LCA) allows researchers to identify distinct groups of respondents that exhibit similar observed behavior, and an unobserved grouping variable is assumed to cause the observed patterns.

We hypothesize that farmers may be categorized into distinct groups based on their opinions on soil health and conservation practices. We in turn assume that the latent grouping variable causes differences in preferences for the discrete choice experiment and estimate a conditional logit for each latent class. Instead of a sample-wide fixed coefficient for each preference parameter, the LCA allows for a fixed coefficient for each class, which allows heterogeneous preferences across classes.

In setting up our latent class analysis, we use best practices outlined in Collins and Lanza (2009). The section preceding the DCE in the survey gathers information on the respondents' attitudes towards soil health. See Table 3.4 for a list of the fourteen questions and possible responses. For seven soil characteristics (water infiltration, organic matter, runoff, erosion, bulk density, compaction, and drainage), we as respondents 1) how important are changes in these characteristics to their base field, and 2) how no-till/strip-till would affect these characteristics on their base field.

Table 3.4: Latent Class Analysis Manifest Variables

Importance of a change in soil health characteristics to your base field:						
	Very Important	Fairly Important	Important	Slightly Important	Not Important	Don't Know
Increasing water infiltration	○	○	○	○	○	○
Increasing organic matter	○	○	○	○	○	○
Decreasing runoff	○	○	○	○	○	○
Decreasing erosion	○	○	○	○	○	○
Decreasing bulk density	○	○	○	○	○	○
Decreasing compaction	○	○	○	○	○	○
Increasing drainage	○	○	○	○	○	○
No-till or strip-till would increase or decrease soil characteristics on your base field:						
	Greatly Increase	Increase	Neither	Decrease	Greatly Decrease	Don't Know
Water infiltration	○	○	○	○	○	○
Organic matter	○	○	○	○	○	○
Runoff	○	○	○	○	○	○
Erosion	○	○	○	○	○	○
Bulk density	○	○	○	○	○	○
Compaction	○	○	○	○	○	○
Drainage	○	○	○	○	○	○

We observe $J = 14$ manifest variables, each of which has $K = 6$ number of possible outcomes for each individual $i = 1, ..., N$. Let $Y_{ijk} = 1$ if respondent i responds with k for manifest variable j, and $Y_{ijk} = 0$ otherwise. Finally, let R be the number of anticipated latent classes, set prior to estimation. We return to the search for the optimal size of R below. Let π_{jrk} be the probability that an individual in class r responds with option k for manifest variable j. Finally, let p_r be the unconditional probability of any given individual belonging to class r. The probability density function of a vector of responses (Y_i) for individual i conditional on π and p is

$$Pr(Y_i|\pi, p) = \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K} (\pi_{jrk})^{Y_{ijk}}, \tag{3.4}$$

and estimation proceeds by iteratively estimating $\hat{\pi}_{jrk}$ and $\hat{p}_r$, and replacing the initial expectations in the log-likelihood function version of equation 3.4 until the difference in log-likelihood between iterations becomes arbitrarily small. We estimate the above probabilities using the R package *poLCA* which identifies equation 3.4 using an expectation-maximization algorithm (Linzer, Lewis et al., 2011; Bandeen-Roche et al., 1997).

The final step in searching for the optimal latent class analysis is to identify the number of classes R. We estimate the above latent class model for classes of size 1 to 5, and examine the model fit across the versions. We search for the log-likelihood closest to zero, and the model that minimizes the Akaike information criterion (AIC) and Bayesian information criterion (BIC) (Akaike, 1998; Schwarz et al., 1978). An LCA with class size $R = 4$ yields the overall best fit across the AIC and BIC statistics.

Figure 3.3 shows the class-conditional response probabilities from the LCA. There are clear patterns that allow us to name the generic latent classes for subsequent analysis. We focus first on the "importance" responses: how important are the soil health characteristics to respondents' base fields? Class 1 has varied predicted response probabilities, but members of this class are most likely to think the seven soil health characteristics are slightly or not at all important to their base field. Member of class 2 are overwhelmingly likely to think the characteristics are very important to their base field. The predicted response probabilities for class 3 are the most balanced, with no dominant response for any of the "importance" questions. Finally, members of class 4 are most likely to not know how important the soil characteristics are for the health of their base field.

Figure 3.3 also shows the class-conditional response probabilities for the "effect" responses: how do respondents think no-till/strip-till would affect the soil health characteristics? Members of class 1 are most likely to think no-till/strip-till would not change any of the characteristics. Members of class 2 are most likely to think no-till would *greatly* increase all characteristics. Members of class 3 are again balanced with roughly neutral opinions, except for water infiltration and organic matter, which members agree would increase under a no-till regime. Members of class 4 are almost certain to not know how no-till would affect the soil health characteristics.

Based on the predicted class-conditional response probabilities, we assign the following labels to the four generic classes. Class 1 is the *soil-apathetic* class, since members are likely to think the seven soil characteristics are not important and likely to think no-till would have no effect on the characteristics. Class 2 is the *soil-conscious* class, since members are likely to think the characteristics are very important and no-till would greatly increase them. Class 3 is the *moderate* class, for balanced and tepid predicted response probabilities. Finally, class 4 is the *uninformed* class, since members are likely to not know either the importance of the characteristics nor the effect no-till would have.

After forming the latent classes and assigning individuals to their respective classes, we estimate the standard conditional logit from the previous section for each latent class separately. Results are shown in Table 3.5. Across all four latent classes, the estimated coefficients are similar in sign and significance to each other and the general results from the conditional logit models of the last section. There are, however, some noticeable differences in magnitude across the latent classes. For example, the uninformed class has a much smaller coefficient on water infiltration than the other classes, and the coefficient on water infiltration is largest for the soil-conscious class. Similar to the conventional-till group from the previous section, the moderate class is most sensitive to rental rates. Despite these differences in coefficient magnitudes, the four classes yield estimates that are generally similar. This suggests – and we confirm below – that attitudes on soil health are not responsible for significant differences in willingness-to-pay for soil quality improvements.

3.4.3 Mixed Logit Analysis

While splitting the sample for the conditional logit model by observable or latent characteristics may account for some important respondent heterogeneity, an alternative approach is to use a mixed logit model. In our mixed logit approach, we assume the indirect utility of respondent i selecting alternative j at choice occasion t to be:

$$V_{ijt} = \alpha_j + X'_{jt}\beta_i + \gamma W_{jt}, \tag{3.5}$$

Figure 3.3: Class-Conditional Response Probabilities, Latent Class Analysis

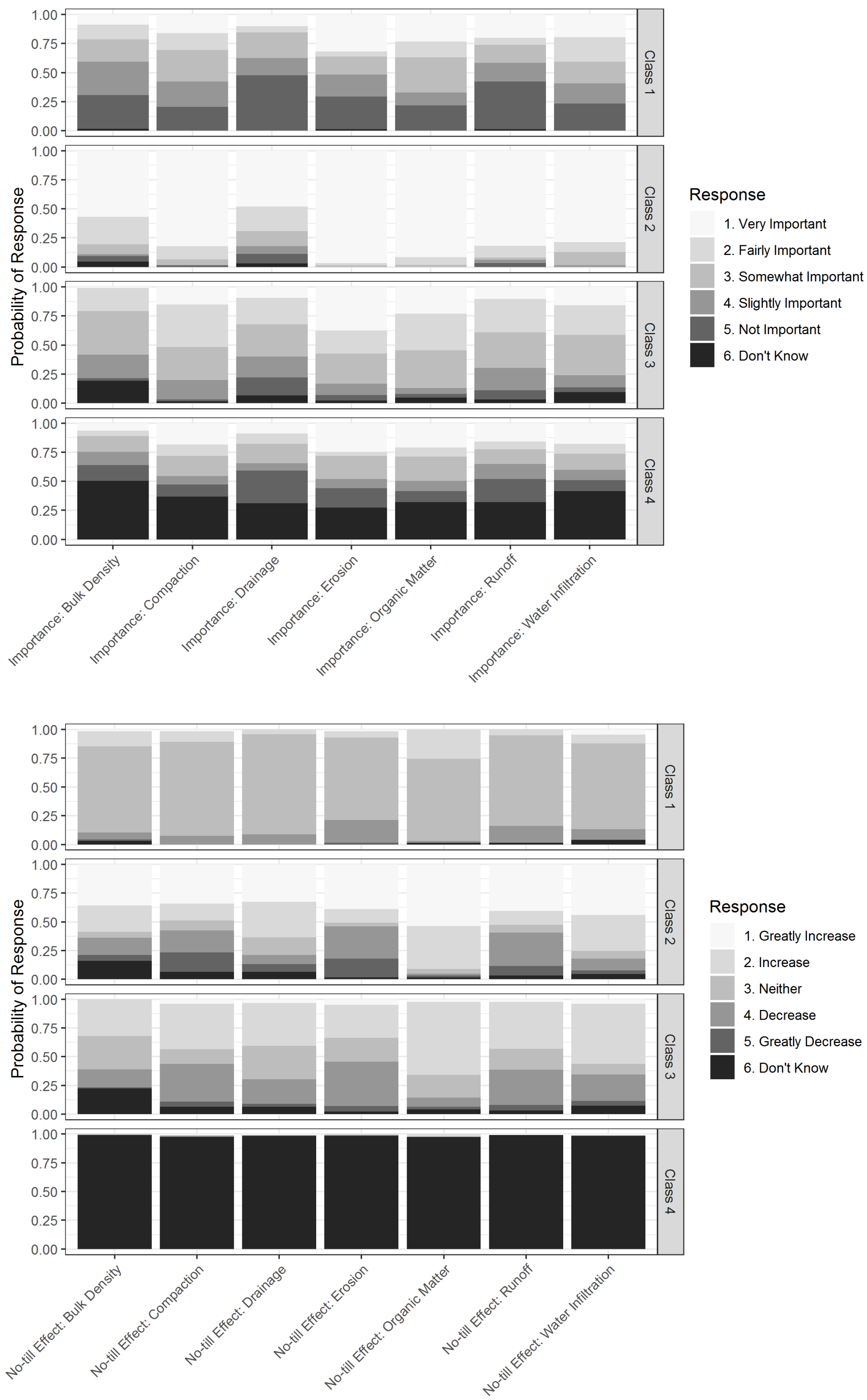


Table 3.5: Conditional Logit Estimation Results By Latent Class

	(1) Soil-Apathetic Class	(2) Soil-Conscious Class	(3) Moderate Class	(4) Uninformed Class
Rental Rate	-0.201^{***}	-0.139^{**}	-0.247^{***}	-0.188^{***}
	(0.082)	(0.073)	(0.053)	(0.066)
Water Infiltration	6.427^{***}	5.799^{***}	5.168^{***}	2.361^{**}
	(1.382)	(1.311)	(0.930)	(1.161)
Organic Matter	0.323^{***}	0.527^{***}	0.369^{***}	0.455^{***}
	(0.113)	(0.113)	(0.075)	(0.092)
High Compaction[a]	-1.610^{***}	-1.963^{***}	-1.710^{***}	-1.554^{***}
	(0.298)	(0.288)	(0.191)	(0.248)
Medium Compaction[a]	-0.639^{***}	-0.787^{***}	-0.508^{***}	-0.402^{***}
	(0.212)	(0.209)	(0.140)	(0.178)
Field A Intercept	-0.560^{**}	0.045	0.258	0.140
	(0.347)	(0.315)	(0.227)	(0.300)
Field B Intercept	-0.580^{**}	−0.128	0.138	0.166
	(0.301)	(0.276)	(0.197)	(0.259)
Observations	1,062	1,296	2,460	1,398
Log-Likelihood	−319.992	−346.346	−703.812	−448.784
AIC	653.985	706.691	1,421.624	911.568

Note: *p<0.1; **p<0.05; ***p<0.01

[a] *Omitted base category: low compaction*

which differs from equation 3.3 by allowing the vector of preference parameters on the soil quality characteristics, β, to vary over individuals. The mixed logit is an extension of the conditional logit (McFadden and Train, 2000), with the full form of utility still comprised of two parts: the deterministic indirect utility of equation 3.5 and a structural error that is still assumed to be distributed i.i.d. type I extreme value. The mixed logit is attractive because it allows for respondent preference heterogeneity and relaxes the restrictive IIA assumption required for the conditional logit (McFadden and Train, 2000). However, because β_i is a random variable, it takes some distribution, i.e. $f(\beta_i|\theta)$, and we make an assumption about the shape of that distribution. Specifically, we recover the parameters in equation 3.5 assuming a normal distribution of β_i for each soil quality characteristic and therefore recover the mean $\bar{\beta}$ and standard deviation σ of $f(\beta_i|\theta)$. We then explore the differences in estimated coefficients.

Recent advances in the mixed logit literature suggest that the typical distributional assumptions of $f(\beta_i|\theta)$ can result in unreliable WTP estimates, and more flexible semi-parametric distributional approaches (polynomials, splines, step-functions) can improve the reliability of WTP estimates (Bazzani, Palma, and Nayga Jr, 2018; Train, 2016; Scarpa, Franceschinis, and Thiene, 2021). Our empirical estimates for MWTP for improvements in manageable soil quality characteristics are, however, novel. We therefore do not have strong prior expectations that heterogeneous preferences exist, much less the shape of the distribution of the preferences. Instead of dialing-in on precise estimates in a highly uncertain environment, our focus in this work is to use three common approaches (conditional logit, mixed logit, and latent class analysis) and identify broad differences in MWTP estimates to better improve future DCE efforts for valuation of manageable soil quality.

Results for the mixed logit models using the sample restrictions above are shown in Table 3.6. Results are similar to the standard conditional logit specification, except for the additional standard deviation parameter. For water infiltration and high compaction, the standard deviation coefficient is significant, suggesting the presence of preference heterogeneity even within the restricted sample. There is also heterogeneity in preferences for organic matter, though not for the low rental rate group, the conventional till group, and the uncertain future plan groups. There is no evidence of preference heterogeneity for medium compaction across any of the models.

Results for the mixed logit models using the estimated latent classes are shown in Table 3.7. Results are again similar, with significant preference heterogeneity in water infiltration and high compaction across the four classes, and less heterogeneity in preferences for water infiltration for the soil-apathetic and soil-conscious classes. There is again no significant preference heterogeneity for medium compaction.

Table 3.6: Mixed Logit Estimation Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		Rental Rate Groups			Tillage Groups		Future Plan Groups			Topography Groups	
	Base	Low Rental Rate	Medium Rental Rate	High Rental Rate	Conventional	No-till/Strip-till	Self	Other	Uncertain	Level	Slope
Means											
Rental Rate	-0.368***	-0.465***	-0.477***	-0.292***	-0.436***	-0.281***	-0.398***	-0.41***	-0.411***	-0.259***	-0.412***
	(0.041)	(0.082)	(0.086)	(0.079)	(0.054)	(0.083)	(0.054)	(0.086)	(0.111)	(0.079)	(0.05)
Water Infiltration	9.686***	9.228***	10.144***	10.758***	9.847***	8.281***	9.806***	6.626***	12.839***	10.1***	9.631***
	(1.008)	(1.895)	(2.078)	(1.831)	(1.366)	(1.856)	(1.356)	(1.919)	(3.518)	(2.023)	(1.257)
Organic Matter	0.687***	0.714***	0.593***	0.614***	0.584***	0.933***	0.713***	0.564***	0.596***	0.79***	0.675***
	(0.068)	(0.128)	(0.135)	(0.142)	(0.08)	(0.161)	(0.092)	(0.142)	(0.179)	(0.144)	(0.084)
High Compaction	-2.853***	-3.025***	-2.603***	-2.601***	-2.749***	-3.183***	-2.632***	-2.962***	-3.706***	-2.563***	-3.037***
	(0.226)	(0.474)	(0.451)	(0.448)	(0.296)	(0.501)	(0.278)	(0.507)	(0.893)	(0.412)	(0.302)
Medium Compaction	-1.038***	-0.912***	-0.956***	-0.863***	-0.966***	-1.124***	-1.095***	-0.803***	-1.252***	-1.081***	-1.029***
	(0.111)	(0.21)	(0.224)	(0.214)	(0.139)	(0.239)	(0.146)	(0.222)	(0.314)	(0.228)	(0.135)
Field A Intercept	0.078	-0.145	0.729**	-0.424	0.393*	-0.814**	-0.13	0.829**	0.335	-0.59*	0.348*
	(0.158)	(0.323)	(0.33)	(0.301)	(0.202)	(0.338)	(0.208)	(0.354)	(0.404)	(0.304)	(0.196)
Field B Intercept	0.213	0.444*	0.635**	-0.391	0.381**	-0.243	0.081	0.904***	0.323	-0.509*	0.535***
	(0.138)	(0.266)	(0.292)	(0.27)	(0.177)	(0.288)	(0.184)	(0.301)	(0.352)	(0.278)	(0.17)
Standard Deviations											
Water Infiltration	8.693***	8.124***	7.79***	8.767***	9.697***	8.952***	9.61***	7.062***	10.698***	10.02***	8.692***
	(0.727)	(1.163)	(1.353)	(1.283)	(0.986)	(1.87)	(1.057)	(1.23)	(2.145)	(1.58)	(0.812)
Organic Matter	0.38***	0.283	0.35**	0.544***	0.185	0.504***	0.434***	0.364**	0.08	0.489***	0.384***
	(0.109)	(0.213)	(0.151)	(0.15)	(0.141)	(0.186)	(0.13)	(0.157)	(0.739)	(0.189)	(0.123)
High Compaction	1.789***	2.023***	1.495***	1.953***	1.785***	1.55***	1.71***	1.621***	1.849***	1.77***	1.867***
	(0.209)	(0.395)	(0.34)	(0.506)	(0.255)	(0.438)	(0.226)	(0.393)	(0.6)	(0.413)	(0.238)
Medium Compaction	0.117	0.145	0.233	0.041	0.009	0.309	0.069	0.069	0.514	0.46	0.104
	(0.137)	(0.191)	(0.304)	(0.298)	(0.186)	(0.336)	(0.168)	(0.251)	(0.406)	(0.346)	(0.141)
Observations	2072	530	466	580	1258	474	1181	428	354	577	1367
Log-Likelihood	-1603.690	-425.047	-354.779	-436.740	-966.634	-379.356	-934.794	-337.904	-242.133	-447.958	-1057.480
AIC	3229.379	872.095	731.557	895.479	1955.268	780.711	1891.588	697.807	506.266	917.916	2136.960

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.7: Mixed Logit Estimation Results: Latent Class Analysis

	(1) Soil-Apathetic Class	(2) Soil-Conscious Class	(3) Moderate Class	(4) Uninformed Class
Means				
Rental Rate	-0.374***	-0.308***	-0.398***	-0.377***
	(0.101)	(0.093)	(0.065)	(0.086)
Water Infiltration	11.322***	11.665***	10.018***	6.360***
	(2.371)	(2.442)	(1.557)	(1.938)
Organic Matter	0.553***	0.848***	0.563***	0.745***
	(0.150)	(0.172)	(0.106)	(0.160)
High Compaction	-2.581***	-3.284***	-2.882***	-2.831***
	(0.488)	(0.526)	(0.366)	(0.496)
Medium Compaction	-1.166***	-1.383***	-0.894***	-0.900***
	(0.269)	(0.267)	(0.170)	(0.233)
Field A Intercept	-0.572	-0.048	0.278	0.263
	(0.376)	(0.355)	(0.251)	(0.340)
Field B Intercept	-0.339	0.018	0.313	0.512*
	(0.327)	(0.317)	(0.219)	(0.299)
Standard Deviations				
Water Infiltration	9.714***	9.242***	7.756***	9.110***
	(1.665)	(1.61)	(1.031)	(1.438)
Organic Matter	0.137	0.476*	0.468***	0.627***
	(0.162)	(0.247)	(0.145)	(0.178)
High Compaction	1.385***	1.727***	1.680***	1.774***
	(0.439)	(0.528)	(0.276)	(0.390)
Medium Compaction	0.090	0.049	0.047	0.044
	(0.261)	(0.271)	(0.206)	(0.304)
Observations	354	432	820	466
Log-Likelihood	-275.696	-303.367	-631.037	-374.968
AIC	573.393	628.734	1284.074	771.936

Note: *p<0.1; **p<0.05; ***p<0.01

3.5 Willingness-To-Pay Simulations

After estimating preference parameters for various subsamples and latent classes, we calculate marginal willingness to pay (MWTP) using a standard approach (Haab and McConnell, 2002):

$$WTP = -\frac{\beta}{\gamma} \cdot \Delta \cdot \rho, \tag{3.6}$$

where β is the coefficient on water infiltration, organic matter, or compaction (or a distribution defined by the mean and standard deviation coefficients on the same parameters), and γ is the coefficient on the rental rate. The ratio of these coefficients is the mean willingness-to-pay (MWTP) for a unit change in a specific attribute. For non-unit changes, we can multiply the ratio by any value of Δ, which is the magnitude of the change. The monetary vehicle in DCEs is typically continuous and thus varies by a natural unit: the dollar. In our case, rental rate varies by increments of \$10, and we must therefore multiply the ratio and Δ by $\rho = 10$ to translate MWTP to normal dollar terms.[12]

Setting $\Delta = 1$ in our context is not necessarily realistic nor helpful, since a full unit change might not be possible in the study area. Instead, we base our MWTP estimates on realistic soil quality changes attained in our study area. A team of soil scientists conducted a longitudinal study of soil quality measures in the watershed of the Brazos River in central Texas, taking repeated measurements of soil quality across farms under different management regimes (Bagnall and Morgan, 2021). See Table 3.8 for median soil quality measures under three different tillage regimes: conventional-till, no-till, and a perennial grass system. Hydraulic conductivity is equivalent to what we present as water infiltration in the DCE. Bulk density is equivalent to what we categorize as compaction.[13] While there is no real difference in bulk density across the regimes, there is a clear increase in water infiltration and organic matter moving from conventional-till to no-till in our study area. It is also important to note that despite the small difference in bulk density across the regimes, it is still possible to achieve significant improvements in compaction levels for farmers in our area.

Table 3.8: Median Soil Quality Indicators For Study Area

Management Regime	Hydraulic Conductivity (cm/hr)	Bulk Density (g/cm^3)	Organic Matter (% by mass)
Conventional	1.33	1.38	2.06
No-till	2.01	1.42	2.29
Perennial	3.04	1.40	4.40

Source: Bagnall and Morgan (2021)

[12]Specifically, $MWTP = -\frac{\beta}{\gamma}$, but since the rental rate changes in tens of dollars, the denominator must be divided by 10 to convert to single dollars. Thus, $MWTP = -\frac{\beta}{\frac{\gamma}{10}} = -\frac{\beta}{\gamma} \cdot 10$

[13]By "equivalent", we mean it measures the same soil quality attribute, even if the units or labels are different.

Based on the in-field measurements, we calculate the willingness-to-pay for a Δ change in soil quality based on the adoption of no-till. That is, we estimate the MWTP for an average conventional-till farm to achieve the average soil quality improvements associated with a typical improvement for farms in the area. We set $\Delta = 0.27$inches/hour for water infiltration.[14] We set $\Delta = 0.04\%$ for organic matter. We leave the compaction indicators equal to one.

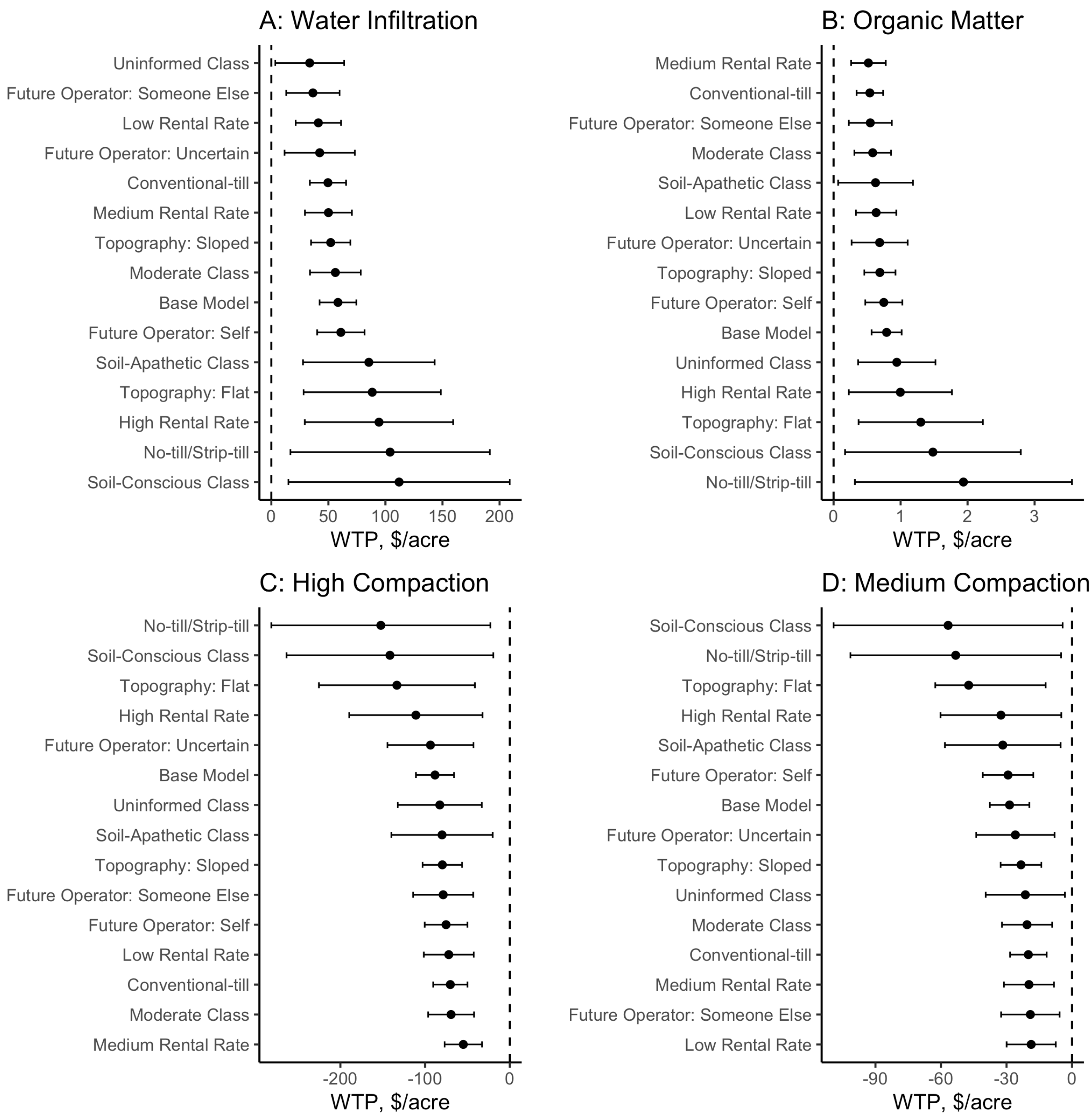
For both the conditional logit and mixed logit models, we calculate the 95% confidence intervals using the delta method. The delta method does not require any simulation for the conditional logit models, but because the mixed logit coefficients define a *distribution*, with the key distributional parameters also subject to sampling variance, the delta method requires both a closed form solution and simulation (Bliemer and Rose, 2013). Specifically, we take 1,000 random draws of the mean and standard error of the distribution of WTP over individuals, and use the average of the drawn means and standard errors of the distributions to define the asymptotic full distribution of WTP. As Bliemer and Rose (2013) note, the nature of variation over the WTP distribution *and* sampling variance of the key distributional parameters typically result in wide confidence intervals, which is what we find.

Results of the MWTP estimates for the conditional logit can be seen in Figure 3.4. Panel A shows MWTP for a realistic improvement in water infiltration by switching from conventional to no-till in our study area. Individuals in the soil-conscious latent class and the no-till/strip-till group are willing to pay the most, about \$100/acre. Respondents operating on flat topography are willing to pay more for water infiltration than respondents on sloped terrain, though the difference is not statistically significant. Indeed, no individual subsample or latent class has a significantly different MWTP than any other group. The same is true for the other soil quality characteristics (Panels B - D). In general, the conditional logit models suggest that farmers in the watershed of the Brazos River in Texas are willing to pay approximately \$50 - \$100/acre to improve water infiltration and approximately \$1 - \$2/acre

[14] We convert the raw difference of 0.68cm/hour to an inch/hour measure by multiplying the measure by 0.393701.

to improve organic matter by adopting no-till. Farmers are willing-to-pay approximately \$100/acre to move from high to low compaction, and approximately \$40/acre to move from medium to low compaction.

Figure 3.4: Marginal Willingness-To-Pay By Model Specification: Conditional Logit



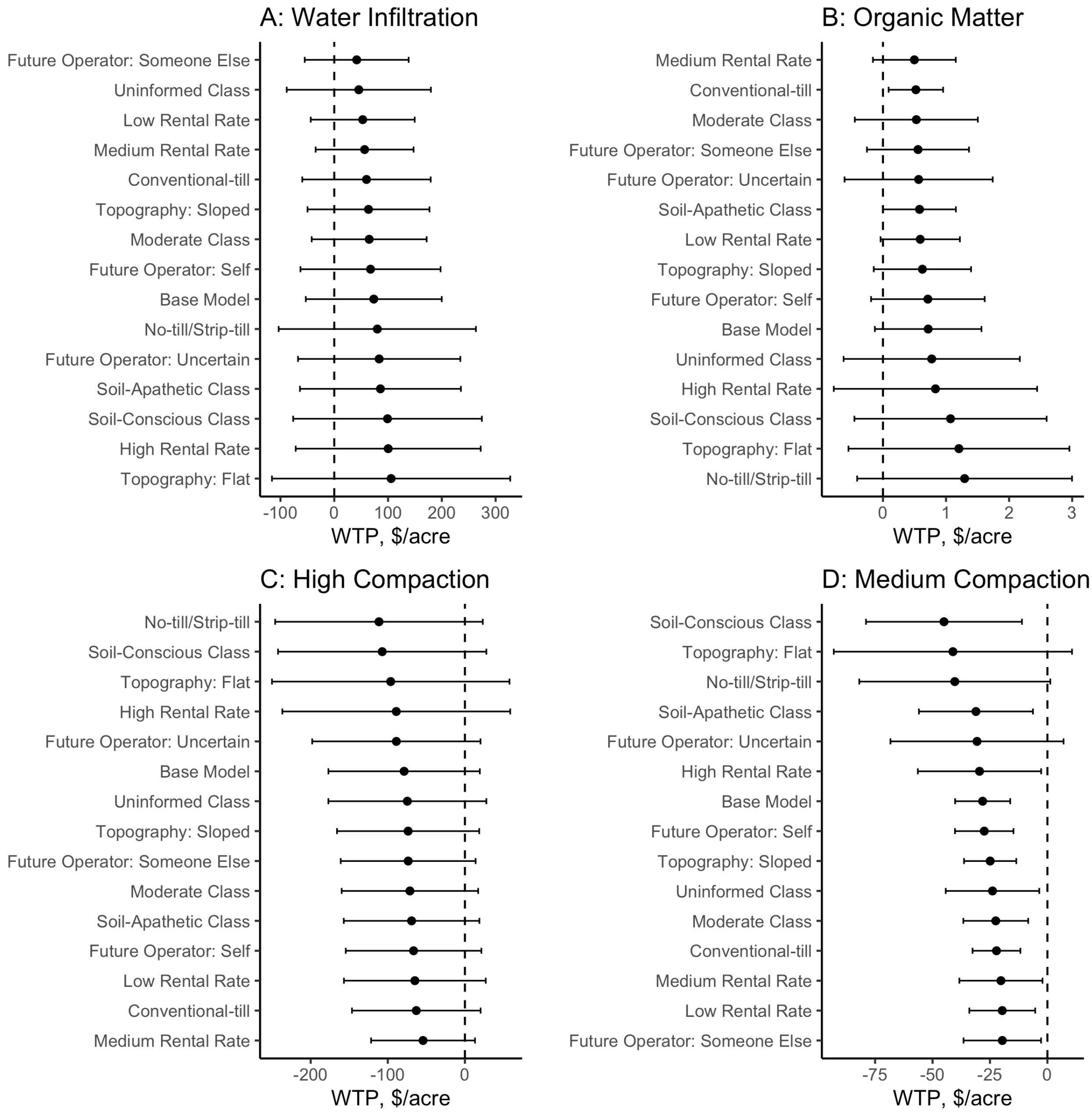
Results of the MWTP estimates for the mixed logit models can be seen in Figure 3.5. Perhaps unsurprisingly, the MWTP estimates are similar to the conditional logit estimates, but with much wider confidence intervals. While all MWTP estimates were significant at the 5% level, few of the MWTP estimates are significant at the same level for the mixed logit models, which is common for mixed logit models Bliemer and Rose (2013). The wide confidence intervals for the mixed logit models are a result of random preference parameters that vary over individuals and the key parameters that define that variation also exhibit natural sampling variance. In the context of the mixed logits, we find no strong evidence that farmers are willing to pay for improvements in soil quality associated with adopting no-till. The full numerical MWTP results for the conditional logit and mixed logit models is presented in Tables B.1 and B.2, respectively.

3.6 Policy Implications

The on-farm value of manageable soil quality characteristics has long been estimated using changes in crop revenues or input costs. This study is the first to directly measure the value of changes in manageable soil quality characteristics independent of changes in agricultural production. Our willingness-to-pay estimates can therefore be considered the perceived benefits of soil quality improvements to farmers in our study area. Our results have important policy implications.

Our results can be compared to realized benefits to identify any significant difference between perceived and realized benefits of improved soil quality. For example, recent work from the Soil Health Institute has focused on partial budget analysis, which estimates the explicit costs and benefits of soil health management. An average soil health management system resulted in some reduced expenses and additional revenue for a total benefit of $93.66 per acre for a typical corn operation in Iowa (Soil Health Institute, 2021). They also estimate a total change in cost of $29.81 per acre. A direct comparison to our WTP results is not appropriate because of significant differences in inherent soil and agricultural characteristics between Texas and Iowa. Nonetheless, estimates from partial budget analyses and studies like

Figure 3.5: Marginal Willingness-To-Pay By Model Specification: Mixed Logit



ours can help policy makers identify potential gaps between potential and realized benefits.

For example, suppose a farmer in the Texas Blackland MLRA is considering switching from a conventional to no-till regime. Suppose further that the farmer has a medium compacted soil, and he expects no-till to result in average improvements in water infiltration, organic matter, and a move from medium to low compaction. According to our work, that farmer (on average) is willing to pay approximately \$87 (\$58 + \$1 + \$28 from the conditional logit, full sample results) per acre to improve his manageable soil quality. Suppose partial budget analysis indicates that the benefits per acre of this change in soil health characteristics is \$150 per acre per year. This would be nearly twice the the stated preference estimate of \$86 and it is only for a single year – soil health benefits in our study should reflect the value of those characteristics for up to five years. In other words, this would suggest that the perceived benefits are much smaller than the best estimates of the true financial benefit. Understanding the difference between these two estimates can help policy makers understand the need for better communication on the benefits of conservation practices and how such practices are marketed, funded, and targeted.

Even without a comparison to partial budget analysis, our results are an estimate of the total value of improvements in soil quality and are therefore important on their own. Partial budget analysis misses non-use values that could be important to farmers and are captured in choice experiments like ours. With or without partial budget analysis, our results should be compared with estimated of the costs of implementing conservation practices to be fully relevant. If the costs of implementing no-till on Texas Blackland Prairie are \$100 per acre, then we find that on average, the marginal benefits of improved soil quality characteristics may not exceed the marginal costs, on average (assuming the \$87 per acre benefit from above). However, we also find that farmers operating on flat topography are willing to pay \$136 (\$88 + \$1 + \$47) per acre to improve water infiltration and organic matter by the average amount from switching to no-till and moving from medium to low compaction. *These* farmers would need less incentive to adopt no-till than other groups, which could lead to more efficient funding decisions from the United States NRCS or Farm Service Agency.

3.7 Discussion and Limitations

We follow a set of best-practices outlined in Johnston et al. (2017), including administering focus groups, performing pilot studies, and refining our experimental design before deploying a survey with a discrete choice experiment to farmers in central Texas. The final survey was administered through the USDA National Agricultural Statistics Service (NASS) to guarantee a representative and random sample in our study area.

We estimate several models to capture both the population-level preferences and the heterogeneous preferences that exist within sub-groups. We find that there are few differences in soil quality and rental rate preferences across groups based on their local rental rates, tillage practices, future plans, topography, or soil health attitudes.

When considering the average change in soil health characteristics that has been achieved by farmers in our study area moving from conventional tillage to no-till, we find that farmers in our study area are willing to pay approximately \$58 per acre to improve water infiltration, \$1 per acre to improve organic matter, \$88 per acre to move from high to low compaction, and \$29 per acre to move from medium to low compaction. The conditional logit results suggest statistically significant MWTP estimates, while the larger standard errors for the mixed logit models yield mostly insignificant estimates.

The significant difference in the estimated MWTP between organic matter and the other soil health characteristics deserves more attention. It may be the case that farmers truly value improvements in water infiltration 60 times more than a comparable improvement in organic matter. If the majority of farmers are satisfied with the prevailing average organic matter levels in their area but are unsatisfied with some levels of water infiltration, this could be the case. Alternatively, farmers may be more comfortable with the way water infiltration was presented in the DCE (as inches absorbed per hour) than organic matter (as percent by soil mass). The way indicators of environmental quality are measured and presented is important (see Boyd et al. (2015) for evidence), and future work should experiment with alternative soil quality indicators to test how soil quality measurement affects valuation efforts.

For many respondents, the estimated WTP measures are larger than the prevailing rental rates across the study area. While this may raise concerns of a possible failure of the scope test, we speculate that it may be rational to expect WTP per acre measures that are larger than the rental rate per acre. The rental rate is an annual cost, while investments in soil health may be fixed, or at least variable for only a short time period. For example, to engage in strip-till, a farmer would need specialized tillage equipment which could be a substantial investment. It is not necessarily helpful to compare the WTP estimates directly to the prevailing rental rates. Rather, future work should be focused on comparing WTP estimates with the marginal costs of adopting soil-health practices that could achieve improvements in manageable soil health characteristics. A rational farmer should equate expected marginal costs with expected marginal benefits, and we only provide estimates of the marginal benefits in this work.

To our knowledge, we are the first study to estimate the willingness-to-pay for changes in soil quality characteristics using a discrete choice experiment. As such, we have learned important lessons that future researchers should keep in mind. First, we have concerns that our pivoted design on the rental rate would be better if we used an *online* survey. In an online format, each rental rate level could immediately pivot from a respondent's response, but in the paper format we were forced to use a generic \$10 pivot and were not able to display an actual dollar amount.[15] This could bias the estimated coefficient(s) on rental rate towards zero, and thus bias the estimates of WTP upwards. Second, we encourage replication of our work in other regions. Our results are not likely to be externally valid for other areas with different inherent soil quality characteristics. The marginal benefits of improvement of manageable characteristics are likely dependent on the local inherent conditions. For example, farmers in Iowa probably value changes in water infiltration quite differently than farmers in our study area.

Nonetheless, in this study we provide novel estimates of the marginal benefits of manageable soil quality improvement for farmers in central Texas. The estimated benefits are not directly tied to production output, as has been the dominant method of valuation for more

[15] On the other hand, we are cognizant that response rates may decline in online surveys, relative to paper surveys.

than half a century (Ciriacy-Wantrup, 1947). We find evidence that farmers are willing to pay to improve manageable soil health, and while some groups are willing to pay more than others, we hope future valuation work continues to acknowledge that there is more to the value of soil than what it produces.

4. THE EFFECT OF A PARTICIPATORY GROUNDWATER MANAGEMENT PROGRAM ON GROUNDWATER LEVELS

4.1 Introduction

Increasing water scarcity and competition for surface water has increased the pressure on the groundwater reserves of the world. The less developed world depends substantially on the groundwater sources because of less reliable public water supply systems. In developing and heavily populated countries like India and China, groundwater sources have reached unsustainable levels of exploitation. For example, India is home to 18 percent of the world's total population and only 4 percent of the world's water resources (Reddy and Reddy, 2020). Nearly 90 percent of India's rural domestic water needs and 70 percent of agricultural needs are fulfilled from drawing water from aquifers (Central Groundwater Board, 2017).

After the Green revolution in the 1960s, India's consumption of groundwater rose significantly on the back of increases in irrigated agriculture. Much of this increase was driven by development of personal irrigation systems - dug wells and borewells - and the use of subsidized electricity to pump groundwater (Zaveri et al., 2016). Between 1950 and 2013, the net irrigated area tripled from 21 million hectares to 68 million hectares and the share of ground water irrigation through wells during the same period rose substantially from 28 percent to 62 percent (Central Groundwater Board, 2017).

Kulkarni, Shah, and Shankar (2015) compare India's Central Groundwater Board's assessment of groundwater depletion across India between 1995 and 2009. The comparison shows that groundwater reserves depleted sharply between 1995-2005, before showing marginal improvement between 2005-2009. According to the National Groundwater Assessment in 2011, 68 percent of a total of 6607 units[1] are safe, 16 percent are over-exploited, 10 percent are semi-critical, and 3 percent are critical (Central Groundwater Board, 2011). The

[1] Units are defined as districts and/or mandals, which are geographically equivalent to the county level and sub-county level, respectively.

fall in groundwater levels has raised concerns about the long-term sustainability of irrigated agriculture in India (Hanjra and Qureshi, 2010).

In a study on the impact of groundwater access on conflict and poverty, Sekhri (2014) argues that when the water table falls below a certain level, users with inadequate access to the technology required to draw water from lower depths find themselves constrained. The results from this study suggests that deepening groundwater depth has led to more disputes among the users and rising rural poverty.

The critical state of India's groundwater resources is not new. Groundwater management programs in India have traditionally focused on developing new groundwater extraction solutions to solve the problem of access. Continued development of new water infrastructure without concurrent policies to moderate groundwater extraction has led to the problem of water scarcity at the aquifer level. There is a growing literature on the importance of "demystifying the science of aquifers" for the stakeholders to regulate groundwater use (Van Steenbergen, 2006; Kulkarni, Shah, and Shankar, 2015; Joshi, Kulkarni, and Aslekar, 2019). Van Steenbergen (2006) uses examples from India, Pakistan, Yemen and Egypt to highlight the effectiveness of self-regulation in moderating groundwater use. Kulkarni, Shah, and Shankar (2015) argue the understanding of groundwater as a common-pool resource is indispensable to sustainable development of groundwater. The authors suggest that the complexity and fungibility of groundwater must be factored into the governance policies to allow for aquifer-based management of the resource. Users often behave myopically due to their poor understanding of groundwater as a common resource. In absence of proper understanding of resource characteristics and its availability, users are likely to pump groundwater at an unsustainable rate. A key aspect of this aquifer-based groundwater management, therefore, is community participation in understanding and managing the common-pool resource.

The knowledge-intensive approach to groundwater governance is often referred to as participatory hydrological monitoring (PHM). PHM involves building capacity to train the groundwater users in collecting information on groundwater availability and underlines the importance of regulating use based on estimated availability. In India, the Andhra Pradesh Farmer Managed Groundwater Systems (APFAMGS) intervention is a well-known participatory hydrological monitoring program. The APFAMGS intervention involved equipping farmers with the necessary training so they can perform data collection, basic analysis and undertake a crop-water budgeting exercise (Reddy and Reddy, 2020). To date, a number of researchers, in addition to the funding organization - the Food and Agriculture Organization of the United Nations - have investigated the extent of impact the APFAMGS program has had on groundwater levels (FAO, 2010; Verma et al., 2012; FAO, 2013; Reddy, Reddy, and Rout, 2014). To our surprise, the assessments could not be more different from each other. While the FAO (2013, 2010) terminal reports conclude that the program had a significant positive impact on groundwater table, Verma et al. (2012) present deeply concerning realities of the program. The existing studies do not conclusively show that the intervention had any favorable impact on groundwater depth during the program period. Two other features of the various reports and studies stand out. First, many of these studies utilize a small set of treatment units (one village in case of Reddy, Reddy, and Rout (2014)) or in some cases district-level data to evaluate the impact of the APFAMGS program. The district-level approach is problematic because the coverage of the program in project districts is less than ten percent of the sown area (Reddy and Reddy, 2020). Our study - to our knowledge - is the first to use a causal-inference approach to assess the impact of the program.

Our study has significant policy implication for a large-scale upcoming groundwater management program in India. In the beginning of 2020, the government of India signed a $500 million loan agreement with the World Bank to support India's groundwater sustainability program (World Bank, 2018). The program known as the Atal Bhujal Yojna or Atal Groundwater Management Program (AGWM), named after a former prime minister, will cover 78 districts across 8 states grappling with the problem of unsustainable groundwater extraction. Participatory hydrological monitoring of groundwater by training the stakeholders to collect and read necessary information on groundwater availability is an integral part of the

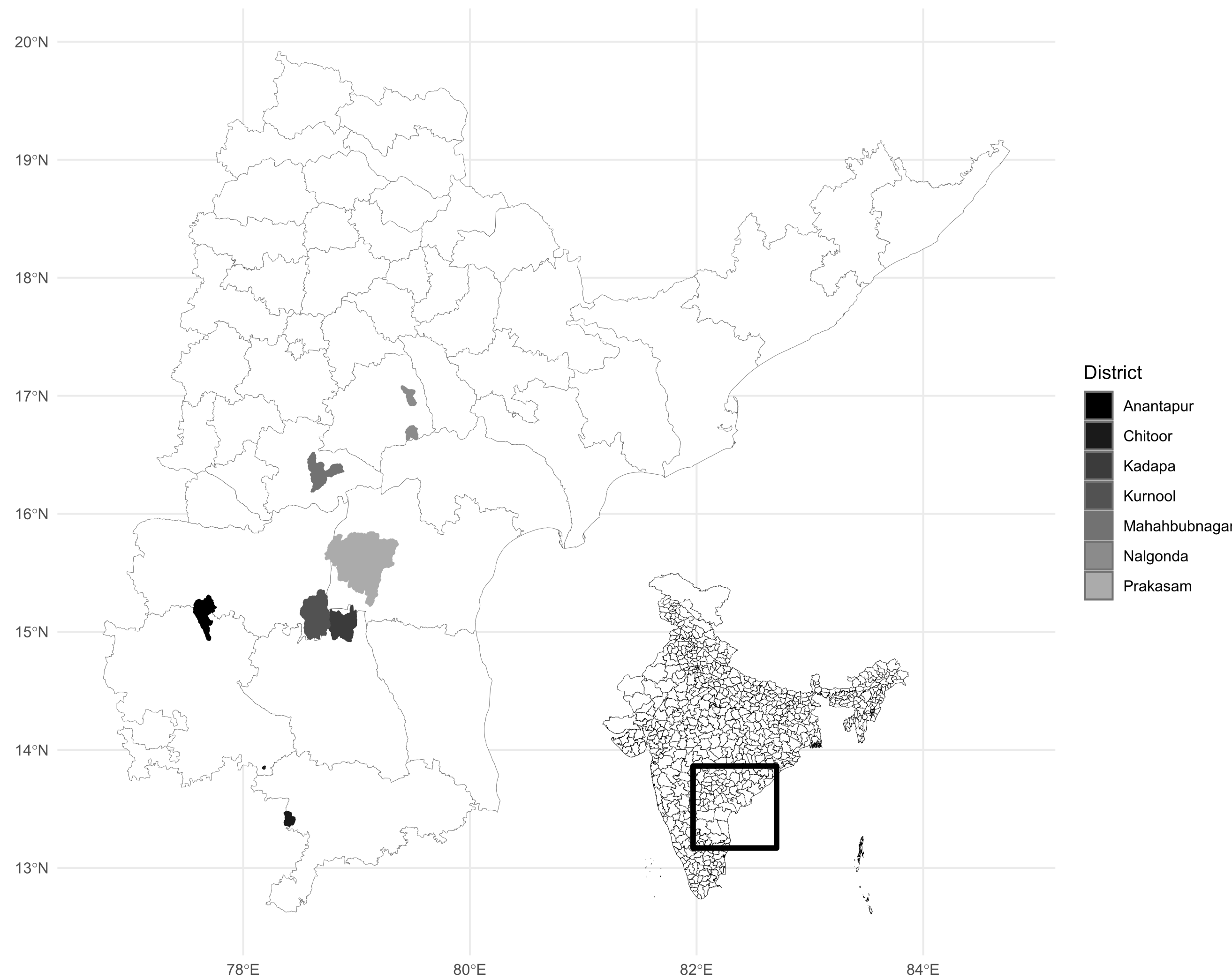
AGWM. In the words of India's Department of Economic Affairs "The AGWM program intends to strengthen the institutional framework for participatory groundwater management and encourage behavioral changes at the community level for sustainable groundwater resource management." Through our study on the causal impact of APFAMGS intervention on depth-to-groundwater in the program region, we offer some warning signs that programs that involve participatory hydrological monitoring may not be effective. In particular, our results suggest that purely participatory programs may have little effect on groundwater reserves.

4.2 Program Background

The Andhra Pradesh Farmer-Managed Groundwater Systems (APFAMGS) program was preceded the Andhra Pradesh Groundwater Borewell Irrigation Schemes (APWELL) launched in 1995 in seven drought prone districts in the state of Andhra Pradesh: Anantapur, Chittoor, Kadapa, Kurnool, Mahbubnagar, Nalgonda and Prakasam (Figure 4.1). The APWELL program was designed to help small holder farmers access irrigation technologies (FAO, 2013). APWELL covered around 14,000 hectares of irrigated agriculture and involved over 14,000 farmers from the seven districts in Andhra Pradesh (Garduño et al., 2009). The project installed community water wells and distribution systems. Prior to the APWELL implementation, it became clear to the agencies that the sustainability of the borewells to be provided to the small- holder farmers as part of the program would become a challenge as groundwater withdrawals outside the APWELL command areas continued to increase (FAO, 2013). It was suggested that self-monitoring of the wells by the users is more viable and sustainable, and thus the idea of participatory hydrological monitoring was developed.

APWELL's introduction of micro-irrigation practices was found to have increased groundwater use efficiency and diversified farming towards less water-consuming crops. The program also had a favorable impact on income of the landless (Garduño et al., 2009). The APWELL program was closed abruptly in 2003. Based on the impact and institutional evaluation of the APWELL program, a new refined project proposal - Andhra Pradesh

Figure 4.1: Map Of The APFAMGS Program Area



Framer-Managed Groundwater Systems (APFAMGS) - was developed in 2003. APFAMGS was written to address groundwater management through crop-water-budgeting, groundwater monitoring, and informational interventions on good groundwater management.

The Bharati Integrated Rural Development Society (BIRDS) was assigned as the primary local agency to oversee the implementation of the APFAMGS program. BIRDS partnered with six other NGOs - each of these NGOs were responsible for supporting the APFAMGS program in a number of areas. APFAMGS was launched in the same seven APWELL drought prone districts of Andhra Pradesh. The APFAMGS intervention built on the APWELL project by strengthening the community based institutions and promoting the use of technology in the participatory hydrological monitoring process. A hydrological monitoring network was established by APFAMGS to create a platform for data collection by farmers. APFAMGS also created a computer-based *Habitation Resource Information System* for storing data collected by the farmers at the partner-NGO level. In addition to these systems, information kiosks were deployed to help farmers understand the impact of their crop choices on the groundwater level in the aquifer. Using the interactive crop-water kiosk, the farmers could input their own and other farmers' crop selection and immediately see the impact of their selection on the groundwater balance in their hydrological unit.

The APFAMGS Project formally ended in 2009, but many of the key activities such as hydrological monitoring and preparation of crop-water budgets continue to take place in several of the APFAMGS hydrological unit even after the FAO withdrew from the project.

4.3 Program Evaluation

We identify the causal effect of APFAMGS on local groundwater reserves, as measured by both the depth to groundwater from year to year and intra-year differences in the depth-to-groundwater level. We focus on intra-year differences primarily because of the natural behavior of groundwater levels in response to the Indian agricultural cycle. There are four agricultural seasons in Andhra Pradesh: rabi (January - March), pre-monsoon (April - June), monsoon (July - September), and kharif (October - December). The kharif harvest follows the monsoon, where surface water is more abundant than in other times of the year. Rabi crops are typically more drought tolerant or more dependent on irrigation, because outside of the monsoon rains, precipitation is exceedingly rare. See Figure 4.2. Panel A shows average rainfall for Andhra Pradesh.[2] Note that because rainfall data is not available for most of the years of interest in our study, we do not use it for any empirical analysis. We present it here to showcase typical rainfall throughout the four seasons in Andhra Pradesh. The grey bars are precipitation totals (average) for Andhra Pradesh, and the approximate boundaries of the the seasons are overlaid in black. The calendar year follows an awkward cycle of cropping, rain, and cropping again, during which the groundwater levels in the area change dramatically. In our analysis, we focus on what we call an *agricultural year*. We let the agricultural year start during the monsoon season and end at the next calendar year's pre-monsoon season. See the bottom part of Panel A in Figure 4.2. Using the agricultural year allows us to compare the natural kharif and rabi seasonal pairs that follow a monsoon. Farmers in the region grow and harvest crops in the kharif and rabi seasons, and use the monsoon and pre-monsoon seasons to prepare their land or engage in other non-growing activities.

In Panel B of Figure 4.2, we show the average depth-to-groundwater levels across the entire program area and across the entire available time period. The depth-to-groundwater levels are generally highest in the monsoon and kharif seasons. While rain falls mostly during the monsoon season, there is some temporal lag as surface water replenishes aquifers, and therefore the groundwater levels remain high during the kharif growing season despite relatively low levels of rainfall. After the kharif season, groundwater levels consistently fall for the remainder of the agricultural year as farmers extract water for their operations.

Importantly, we do not observe the timing of the seasonal monsoon, nor do we observe when the seasonal average groundwater levels were taken. That is, an individual well's observation for monsoon-level depth-to-groundwater could be an average of measurements early or late in the monsoon, and we do not observe the timing of the measurements used in the provided seasonal average from India WRIS.

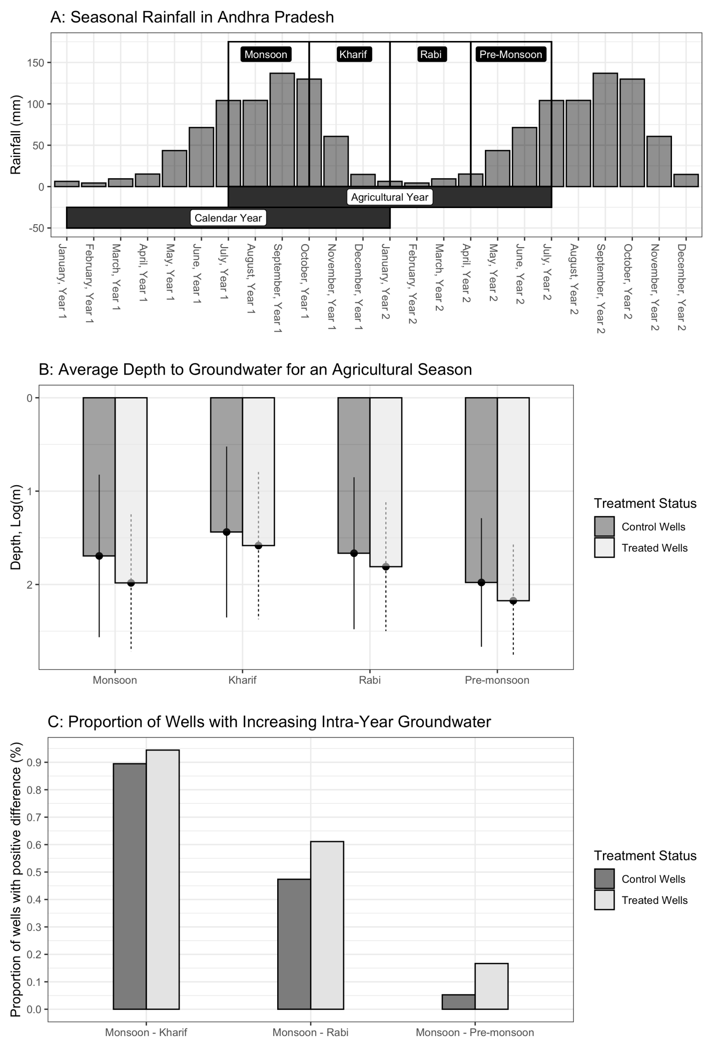
In Panel C of Figure 4.2, we show the proportion of wells that have rising intra-year

[2]The average is provided directly by India WRIS, but the average is filtered to include only districts in the treatment and control areas.

groundwater levels by comparing the monsoon season with the other three seasons. There is a clear pattern (that can also be seen in Panel B). Between the monsoon and kharif seasons, groundwater is generally being replenished. Then groundwater levels begin to fall, and about half of wells have groundwater levels that are slightly lower than in the monsoon season. By the pre-monsoon season, a large majority of wells have groundwater levels lower than in the monsoon season.

Groundwater levels fluctuate throughout the agricultural year, following the natural hydrological cycles of the area. See Figure 4.3. Panel A shows the depth-to-groundwater levels for wells within the APFAMGS treated area and control wells outside the treatment area, in logged meters. Prior to the start of APFAMGS in 2006, the treated and control wells followed similar booms and busts cycles. After 2006, they appear to continue to follow each other closely until 2014 when there is a slight divergence where the depth to groundwater actually *increases* (the opposite of the intended effect) for treated wells compared to control wells. We empirically test whether the average depth-to-groundwater is different between the two groups before and after 2006 below. Because of the natural fluctuations in groundwater levels, we are primarily interested in the intra-year changes in groundwater levels. Panels B-D in Figure 4.3 show the *difference* (in logged meters) between the monsoon and other seasons depth-to-groundwater levels in absolute value. For example, a value of 0.5 in Panel B of Figure 4.3 indicates a difference of $e^{0.5} = 1.65$ meters between the monsoon and kharif depth-to-groundwater levels. A large value for this difference measure suggests severe fluctuation between seasons, while a small value suggests little fluctuation. We hypothesize that if natural rainfall fully recharges an aquifer and farmers over-pump from the aquifer during the cropping seasons, the difference in groundwater levels between the seasons would be large. Meanwhile, more conservative pumping would likely result in a smaller difference in depth-to-groundwater during the year. Indeed, there appears to be evidence of this in Panels B-D in Figure 4.3. The seasonal differences in depth-to-groundwater for treatment and control groups is similar prior to 2006, but sometime after APFAMGS, the treated wells

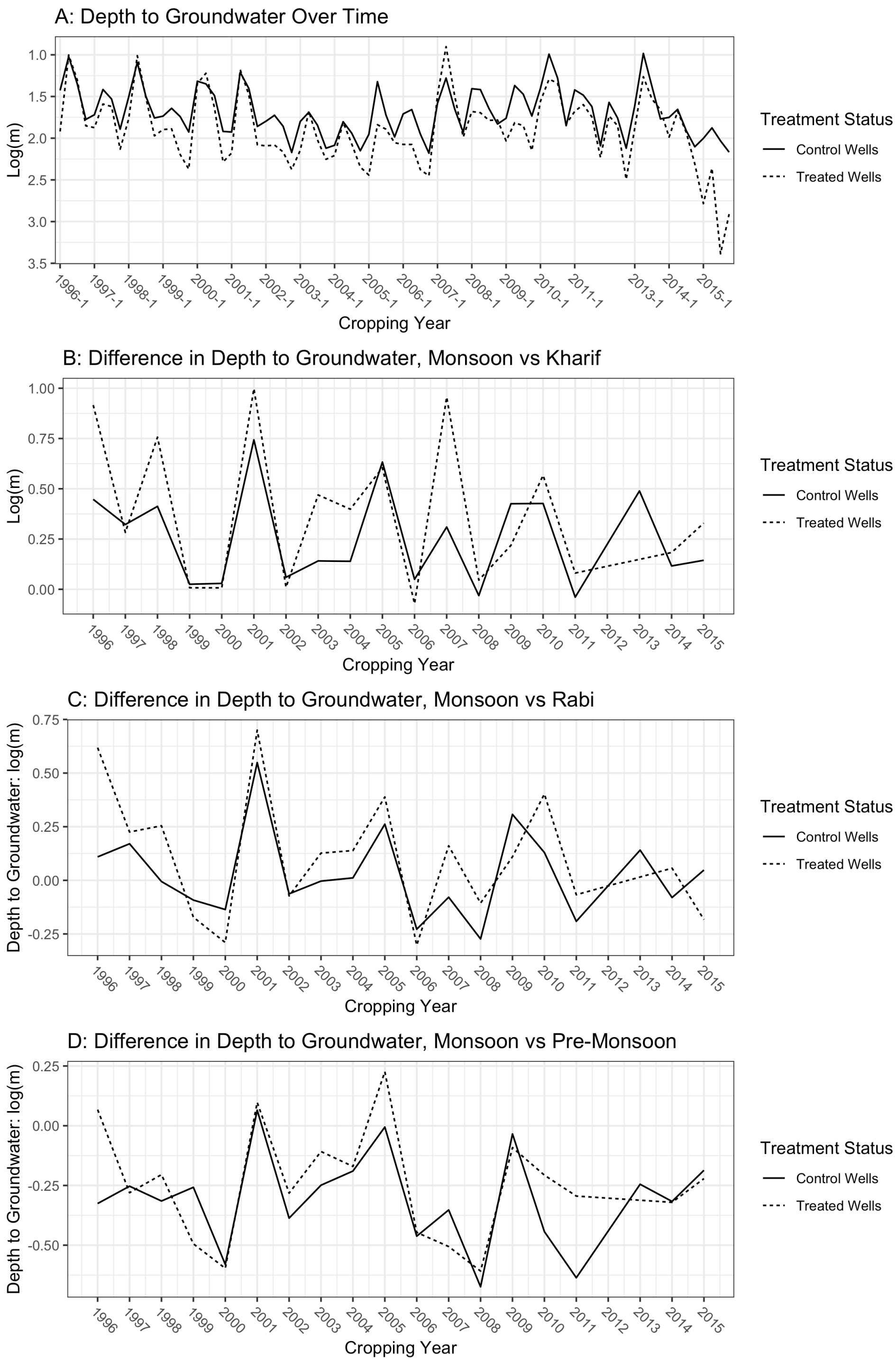
Figure 4.2: Seasonal Rainfall And Groundwater Levels In Andhra Pradesh



Source: India WRIS. Note that rainfall data is not available prior to 2014. Groundwater data is not available in monthly increments. In Panel B, the average and standard deviations are presented.

appear to show a smaller (however slight) seasonal differences in depth-to-groundwater. This is particularly visible in the difference between the solid and dashed lines in Panels C and D: the rabi and pre-monsoon seasons. We empirically test whether the seasonal difference in depth-to-groundwater is different between the treatment and control groups before and after 2006 below.

Figure 4.3: Groundwater Levels In Program Area



Source: Groundwater data from India WRIS

4.3.1 Empirical Model

Let d_{0it} be some distance-to-groundwater outcome (either depth or intra-year difference in depth) for a unit i not exposed to APFAMGS at time t, and d_{1it} be the same measure for i when exposed to APFAMGS. We use the following identifying assumptions in a difference-in-differences framework:

$$E[d_{0it}|\alpha_i, \tau_t, T_{it}] = E[d_{0it}|\alpha_i, \tau_t] \tag{4.1}$$

The identifying assumption states that the distance outcome is independent of treatment status T_{it} after conditioning on unit and time fixed-effects. This is a common identifying assumption for two-way fixed-effect models, and while there are established problems with two-way fixed effects models when treatment time varies across units (see, for example, Goodman-Bacon (2018) for a thorough treatment of the problems), we assume APFAMGS was initiated simultaneously across the program areas.[3] We therefore avoid the pitfalls of differential treatment times across units and the subsequent bias of the treatment effect.

As explained in the previous section, we are more interested in the intra-year difference in depth-to-groundwater between the monsoon and other seasons. Our identifying assumption for this intra-year measure is identical to equation 4.1: the outcome of interest is independent of treatment status conditional on unit and time fixed effects. Note that the intra-year difference in groundwater depth is defined as:

$$d_{it} = d_{it}^M - d_{it}^S \tag{4.2}$$

where d_{it}^M is the depth-to-groundwater for unit i in agricultural year t during the monsoon season M, and d_{it}^S is the same measure for any of the three other seasons, S. A positive value for d_{it} indicates a groundwater level that is higher in the season S than in the monsoon, and a negative value indicates a groundwater level that is higher in the monsoon than the

[3] This is an assumption because we do not know with certainty when each program started, nor do we have a definition of what *started* means. Programs like APFAMGS start slowly; they are not magical light-switches.

season S. Recall from Figure 4.2 that groundwater levels in the kharif season are mostly higher than in the monsoon, approximately equal between the rabi and monsoon seasons, and greater in the monsoon than in the pre-monsoon season.

Under identifying assumption 4.1, we estimate the following model:

$$\ln(d_{it}) = \alpha_i + \tau_t + \delta T_{it} + \varepsilon_{it}, \tag{4.3}$$

where $\ln(d_{it})$ is the logged outcome variable (depth-to-groundwater or intra-year difference of depth-to-groundwater) for unit i in agricultural year/season t, α_i is a unit-specific fixed effect, τ_t is a time fixed effect, and T_{it} is an indicator for APFAMGS treatment status for unit i agricultural year/season t.

In equation 4.3, the treatment effect of APFAMGS is captured by δ. We recover several versions of δ. The two-way fixed effects estimator takes the difference across treatment and control units of the difference of our outcomes across time. We are able to define the length of time over which to average the outcome variables. We therefore estimate $\hat{\delta}_5$, which is the average treatment effect using a five year post-treatment window, and $\hat{\delta}_{10}$, which uses a 10-year window. We also estimate a dynamic version which isolates the treatment effect for each agricultural year after treatment individually. However, as we explain later, because of our sparse data we are less confident of the dynamic version of the model.

For the average depth-to-groundwater we expect $\delta < 0$ since we expect units in APFAMGS-treated areas to pump more conservatively than untreated areas, which would increase the level of the water table and thus *decrease* the depth-to-groundwater level.

The intra-year difference model is more complicated. Because groundwater almost uniformly increases between the monsoon and kharif seasons, we expect treated areas to recharge faster and thus we expect $\delta > 0$, meaning the recharge difference is larger for treated than untreated areas. One reason we expect this is because treated areas may seek to conserve groundwater and stockpile a large reserve for use later in the season. We have no strong expectation for the rabi season. We expect $\delta < 0$ during the pre-monsoon intra-year difference, because this indicates that the treated areas experience less of a seasonal drop in groundwater levels than untreated areas. Recall that groundwater levels decrease between the monsoon and pre-monsoon seasons.

4.3.2 Data

Information on groundwater depth comes from the *Water Resources Information System* (India-WRIS), an online database managed by the Indian Department of Water Resources, River Development, and Ganga Rejuvenation. India-WRIS consolidates a variety of water quality and quantity data from several state and regional agencies. We observe groundwater monitoring stations across what was then the state of Andhra Pradesh; in 2014 a northeast section of the state split into its own state known as Telangana. The administrative bodies of India, in decreasing geographic size are: country, state, district, mandal, blocks/villages. APFAMGS was administered in 63 internally-defined hydrological units across 9 mandals in Andhra Pradesh and Telangana, but there is potential for substantial geographic spillovers. The program was a combination of education workshops and technical assistance, and farmers from outside the APFAMGS program area could have attended parts of the program and enjoyed some benefits from improved groundwater management.

In an ideal setting, we would observe groundwater levels and other characteristics for every unit (e.g. farm) across every season for a number of years before and after treatment. In reality, while we observe some groundwater levels for some wells for some years, there are substantial holes in our data. We therefore take two approaches to estimation. First, we accept the missing data and estimate a two-way fixed effects model with an unbalanced panel. Second, we create a fine grid of points across the landscape and use the spatial interpolation technique known as kriging to first estimate groundwater levels for every grid point for every year, and then estimate a two-way fixed effects model using the grid points as opposed to the sparse well-level observations. We describe the kriging approach in the next section.

4.3.3 Spatial Interpolation

As in many settings, we face sparse spatial-temporal data. The field of geostatistics has developed an approach to interpolating missing data that is governed by some spatial-temporal process (Calder and Cressie, 2009; Matheron, 1963). The approach is known as kriging. Specifically, let there be some set of spatial data $Z(\cdot)$ defined at points s. For example, we observe groundwater levels across Andhra Pradesh for a set of wells, $S = \{s_1, s_2, ..., s_S\}$. Our goal is to predict the groundwater level $Z(s_0)$ for some unobserved point in space, s_0. Let $Z(s) = Y(s) + \varepsilon(s)$, where $Y(s)$ is some process with mean $E[Y(s)] = \mu_Y(s)$ and mean-zero error $E[\varepsilon(s)] = 0$. Further, let the covariance between two points s_i and s_j for the $Y(s)$ process be $C_Y(s_i, s_j) = cov(Y(s_i), Y(s_j))$. Then, we predict $\hat{Y}(s_0)$ as

$$\hat{Y}(s_0) = \mu_Y(s_0) + \tilde{c}_Y(s_0)'\Sigma^{-}1_Z(Z - \mu_Y), \tag{4.4}$$

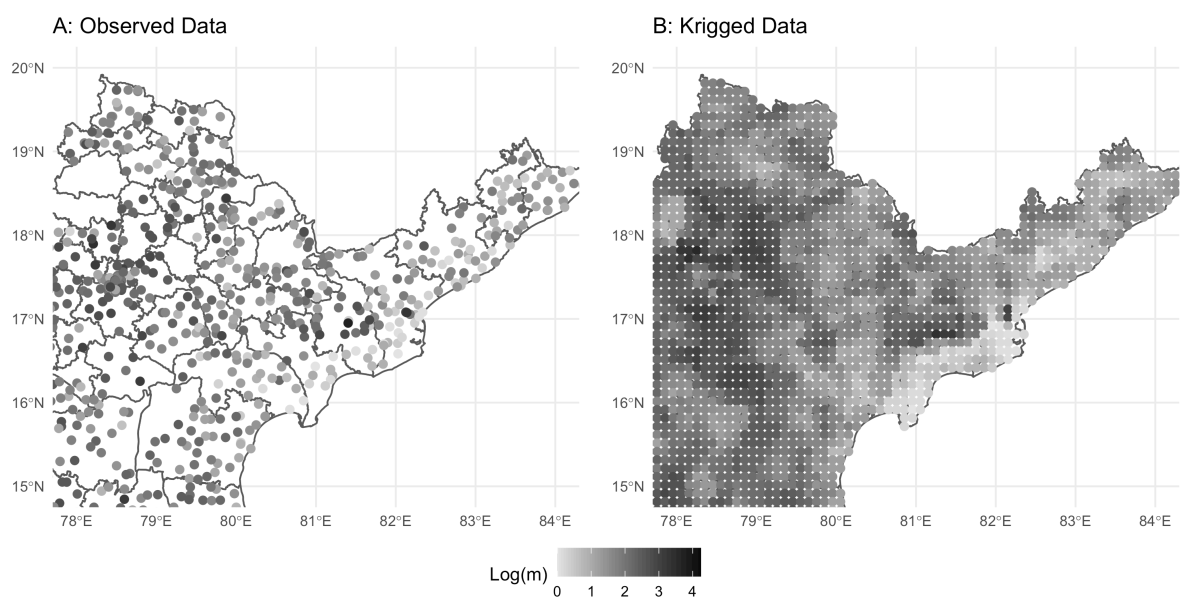
where $\mu_Y(s_0)$ is the expected value of $Y(\cdot)$ at the unobserved point, $\tilde{c}_Y(s_0)$ is the covariance between $Y(s_0)$ and Z, and Z and μ_Y are observed for each point $s \in S$. The matrix Σ_Z is the variance-covariance matrix of $C_Y(s_i, s_j)$. See Calder and Cressie (2009) for a more detailed explanation of the kriging process and the above notation.

Kriging can proceed to predict values for, say, groundwater levels according to equation 4.4 if we observe $C_Y(s_i, s_j)$ and the variance of the $\varepsilon(s)$. However, both are latent. We therefore use a variogram which estimates the covariance function according to the distance between any two points.

Our specific approach is as follows: we use the observed spatial-temporal data to fit a variogram model to predict the covariance in groundwater levels between any two points according to their distance from each other. We then lay a grid of points over the study area, where each grid point is 11.1km (0.1 degrees) away from its nearest four neighbors. We then use the variogram model to predict the groundwater levels for each grid point in space. We then repeat this process for each year to arrive at a balanced panel of groundwater levels across the entire study area for each season-year.

See Figure 4.4 for a demonstration of the benefit of the krigging process. In Panel A, we show the observed measurements for the 2004 rabi cropping season. In Panel B, we show the krigged data for the grid overlaid on the program area. Krigging affords us a substantial increase in sample size, but comes at an obvious cost: the dependent variable in our two-way fixed effects model is a predicted value. Fortunately, krigging is a much more sophisticated process than simple interpolation based on, say, nearest-neighbor matching. Nonetheless, krigging results in smooth spatial predictions and is thus not suited for strong spatial discontinuities or anomalies. Since groundwater also does not exhibit strong spatial discontinuity, we assume krigging captures the true but otherwise unobserved groundwater levels across our study area.

Figure 4.4: Depth To Water Table, Observed Versus Krigged Data



Recall our empirical model for estimating the effect of APFAMGS on depth-to-groundwater: $\ln(d_{it}) = \alpha_i + \tau_t + \delta T_{it} + \varepsilon_{it}$, where $\ln(d_{it})$ is the logged depth-to-groundwater (in meters) for well i at time t and δ captures the effect (in percentage terms) of APFAMGS on depth-to-groundwater. In the krigging setting, we use the observed data to make spatial interpolations across a grid covering the study area. Let $Y_{it} = \ln(d_{it})$ be the true observed value. When we use krigged values, we are actually estimating the following model:

$$\tilde{Y}_{it} = \alpha_i + \tau_t + \delta T_{it} + \varepsilon_{it} + v_{it}, \tag{4.5}$$

where

$$\tilde{Y}_{it} = Y_{it} + v_{it}. \tag{4.6}$$

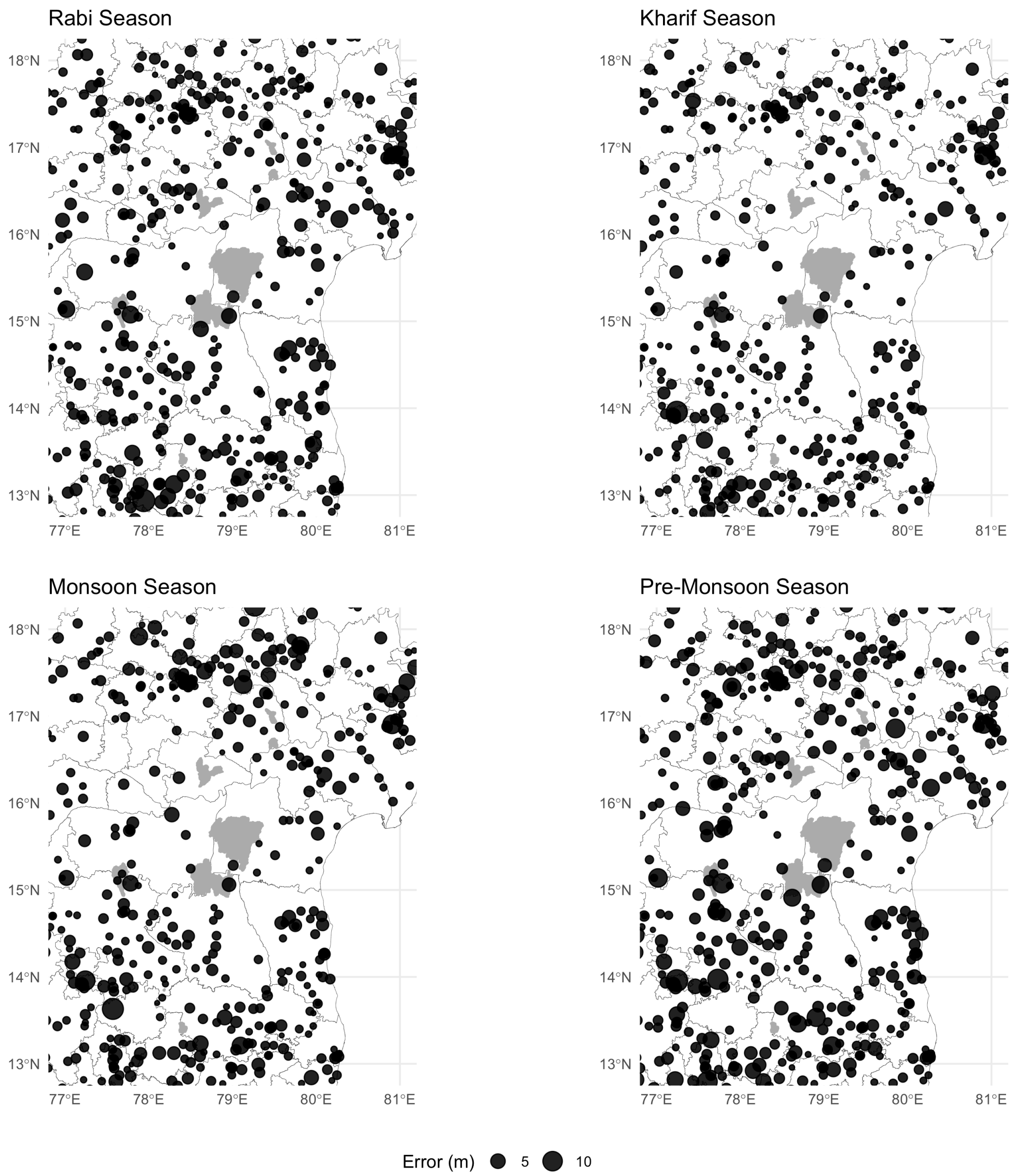
That is, we observe some measurement error in the dependent variable of our two-way fixed effects model.[4] Measurement error in the dependent variable of a linear model does not effect affect consistent estimation of the desired treatment effect, provided the error v_{it} in equation 4.6 is uncorrelated to T_{it} in equation 4.3 (Wooldridge, 2010). We provide some evidence of no correlation in a cross-validation exercise.

We do not observe real groundwater levels for any of the grid points from the krigging approach. However, we perform a cross-validation test of the krigging process to both test the size of prediction error and correlation to treatment status. Specifically, we isolate real well-level observations of groundwater levels for each season in each year. We then randomly select two-thirds of the data for training, and one-third for testing. We train a krigging model on the training data set, and predict groundwater levels for each season on the testing set. We then take the difference between the predicted and real values of the test data set to understand both the size and distribution of the error, v_{it}.

See Figure 4.5 for a map of the prediction error of the krigging process. We use the year 2006 as an example here, though results are similar across all years. Mean, median, and 90^{th} percentile prediction errors (in meters) for the cross-validation are presented in Table 4.1. On

[4]Note that this measurement error does not effect consistent identification of our causal model, but the standard errors in our main model are larger than they would be without measurement error, because of the presence of v_{it}. The measurement error, v_{it} is naturally unobserved, making a correction to the standard errors in our OLS model impossible. We know that there is measurement error which increases the size of the standard errors in our OLS model, but we do not know the magnitude of that increase.

Figure 4.5: Depth To Water Table, Kriging Measurement Error



average, the kriged depth-to-groundwater levels are off by approximately 2 meters across all four seasons. This roughly corresponds to a measurement error of 23 - 31% depending on the season. The size of the error is concerning. Recall that the cross-validation error is a measure of v_{it}. The relatively large size of v_{it} does not prevent consistent estimation of δ (the treatment effect of APFAMGS). However, as v_{it} increases, the probability of making a type II error in our statistical inference increases. That is, the large measurement error we observe in cross-validating the krigging process suggests that in our difference-in-differences estimator may lead us to an insignificant treatment effect that might be significant in reality.

Table 4.1: Krigging Cross-Validation Summary Statistics

	Rabi	Kharif	Pre-Monsoon	Monsoon
Mean Error (m)	2.03	2.17	1.87	2.04
Median Error (m)	1.60	1.67	1.55	1.62
90^{th} Percentile Error (m)	3.32	3.77	3.03	3.43
Median Error (%)	30.68	31.21	22.98	26.10
corr(v_{it}, T_{it})	0.023	0.049	0.038	0.008
	(0.717)	(1.453)	(1.203)	(0.225)

Note:
t-statistic from Pearson's product-moment correlation in parentheses
*p<0.1; **p<0.05; ***p<0.01

In the final step of the cross-validation process, we test the correlation between measurement error v_{it} and treatment status of a well T_{it}. So long as the two measures are uncorrelated, we can recover a consistent treatment effect of APFAMGS on groundwater levels using krigged data. The last row of Table 4.1 reports the Pearson's product-moment correlation measure, with the t-statistic for the hypothesis of zero correlation in parentheses. We fail to reject zero correlation across all seasons, suggesting the measurement error is uncorrelated with treatment status of a well.

We therefore are comfortable using krigged data to measure the treatment effect of APFAMGS on groundwater levels, though we are wary of the relatively large measurement error and the implication large measurement error has on the probability of making a type II error. Insignificant treatment effects may be truly insignificant or the result of sampling or prediction error.

4.4 Results

We first estimate the average treatment effect of APFAMGS on the average annual depth-to-groundwater across a five and ten year window. The five year window uses data only from the first five years after the program, and the ten year window uses data only from the first ten years after the program. Results are presented in Table 4.2. Using the krigged data, we find that APFAMGS-treated areas experienced average depth-to-groundwater levels that were approximately 4.5% lower than untreated areas (i.e. higher water tables). The ten year average effect is dampened to approximately 1.9%, although neither effect is statistically different from zero. We find a smaller average effect for the five year window using the raw data, and a slightly positive effect for the ten year window. The estimated treatment effects using the raw data are again not statistically different from zero.[5]

Next we explore the intra-year differences in depth-to-groundwater. The intra-year difference in depth-to-groundwater is the difference between the monsoon average depth-to-groundwater and another season's average depth-to-groundwater. We calculate these seasonal differences between the monsoon season and the kharif and rabi cropping seasons and the pre-monsoon season. Results for are shown in Table 4.3. Using the krigged data, we find that the difference in groundwater levels between the monsoon and kharif season are 9% and 9.8% smaller in APFAMGS-treated areas than untreated areas, using a five and ten year post-treatment window respectively. Using only the raw data, we find a 16.1% reduction in the kharif seasonal difference using a five year window, and a 14.4% reduction in the difference using a ten year window.

The negative coefficient for the kharif seasonal difference coupled with the fact that groundwater levels rise between the monsoon season and kharif season means that APFAMGS treated areas experience *less* recharge than untreated areas. The negative coefficient is the opposite of our *a priori* expectation.

[5] In an alternate specification we find some significant effects when we expand the treatment area to allow for spatial spillovers. However, the results are sensitive to the size of the expanded area, and we are not aware of the extent or magnitude of spatial spillover of the APFAMGS program

Table 4.2: Effect Of APFAMGS On Annual Depth To Groundwater

	Five-Year Effect	Ten-Year Effect
	(1)	(2)
Krigged Data		
$\hat{\delta}$	-0.045	-0.019
	(0.043)	(0.033)
Observations	2,430	3,078
R^2	0.001	0.0002
F Statistic	1.906	0.435
Raw Data		
$\hat{\delta}$	-0.011	0.016
	(0.066)	(0.031)
Observations	1,929	2,357
R^2	0.00001	0.00002
F Statistic	0.018	0.040
Note:		*p<0.1; **p<0.05; ***p<0.01

We find no intra-year differences between the monsoon and rabi cropping seasons. Rabi groundwater levels are, on average, similar to monsoon groundwater levels, as shown in Figure 4.2.

Using the raw data, we find that the difference between monsoon and pre-monsoon groundwater levels is 21.6% and 16% lower for APFAMGS areas compared to untreated areas, using a five and ten year treatment window respectively. The negative coefficient coupled with the fact that groundwater levels fall between the monsoon and pre-monsoon seasons means that APFAMGS areas experience a less severe groundwater draw-down compared to untreated areas.

In our final specification, we estimate the effect of APFAMGS on the seasonal differential depth-to-groundwater using dynamic treatment effects. In the dynamic specification, we estimate a treatment effect for each agricultural year separately, instead of taking an average

Table 4.3: Effect Of APFAMGS On Intra-Year Depth To Groundwater

	Kharif[a]	Rabi[a]	Pre-Monsoon[a]
	(1)	(2)	(3)
A: Krigged Data, Five-Year Window			
$\hat{\delta}$	-0.090^{**}	0.022	-0.030
	(0.039)	(0.045)	(0.042)
Observations	2,430	2,430	2,430
R^2	0.001	0.0001	0.0002
F Statistic	3.329^{*}	0.177	0.397
B: Krigged Data, Ten-Year Window			
$\hat{\delta}$	-0.098^{***}	0.003	-0.034
	(0.037)	(0.040)	(0.037)
Observations	3,078	3,078	3,078
R^2	0.002	0.00000	0.0002
F Statistic	5.191^{**}	0.003	0.684
C: Raw Data, Five-Year Window			
$\hat{\delta}$	-0.161^{**}	-0.167	-0.216^{*}
	(0.079)	(0.103)	(0.112)
Observations	2,278	2,214	2,046
R^2	0.001	0.001	0.001
F Statistic	1.344	1.523	2.565
D: Raw Data, Ten-Year Window			
$\hat{\delta}$	-0.144^{*}	-0.112	-0.160^{**}
	(0.074)	(0.093)	(0.078)
Observations	2,889	2,730	2,581
R^2	0.001	0.0004	0.001
F Statistic	1.351	0.823	1.743

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

[a] *Difference between the monsoon season and the labeled season*

across the pre- and post-treatment years. That is, we estimate:

$$\ln(d_{it}) = \alpha_i + \tau_t + \delta_t T_{it} + \varepsilon_{it}, \tag{4.7}$$

which is identical to our main equation 4.3, except we now estimate a treatment effect, $\hat{\delta}_t$ for each year after APFAMGS started.

Results for this specification using the raw data are shown in Table 4.4. When we isolate the treatment into yearly indicators, we see that the benefits of APFAMGS are not consistent over time. Rather, the benefits oscillate from year to year. The estimated signs of the coefficients are generally negative, consistent with the average treatment effects presented above. However, the average treatment effect appear to be driven by the 2006 and 2009 agricultural years, and to a lesser extent the 2007 and 2014 pre-monsoon seasons.

The dynamic treatment effects again suggest that APFAMGS areas experience less groundwater recharge between the monsoon and kharif seasons, and less groundwater draw-down between the monsoon and pre-monsoon seasons.

4.5 Discussion

Participatory hydrological monitoring programs aim to educate groundwater users about the nature of aquifers and its key properties. The users are trained to monitor the groundwater levels and base their extraction decisions based on the estimated groundwater balance. The APFAMGS intervention in seven drought prone districts of Andhra Pradesh and Telangana in India is an example of a participatory hydrological monitoring program. A previous regional program (APWELL) established in 1995 increased the use of irrigation and thus increased the pressure on groundwater reserves for the area. APFAMGS was enacted to help conserve groundwater reserves by encouraging optimal groundwater extraction.

In this study, we utilize well-level data on groundwater to understand the causal impact of the APFAMGS intervention on depth-to-groundwater and seasonal variation in depth-to-groundwater in the program area. While the initial evaluations by the funding agency

Table 4.4: Intra-Year Effects, Dynamic, Raw Data

	Kharif[a]	Rabi[a]	Pre-Monsoon[a]
	(1)	(2)	(3)
$\hat{\delta}_{2006}$	-0.392^{***}	-0.335^{**}	-0.222
	(0.150)	(0.141)	(0.144)
$\hat{\delta}_{2007}$	0.394	-0.009	-0.385^{*}
	(0.303)	(0.268)	(0.213)
$\hat{\delta}_{2008}$	-0.144	-0.102	-0.066
	(0.197)	(0.174)	(0.221)
$\hat{\delta}_{2009}$	-0.413^{***}	-0.334^{*}	-0.253
	(0.130)	(0.199)	(0.454)
$\hat{\delta}_{2010}$	-0.038	0.232	0.176
	(0.188)	(0.245)	(0.115)
$\hat{\delta}_{2011}$	-0.142	0.024	0.200
	(0.114)	(0.143)	(0.139)
$\hat{\delta}_{2014}$	-0.156	0.009	-0.216^{**}
	(0.141)	(0.161)	(0.088)
$\hat{\delta}_{2015}$	0.036		-0.262
	(0.204)		(0.188)
Observations	2,889	2,730	2,581
R^2	0.004	0.002	0.003
F Statistic	1.175	0.610	0.756

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

[a] *Difference between the monsoon season and the labeled season*

- Food and Agriculture Organization - regard the program as a success, some of the later analyses have raised concerns regarding the voluntary no-incentive nature of the program. Reddy, Reddy, and Rout (2014) suggest that participatory programs are more effective when supplemented with supply side interventions such as construction of injection wells. In a survey by Reddy, Reddy, and Rout (2014), several personnel associated with the APFAMGS programs expressed concern about the no-regulation no-sanction policy of the APFAMGS intervention. In the absence of a compulsory regulation, farmers are left on their own to decide whether or not they follow the procedures of the program.

Reddy, Reddy, and Rout (2014) argue that the gains from APFAMGS program are largely limited to large-scale farmers who own wells. The most recent long-term assessment of the APFAMGS program (Reddy and Reddy, 2020) suggests that treatment villages are more resilient to drought risks primarily because farmers in the treated region are more aware of the groundwater situation and sustainable management. The intervention has supposedly led to a significant increase in communities capacity to understand and apply crop diversification and improved irrigation practices. Importantly, the study suggests that an integrated approach that combines building knowledge capacity with economic incentives and regulation is required to make participatory hydrological monitoring.

In our study, we find that APFAMGS has very little effect on average groundwater levels, but had some success at changing the seasonal variation in groundwater levels. Specifically, we find that APFAMGS-treated areas experience less seasonal groundwater recharge after the seasonal rains and less groundwater draw-down after the cropping seasons than untreated areas. While we expected greater levels of recharge and less draw-down, we speculate that the consistent monitoring of groundwater levels removes information asymmetry on groundwater levels and encourages groundwater extraction in times of plenty and conservation in lean times.

We use well-level data to estimate the causal effect of the program, and we also use spatial interpolation (krigging) to expand our ability to measure the effect of the program. The kriging process allows us to observe a more balanced panel of treated wells, but ironically likely reduces our statistical power. In cross-validating the krig-predicted groundwater levels across the program area, we find an average measurement error of nearly 30%. While the measurement error does not affect the consistency of our causal estimates, it does affect the efficiency. The insignificant results we find using krigged data therefore should be interpreted cautiously: we are not confident that insignificant results are actually insignificant or just a type-II error.

Better data availability, particularly in groundwater levels and rainfall, would improve this work. Having more frequent and more spatially diverse measurements of groundwater would allow us to more accurately identify the causal effect of APFAMGS.

Using the best data available on groundwater levels in the region, we find no strong evidence that APFAMGS had any effect on average groundwater levels, but some sporadic success in decreasing seasonal variation in groundwater levels. Participatory hydrological monitoring programs such as APFAMGS have, by design, no real economic incentives to participate. While theory suggests that these types of programs should reduce individuals' groundwater extraction rates and thus increase groundwater reserves availability, we find that APFAMGS had modest effects and policy makers should be wary of extending costly participatory programs without stronger evidence of benefits.

5. SUMMARY AND CONCLUSION

This research estimates the benefits of improving soil and water quality characteristics, and the effect of a groundwater management program on groundwater levels. I find that anglers are willing to pay more for long-term changes in water quality than short-term changes. This intuitive result is novel and important as the use of panel data in recreational demand modeling becomes more popular. I find that farmers in central Texas value the increase in water infiltration much more than the increase in organic matter associated with the adoption of no-till. This result is robust to all sub-samples and groups of farmers in the area and may be a local phenomenon or perhaps a function of how soil quality characteristics are measured and communicated. Further research is needed to understand the reasons why there is such a divergence in the willingness-to-pay between the two measures.

I also find that APFAMGS - a participatory groundwater management program - had no effect on average groundwater levels, but upon further inspection did result in treated areas withdrawing more groundwater in times of plenty and less groundwater in lean times. Nonetheless, the effect was driven by a few select years and for most years the program had no effect even on these intra-year measurements of groundwater.

This research would be useful to policy makers interested in making good economic decisions. There is an entire class of non-market benefits (the value of changing long-term conditions) that I show can and should be included in the accounting of general costs and benefits when considering environmental policy.

The USDA Natural Resource Conservation Service (NRCS) works extensively with farmers to provide technical and financial assistance, and ultimately helps farmers decide whether or not to engage in conservation practices. While the costs of conservation are easy to calculate, benefits are less so. The second essay of this work provides some evidence that farmers are willing-to-pay for increases in some soil quality characteristics.

These essays, in their novelty, open more doors than they close. The first essay demands more attention on the difference between willingness-to-pay for short- and long-term changes. The results should be tested in a more data-rich setting, and across a variety of settings. The mechanism for the difference is likely the way anglers perceive water quality, which could be tested more directly to provide more support for the results. The second essay produces significantly different willingness-to-pay estimates between two characteristics, and further research is needed to identify the cause of these differences. If the estimated benefits are sensitive to the way information is presented to farmers, then more care should be taken in finding the best way (i.e. the best linking indicators) to communicate technical measurements to farmers who are making large financial decision. From the perspective of the USDA, the discrete choice experiment should be replicated in a variety of agricultural settings to understand the heterogeneity in value; farmers in Iowa may value very different soil attributes than farmers in Texas. Understanding how the benefits of soil improvement change across operations could help the USDA fine-tune their approach to technically and financially assisting farmers. Finally, the new participatory groundwater management program intended for Mahahrashtra in India should strive for richer data in monitoring the success or failure of the program. Understanding the pumping behavior of individual farmers, for example, could vastly improve our understanding of the effectiveness of these types of participatory programs.

In short, these essays should make regulatory decisions more complicated for policy makers.

REFERENCES

Aizaki, H. 2012. "Basic Functions For Supporting An Implementation Of Choice Experiments In R." *Journal Of Statistical Software* 50:1–24.

Akaike, H. 1998. "Information Theory And An Extension Of The Maximum Likelihood Principle." In *Selected Papers Of Hirotugu Akaike*. Springer, pp. 199–213.

Alabaster, J.S., and R.S. Lloyd. 2013. *Water Quality Criteria For Freshwater Fish*. Elsevier.

Allison, P.D. 2009. *Fixed Effects Regression Models*, vol. 160. SAGE publications.

Bagnall, D.K., W.A. McIntosh, C.L. Morgan, R.T. Woodward, M. Cisneros, M. Black, E.M. Kiella, and S. Ale. 2020. "Farmers Insights On Soil Health Indicators And Adoption." *Agrosystems, Geosciences & Environment* 3:e20066.

Bagnall, D.K., and C.L. Morgan. 2021. "SLAKES And 3D Scans Characterize Management Effects On Soil Structure In Farm Fields." *Soil And Tillage Research* 208:104893.

Bandeen-Roche, K., D.L. Miglioretti, S.L. Zeger, and P.J. Rathouz. 1997. "Latent Variable Regression For Multiple Discrete Outcomes." *Journal of the American Statistical Association* 92:1375–1386.

Barbier, E.B. 1990. "The Farm-Level Economics Of Soil Conservation: The Uplands Of Java." *Land Economics* 66:199–211.

Barrett, S. 1991. "Optimal Soil Conservation And The Reform Of Agricultural Pricing Policies." *Journal of Development Economics* 36:167–187.

Bazzani, C., M.A. Palma, and R.M. Nayga Jr. 2018. "On The Use Of Flexible Mixing Distributions In WTP Space: An Induced Value Choice Experiment." *Australian Journal of Agricultural and Resource Economics* 62:185–198.

Bliemer, M.C., and J.M. Rose. 2013. "Confidence Intervals Of Willingness-to-pay For Random Coefficient Logit Models." *Transportation Research Part B: Methodological* 58:199–214.

Boyd, J., P. Ringold, A. Krupnick, R. Johnson, M. Weber, and K.M. Hall. 2015. "Ecosystem

Services Indicators: Improving The Linkage Between Biophysical And Economic Analyses." *Resources For The Future Discussion Paper*, pp. 15–40.

Bünemann, E.K., G. Bongiorno, Z. Bai, R.E. Creamer, G. De Deyn, R. de Goede, L. Fleskens, V. Geissen, T.W. Kuyper, P. Mäder, et al. 2018. "Soil Quality–A Critical Review." *Soil Biology and Biochemistry* 120:105–125.

Burt, O.R. 1981. "Farm Level Economics Of Soil Conservation In The Palouse Area Of The Northwest." *American Journal Of Agricultural Economics* 63:83–92.

Calder, C., and N. Cressie. 2009. "Kriging and Variogram Models." In R. Kitchin and N. Thrift, eds. *International Encyclopedia of Human Geography*. Oxford: Elsevier, pp. 49–55.

Carson, R.T., and J.J. Louviere. 2011. "A Common Nomenclature For Stated Preference Elicitation Approaches." *Environmental And Resource Economics* 49:539–559.

Central Groundwater Board. 2011. "Ground Water Assessment." Working paper, Ministry of Water Resources, Government of India.

—. 2017. "Report Of The Ground Water Resource Estimation Committee." Working paper, Ministry of Water Resources, Government of India.

Chinedu, O., E. Sanou, J. Tur-Cardona, F. Bartolini, G. Gheysen, and S. Speelman. 2018. "Farmers Valuation Of Transgenic Biofortified Sorghum For Nutritional Improvement In Burkina Faso: A Latent Class Approach." *Food Policy* 79:132–140.

Ciriacy-Wantrup, S.V. 1947. "Capital Returns From Soil-conservation Practices." *Journal Of Farm Economics* 29:1181–1196.

Collins, L.M., and S.T. Lanza. 2009. *Latent Class And Latent Transition Analysis: With Applications In The Social, Behavioral, And Health Sciences*, vol. 718. John Wiley & Sons.

de Bekker-Grob, E.W., B. Donkers, M.F. Jonker, and E.A. Stolk. 2015. "Sample Size Requirements For Discrete-Choice Experiments In Healthcare: A Practical Guide." *The Patient-Patient-Centered Outcomes Research* 8:373–384.

Dominati, E., A. Mackay, S. Green, and M. Patterson. 2014. "A Soil Change-Based Methodology For The Quantification And Valuation Of Ecosystem Services From Agro-ecosystems: A Case Study Of Pastoral Agriculture In New Zealand." *Ecological Economics* 100:119–129.

Dundas, S.J., R.H. von Haefen, and C. Mansfield. 2018. "Recreation Costs Of Endangered Species Protection: Evidence From Cape Hatteras National Seashore." *Marine Resource Economics* 33:1–25.

FAO. 2010. "Andhra Pradesh Farmer Managed Groundwater Systems Project (APFAMGS Project)." Working paper, Food and Agriculture Organization of the United Nations.

—. 2013. "Smallholders And Sustainable Wells A Retrospect: Participatory Groundwater Management In Andhra Pradesh (India)." Working paper, Food and Agriculture Organization of the United Nations.

Garduño, H., S. Foster, P. Raj, and F. van Steenbergen. 2009. "Addressing Groundwater Depletion Through Community-Based Management Actions In The Weathered Granitic Basement Aquifer Of Drought-Prone Andhra Pradesh-India." *GW-MATE Case Profile Collection* 19.

Goodman-Bacon, A. 2018. "Difference-In-Differences With Variation In Treatment Timing." Working paper, National Bureau Of Economic Research.

Groothuis, P.A., T.M. Mohr, J.C. Whitehead, and K. Cockerill. 2017. "Endogenous Consequentiality in Stated Preference Referendum Data: The Influence of the Randomly Assigned Tax Amount." *Land Economics* 93:258–268.

Haab, T.C., and K.E. McConnell. 2002. *Valuing Environmental and Natural Resources : The Econometrics Of Non-market Valuation*. New Horizons in Environmental Economics, Cheltenham, U.K.: Edward Elgar Publishing.

Hanjra, M.A., and M.E. Qureshi. 2010. "Global Water Crisis And Future Food Security In An Era Of Climate Change." *Food Policy* 35:365–377.

Hanley, N., D. MacMillan, R.E. Wright, C. Bullock, I. Simpson, D. Parsisson, and B. Crabtree. 1998. “Contingent Valuation Versus Choice Experiments: Estimating The Benefits Of Environmentally Sensitive Areas In Scotland.” *Journal Of Agricultural Economics* 49:1–15.

Hansen, L., and D. Hellerstein. 2007. “The Value Of The Reservoir Services Gained With Soil Conservation.” *Land Economics* 83:285–301.

Hansen, L., and M. Ribaudo. 2008. “Economic Measures Of Soil Conservation Benefits: Regional Values For Policy Assessment.” *USDA Technical Bulletin* 1922.

Hensher, D.A., and J.M. Rose. 2007. “Development Of Commuter And Non-Commuter Mode Choice Models For The Assessment Of New Public Transport Infrastructure Projects: A Case Study.” *Transportation Research Part A: Policy and Practice* 41:428–443.

Hess, S., and J.M. Rose. 2009. “Should Reference Alternatives In Pivot Design SC Surveys Be Treated Differently?” *Environmental and Resource Economics* 42:297–317.

Hole, A.R. 2015. “DCREATE: Stata Module To Create Efficient Designs For Discrete Choice Experiments.” Statistical Software Components, Boston College Department Of Economics.

Holmes, T.P., W.L. Adamowicz, and F. Carlsson. 2017. “Choice Experiments.” In K. J. Boyle, T. C. Brown, and P. A. Champ, eds. *A Primer On Nonmarket Valuation*. Springer Netherlands, pp. 133–186.

Ji, Y., J.A. Herriges, and C.L. Kling. 2016. “Modeling Recreation Demand When The Access Point Is Unknown.” *American Journal Of Agricultural Economics* 98:860–880.

Ji, Y., D.A. Keiser, and C.L. Kling. 2020. “Temporal Reliability of Welfare Estimates from Revealed Preferences.” *Journal Of The Association Of Environmental And Resource Economists* 7:659–686.

Johnson, F.R., B. Kanninen, M. Bingham, and S. Özdemir. 2006. “Experimental Design For Stated-Choice Studies.” In *Valuing Environmental Amenities Using Stated Choice Studies*. Springer, pp. 159–202.

Johnston, R.J., K.J. Boyle, W. Adamowicz, J. Bennett, R. Brouwer, T.A. Cameron, W.M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, et al. 2017. “Contemporary Guidance For

Stated Preference Studies." *Journal Of The Association Of Environmental And Resource Economists* 4:319–405.

Jonsson, J.O.G., and B. Daviosdottir. 2016. "Classification And Valuation Of Soil Ecosystem Services." *Agricultural Systems* 145:24–38.

Joshi, D., H. Kulkarni, and U. Aslekar. 2019. "Bringing Aquifers And Communities Together: Decentralised Groundwater Governance In Rural India." In *Water Governance: Challenges And Prospects*. Springer, pp. 157–185.

Keiser, D.A. 2019. "The Missing Benefits Of Clean Water And The Role Of Mismeasured Pollution." *Journal Of The Association Of Environmental And Resource Economists* 6:669–707.

King, D.A., and J. Sinden. 1988. "Influence Of Soil Conservation On Farm Land Values." *Land Economics* 64:242–255.

Klaiber, H.A., and R.H. von Haefen. 2011. "Do Random Coefficients And Alternative Specific Constants Improve Policy Analysis? An Empirical Investigation Of Model Fit And Prediction." *Environmental And Resource Economics*, pp. 1–17.

Kulkarni, H., M. Shah, and P.V. Shankar. 2015. "Shaping The Contours Of Groundwater Governance In India." *Journal Of Hydrology: Regional Studies* 4:172–192.

Kuwayama, Y., S.M. Olmstead, D.C. Wietelman, and J. Zheng. 2020. "Trends In Nutrient-Related Pollution As A Source Of Potential Water Quality Damages: A Case Study Of Texas, USA." *Science Of The Total Environment* 724:137962.

Lal, R. 2014. "Societal Value Of Soil Carbon." *Journal Of Soil And Water Conservation* 69:186A–192A.

Lee, L.K. 1980. "The Impact Of Landownership Factors On Soil Conservation." *American Journal Of Agricultural Economics* 62:1070–1076.

Linzer, D.A., J.B. Lewis, et al. 2011. "poLCA: An R Package For Polytomous Variable Latent Class Analysis." *Journal Of Statistical Software* 42:1–29.

Matheron, G. 1963. "Principles Of Geostatistics." *Economic Geology* 58:1246–1266.

Mazurkiewicz, S.M., K.J. Boyle, M.F. Teisl, K.I. Morris, and A.G. Clark. 1996. "Recall Bias And Reliability Of Survey Data: Moose Hunting In Maine." *Wildlife Society Bulletin (1973-2006)* 24:140–148.

McConnell, K.E. 1983. "An Economic Model Of Soil Conservation." *American Journal Of Agricultural Economics* 65:83–89.

McFadden, D. 1974. "Conditional Logit Analysis Of Qualitative Choice Behavior." *Frontiers In Econometrics*, pp. 105–142.

McFadden, D., and K. Train. 2000. "Mixed MNL Models For Discrete Response." *Journal Of Applied Econometrics* 15:447–470.

Melstrom, R.T., F. Lupi, P.C. Esselman, and R.J. Stevenson. 2015. "Valuing Recreational Fishing Quality At Rivers And Streams." *Water Resources Research* 51:140–150.

Miller, R.L., W.L. Bradford, and N.E. Peters. 1988. *Specific Conductance: Theoretical Considerations And Application To Analytical Quality Control*. US Government Printing Office.

Moeltner, K., and R. Von Haefen. 2011. "Microeconometric Strategies For Dealing With Unobservables And Endogenous Variables In Recreation Demand Models." *Annual Review Of Resource Economics* 3:375–396.

Moriasi, D.N., L.F. Duriancik, E.J. Sadler, T. Tsegaye, J.L. Steiner, M.A. Locke, T.C. Strickland, and D.L. Osmond. 2020. "Quantifying The Impacts Of The Conservation Effects Assessment Project Watershed Assessments: The First Fifteen Years." *Journal of Soil and Water Conservation* 75:57A–74A.

Murdock, J. 2006. "Handling Unobserved Site Characteristics In Random Utility Models Of Recreation Demand." *Journal Of Environmental Economics And Management* 51:1–25.

Murphy, J.J., P.G. Allen, T.H. Stevens, and D. Weatherhead. 2005. "A Meta-Analysis Of Hypothetical Bias In Stated Preference Valuation." *Environmental And Resource Economics* 30:313–325.

Parsons, G.R., A.K. Kang, C.G. Leggett, and K.J. Boyle. 2009. "Valuing Beach Closures On

The Padre Island National Seashore." *Marine Resource Economics* 24:213–235.

Pendleton, L.H., and R. Mendelsohn. 1998. "Estimating The Economic Impact Of Climate Change On The Freshwater Sportsfisheries Of The Northeastern US." *Land Economics*, pp. 483–496.

Plaas, E., F. Meyer-Wolfarth, M. Banse, J. Bengtsson, H. Bergmann, J. Faber, M. Potthoff, T. Runge, S. Schrader, and A. Taylor. 2019. "Towards Valuation Of Biodiversity In Agricultural Soils: A Case For Earthworms." *Ecological Economics* 159:291–300.

Plümper, T., and V.E. Troeger. 2007. "Efficient Estimation Of Time-Invariant And Rarely Changing Variables In Finite Sample Panel Analyses With Unit Fixed Effects." *Political Analysis*, pp. 124–139.

Reddy, V.R., and M.S. Reddy. 2020. "Can Participatory Groundwater Management Enhance Drought Resilience? The Case of the Andhra Pradesh Farmer-Managed Groundwater Systems Project." Working paper, The World Bank.

Reddy, V.R., M.S. Reddy, and S.K. Rout. 2014. "Groundwater Governance: A Tale Of Three Participatory Models In Andhra Pradesh, India." *Water Alternatives* 7.

Rose, J.M., and M.C. Bliemer. 2009. "Constructing Efficient Stated Choice Experimental Designs." *Transport Reviews* 29:587–617.

Scarpa, R., D. Campbell, and W.G. Hutchinson. 2007. "Benefit Estimates For Landscape Improvements: Sequential Bayesian Design And Respondents Rationality In A Choice Experiment." *Land Economics* 83:617–634.

Scarpa, R., C. Franceschinis, and M. Thiene. 2021. "Logit Mixed Logit Under Asymmetry And Multimodality Of WTP: A Monte Carlo Evaluation." *American Journal of Agricultural Economics* 103:643–662.

Schwarz, G., et al. 1978. "Estimating The Dimension Of A Model." *The Annals Of Statistics* 6:461–464.

Seitz, W.D., and E.R. Swanson. 1980. "Economics Of Soil Conservation From The Farmer's Perspective." *American Journal of Agricultural Economics* 62:1084–1088.

Sekhri, S. 2014. "Wells, Water, And Welfare: The Impact Of Access To Groundwater On Rural Poverty And Conflict." *American Economic Journal: Applied Economics* 6:76–102.

Smith, D., R. Harmel, and R. Haney. 2020. "Long-Term Agro-Economic And Environmental Assessment Of Adaptive Nutrient Management On Cropland Fields With Established Structural Conservation Practices." *Journal Of Soil And Water Conservation* 75:416–425.

Soil Health Institute. 2021. "Economics of Soil Health Systems in Iowa." Working paper, Soil Health Institute.

Train, K. 2016. "Mixed Logit With A Flexible Mixing Distribution." *Journal Of Choice Modelling* 19:40–53.

Train, K., and W.W. Wilson. 2008. "Estimation On Stated-Preference Experiments Constructed From Revealed-Preference Choices." *Transportation Research Part B: Methodological* 42:191–203.

Unger, P.W., and T.C. Kaspar. 1994. "Soil compaction and root growth: a review." *Agronomy Journal* 86:759–766.

United States Department of Agriculture, Natural Resources Conservation Service. 2015. "Illustrated Guide to Soil Taxonomy." Working paper, United States Department of Agriculture.

—. 2006. "Land Resource Regions and Major Land Resource Areas of the United States, the Caribbean, and the Pacific Basin." Working paper, U.S. Department of Agriculture Handbook 296.

United States Fish and Wildlife Service. 2001. "National Survey Of Fishing, Hunting, And Wildlife-Associated Recreation." Report.

—. 2006. "National Survey Of Fishing, Hunting, And Wildlife-Associated Recreation." Report.

—. 2011. "National Survey Of Fishing, Hunting, And Wildlife-Associated Recreation." Report.

Van Steenbergen, F. 2006. "Promoting Local Management In Groundwater." *Hydrogeology*

Journal 14:380–391.

Verma, S., S. Krishnan, A. Reddy, and K. Reddy. 2012. "Andhra Pradesh [India] Farmer Managed Groundwater Systems (APFAMGS): A Reality Check." *IWMI-Tata Water Policy Research Highlight* 37.

Von Haefen, R.H., and D.J. Phaneuf. 2008. "Identifying Demand Parameters In The Presence Of Unobservables: A Combined Revealed And Stated Preference Approach." *Journal Of Environmental Economics And Management* 56:19–32.

Whitehead, J.C., T. Haab, S.L. Larkin, J.B. Loomis, S. Alvarez, and A. Ropicki. 2018. "Estimating Lost Recreational Use Values Of Visitors To Northwest Florida Due To The Deepwater Horizon Oil Spill Using Cancelled Trip Data." *Marine Resource Economics* 33:119–132.

Wooldridge, J.M. 2010. *Econometric Analysis Of Cross Section And Panel Data*. MIT Press.

World Bank. 2018. "Program Appraisal Document On A Proposed Loan In The Amount US $450 Million To The Republic Of India For An Atal Bhujal Yojana (ABHY) National Groundwater Management Improvement Program." Working paper, Water Global Practice, South Asia Region, World Bank.

Yi, D., and J.A. Herriges. 2017. "Convergent Validity And The Time Consistency Of Preferences: Evidence From The Iowa Lakes Recreation Demand Project." *Land Economics* 93:269–291.

Zaveri, E., D.S. Grogan, K. Fisher-Vanden, S. Frolking, R.B. Lammers, D.H. Wrenn, A. Prusevich, and R.E. Nicholas. 2016. "Invisible Water, Visible Impact: Groundwater Use And Indian Agriculture Under Climate Change." *Environmental Research Letters* 11:084005.

APPENDIX A

APPENDIX FOR CHAPTER 1

Table A.1: Alternate Specifications: First Stage Using "Most-Often"

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Travel Cost	-0.0116***	-0.0153***	-0.0244***	-0.0116***	-0.0118***	-0.0106***	-0.0114***
	(0.000571)	(0.000772)	(0.000969)	(0.000576)	(0.000615)	(0.00172)	(0.000626)
Dissolved Oxygen	-0.113	-0.114	-0.0610	-0.129*	-0.0965	-0.0898	-0.0209
	(0.0776)	(0.0775)	(0.0959)	(0.0775)	(0.0862)	(0.217)	(0.0914)
pH	1.040***	1.067***	0.764**	1.052***	0.905***	1.374	0.813***
	(0.255)	(0.253)	(0.321)	(0.256)	(0.273)	(1.002)	(0.291)
Specific Conductance	-0.293**	-0.299**	-0.488**	-0.294**	0.161	-0.454	-0.312**
	(0.131)	(0.129)	(0.218)	(0.132)	(0.254)	(0.348)	(0.141)
Transparency	0.195	0.235	0.255	0.204	0.0607	2.452***	0.256
	(0.193)	(0.194)	(0.247)	(0.194)	(0.201)	(0.731)	(0.212)
Observations	179691	179691	41742	177730	117450	4629	146537
AIC	6223.3	6156.3	3517.4	6162.8	5127.9	424.3	5142.5
Log-likelihood	-2903.6	-2870.1	-1550.7	-2873.4	-2430.0	-135.2	-2363.2

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

Notes: Standard errors in parentheses. Specifications/columns:

1. *Most-often base model*
2. *Alternate travel cost specification*
3. *Distance of $\leq$ 150 miles*
4. *Only including anglers who reported 1 trip or more*
5. *Lakes only*
6. *Rivers only*
7. *Drop 2001 survey*

Table A.2: Alternate Specifications: First Stage Using "Typical-Trip"

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Travel Cost	-0.00609***	-0.00792***	-0.0163***	-0.00633***	-0.00605***	-0.00697***	-0.00934***	-0.00812***	-0.00745***
	(0.000222)	(0.000288)	(0.000648)	(0.000235)	(0.000233)	(0.000837)	(0.000442)	(0.000441)	(0.000530)
Dissolved Oxygen	0.102*	0.102*	-0.0170	0.0550	0.137**	0.272**	-0.0865	0.133	-0.101
	(0.0586)	(0.0588)	(0.0768)	(0.0607)	(0.0654)	(0.138)	(0.0736)	(0.0818)	(0.0702)
pH	0.537***	0.554***	0.505*	0.751***	0.585***	-1.063	0.434*	0.458	0.494**
	(0.204)	(0.203)	(0.260)	(0.210)	(0.223)	(0.819)	(0.256)	(0.285)	(0.244)
Specific Conductance	-0.311***	-0.315***	-0.405***	-0.341***	0.116	-0.523***	-0.207*	-0.123	-0.309***
	(0.107)	(0.107)	(0.155)	(0.115)	(0.179)	(0.200)	(0.115)	(0.152)	(0.118)
Transparency	0.316***	0.331***	-0.151	0.276**	0.317***	0.646	0.125	0.191	0.113
	(0.112)	(0.112)	(0.219)	(0.111)	(0.115)	(0.418)	(0.164)	(0.177)	(0.159)
Observations	262103	262103	50981	244541	170402	7478	192926	133043	68824
AIC	11978.2	11906.3	5394.1	11035.3	10061.8	821.7	7380.1	5742.0	7013.5
Log-likelihood	-5781.1	-5745.2	-2489.0	-5309.7	-4896.9	-332.8	-3482.1	-2663.0	-3298.8

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

Notes: Standard errors in parentheses. Specifications/columns:

1. *Base Model*
2. *Alternate travel cost specification*
3. *Distance of ≤ 150 miles*
4. *Only including anglers who reported 1 trip or more*
5. *Lakes only*
6. *Rivers only*
7. *Drop 2001 survey*
8. *Day-trips only*
9. *Distance error (calculated - reported) of ≤ 100 miles*

Table A.3: Alternate Specifications: Second Stage Using "Typical-Trip"

Panel A: Without Regional Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dissolved Oxygen	-0.0486	-0.0419	-0.0234	0.131	0.320	-0.265	0.0516	0.232	-0.0204
	(0.551)	(0.559)	(0.566)	(0.525)	(0.790)	(0.711)	(0.544)	(0.486)	(0.567)
pH	4.473**	4.558**	3.737*	4.137**	1.022	5.630*	3.265*	6.041***	3.553*
	(1.908)	(1.941)	(1.922)	(1.878)	(2.886)	(3.054)	(1.915)	(2.046)	(1.894)
Specific Conductance	-1.788**	-1.793**	-1.689**	-1.778***	-0.652	-2.057***	-1.619**	-2.123***	-1.784**
	(0.692)	(0.698)	(0.702)	(0.667)	(1.068)	(0.729)	(0.646)	(0.721)	(0.691)
Transparency	3.042***	3.065**	3.455***	3.033***	1.775	4.154**	2.904***	2.429*	3.166***
	(1.166)	(1.184)	(1.229)	(1.131)	(1.373)	(2.027)	(1.100)	(1.249)	(1.143)
Constant	-33.08**	-33.91**	-27.18**	-32.01**	-13.88	-39.41*	-24.46*	-46.90***	-30.00**
	(13.02)	(13.27)	(12.88)	(12.79)	(18.80)	(21.38)	(13.31)	(13.72)	(13.20)
Observations	198	198	197	198	126	72	198	198	198
R^2	0.089	0.088	0.085	0.090	0.020	0.128	0.081	0.086	0.085
Adjusted R^2	0.070	0.069	0.066	0.071	-0.012	0.076	0.061	0.067	0.066

Panel B: With Regional Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dissolved Oxygen	-0.0924	-0.0879	-0.302	0.0594	0.490	-0.365	-0.186	-0.148	0.0774
	(0.650)	(0.660)	(0.688)	(0.611)	(0.792)	(0.875)	(0.654)	(0.569)	(0.655)
pH	5.841*	5.911*	7.105**	5.916**	4.727	1.248	7.297**	10.32***	5.540*
	(2.984)	(3.035)	(3.104)	(2.965)	(4.921)	(4.140)	(3.165)	(3.192)	(2.931)
Specific Conductance	-1.677*	-1.673*	-1.206	-1.619*	-1.179	-0.308	-1.086	-1.458	-2.017**
	(0.986)	(0.993)	(1.022)	(0.973)	(1.377)	(1.396)	(0.896)	(1.054)	(0.958)
Transparency	2.826**	2.845**	3.105**	2.795**	3.359**	-0.0691	2.816**	2.129	2.620**
	(1.259)	(1.277)	(1.345)	(1.237)	(1.526)	(2.619)	(1.247)	(1.303)	(1.225)
Constant	-50.19**	-50.76**	-50.17**	-52.77**	-50.76	-20.45	-65.24**	-90.24***	-62.92**
	(25.31)	(25.63)	(25.30)	(24.98)	(36.76)	(33.38)	(27.28)	(27.00)	(26.49)
Observations	193	193	192	193	121	72	193	193	193
R^2	0.157	0.156	0.151	0.155	0.179	0.360	0.163	0.173	0.174
Adjusted R^2	0.048	0.047	0.040	0.046	0.024	0.091	0.054	0.066	0.068

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

Notes: Standard errors in parentheses. Specifications/columns:
1: Typical-trip base model
2: Alternate first stage travel cost
3: Distance of ≤ 150 miles
4: Only including anglers who reported 1 trip or more
5: Lakes only
6: Rivers only
7: Drop 2001 survey
8: Day-trips only
9: Distance error (calculated - reported) of ≤ 100 miles
Regional fixed-effects control for a site's major river basin.

Table A.4: Alternate Specifications: Second Stage Using "Most-Often"

Panel A: Without Regional Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dissolved Oxygen	0.0628	0.0669	0.265	0.129	0.102	1.039	-0.00116
	(0.513)	(0.548)	(0.530)	(0.507)	(0.820)	(0.648)	(0.587)
pH	2.672	2.914	2.665	2.558	1.636	-2.900	3.714
	(1.977)	(2.126)	(2.083)	(1.989)	(3.067)	(2.734)	(2.254)
Specific Conductance	-1.956***	-2.044***	-1.887***	-1.973***	-1.466	-1.281	-1.656**
	(0.689)	(0.733)	(0.718)	(0.683)	(1.166)	(0.805)	(0.751)
Transparency	2.945**	3.082**	4.069***	2.732**	2.477*	-1.781	3.206**
	(1.220)	(1.300)	(1.303)	(1.203)	(1.466)	(2.155)	(1.369)
Constant	-19.31	-21.43	-22.08	-18.79	-12.75	10.16	-30.39*
	(13.44)	(14.48)	(14.00)	(13.48)	(20.02)	(19.62)	(15.58)
Observations	198	198	197	198	126	71	198
R^2	0.079	0.075	0.097	0.075	0.043	0.068	0.068
Adjusted R^2	0.059	0.056	0.078	0.056	0.012	0.011	0.049

Panel B: With Regional Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dissolved Oxygen	-0.106	-0.116	-0.0599	-0.0757	0.156	1.501*	-0.408
	(0.596)	(0.637)	(0.624)	(0.590)	(0.744)	(0.808)	(0.708)
pH	6.308**	6.721**	7.032**	6.440**	8.344*	-5.750	9.388**
	(3.056)	(3.284)	(3.219)	(3.047)	(4.976)	(4.585)	(3.702)
Specific Conductance	-1.991**	-2.065**	-1.922*	-1.916**	-2.398	-1.847	-0.982
	(0.955)	(1.014)	(1.041)	(0.967)	(1.463)	(1.765)	(1.068)
Transparency	3.415**	3.582**	3.915***	3.162**	4.168***	-3.816	3.453**
	(1.319)	(1.403)	(1.372)	(1.308)	(1.555)	(3.614)	(1.509)
Constant	-54.04**	-57.83**	-60.88**	-55.79**	-72.21*	30.63	-85.15***
	(25.55)	(27.38)	(26.82)	(25.53)	(37.77)	(36.90)	(31.44)
Observations	193	193	192	193	121	71	193
R^2	0.162	0.159	0.207	0.154	0.237	0.260	0.150
Adjusted R^2	0.053	0.050	0.104	0.044	0.094	-0.057	0.040

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

Notes: Standard errors in parentheses. Specifications/columns:

1: Most-often base model

2: Alternate first stage travel cost

3: Distance of $\leq$ 150 miles

4: Only including anglers who reported 1 trip or more

5: Lakes only

6: Rivers only

7: Drop 2001 survey

Regional fixed-effects control for a site's major river basin.

APPENDIX B

APPENDIX FOR CHAPTER 2

B.1 Full Willingness-to-Pay Estimates

Table B.1: Marginal Willingness-To-Pay By Model/Group: Conditional Logit

Model/Group	Water Infiltration	Organic Matter	High Compaction[a]	Medium Compaction[a]
Base Model	58.40	0.79	-88.26	-28.60
	(42.27, 74.53)	(0.57, 1.02)	(-110.78, -65.75)	(-37.66, -19.55)
Low Rental Rate	41.23	0.64	-72.02	-18.70
	(21.33, 61.12)	(0.34, 0.94)	(-101.54, -42.51)	(-29.95, -7.44)
Medium Rental Rate	50.06	0.52	-54.87	-19.72
	(29.49, 70.63)	(0.26, 0.78)	(-76.93, -32.81)	(-31.15, -8.28)
High Rental Rate	94.35	1	-110.87	-32.55
	(29.31, 159.38)	(0.23, 1.77)	(-189.58, -32.17)	(-60.2, -4.9)
Conventional-till	49.62	0.54	-70.20	-19.99
	(33.75, 65.49)	(0.34, 0.74)	(-90.47, -49.93)	(-28.38, -11.6)
No-till/Strip-till	104.09	1.94	-152.25	-53.26
	(16.72, 191.46)	(0.32, 3.56)	(-281.6, -22.89)	(-101.47, -5.04)
Future Operator: Self	60.95	0.75	-75.23	-29.30
	(40.21, 81.68)	(0.47, 1.03)	(-100.48, -49.98)	(-40.89, -17.72)
Future Operator: Someone Else	36.45	0.55	-78.65	-19.08
	(13.05, 59.85)	(0.23, 0.87)	(-114.18, -43.12)	(-32.49, -5.67)
Future Operator: Uncertain	42.39	0.69	-93.65	-25.94
	(11.58, 73.19)	(0.27, 1.11)	(-144.46, -42.85)	(-43.89, -7.98)
Topography: Flat	88.45	1.30	-133.33	-47.33
	(28.28, 148.62)	(0.37, 2.23)	(-225.42, -41.24)	(-62.61, -12.05)
Topography: Sloped	52.05	0.69	-79.63	-23.36
	(34.86, 69.24)	(0.46, 0.92)	(-102.96, -56.29)	(-32.69, -14.04)
Soil-Apathetic Class	85.47	0.63	-79.95	-31.70
	(27.76, 143.17)	(0.07, 1.18)	(-139.77, -20.14)	(-58.2, -5.21)
Soil-Conscious Class	111.97	1.48	-141.52	-56.74
	(14.98, 208.97)	(0.17, 2.79)	(-263.59, -19.45)	(-109.19, -4.28)
Moderate Class	56.13	0.58	-69.34	-20.61
	(33.85, 78.41)	(0.31, 0.86)	(-96.43, -42.25)	(-32.1, -9.12)
Uninformed Class	33.62	0.94	-82.64	-21.36
	(3.46, 63.79)	(0.37, 1.52)	(-132.25, -33.03)	(-39.52, -3.21)

Note: 95% confidence intervals in parentheses.
[a] *Willingness-to-pay to move to low compaction*

Table B.2: Marginal Willingness-To-Pay By Model/Group: Mixed Logit

Model/Group	Water Infiltration	Organic Matter	High Compaction[a]	Medium Compaction[a]
Base Model	73.49	0.72	-78.84	-28.19
	(-52.90, 199.89)	(-0.13, 1.56)	(-177.05, 19.38)	(-40.18, -16.19)
Low Rental Rate	52.94	0.59	-64.91	-19.65
	(-43.73, 149.61)	(-0.04, 1.22)	(-156.9, 27.08)	(-33.99, -5.31)
Medium Rental Rate	56.44	0.50	-54.30	-20.23
	(-34.61, 147.49)	(-0.16, 1.15)	(-121.7, 13.11)	(-38.34, -2.13)
High Rental Rate	100.30	0.83	-89.00	-29.56
	(-71.68, 272.29)	(-0.78, 2.44)	(-236.78, 58.79)	(-56.40, -2.71)
Conventional-till	59.96	0.52	-63.04	-22.18
	(-59.45, 179.36)	(0.09, 0.95)	(-146.47, 20.39)	(-32.6, -11.76)
No-till/Strip-till	80.02	1.29	-111.45	-40.31
	(-103.37, 263.41)	(-0.41, 3)	(-246.09, 23.19)	(-81.89, 1.26)
Future Operator: Self	67.44	0.71	-66.59	-27.50
	(-62.81, 197.70)	(-0.19, 1.61)	(-154.55, 21.37)	(-40.22, -14.77)
Future Operator: Someone Else	41.74	0.56	-73.56	-19.63
	(-54.86, 138.34)	(-0.25, 1.36)	(-161.04, 13.91)	(-36.51, -2.75)
Future Operator: Uncertain	83.47	0.57	-88.91	-30.64
	(-67.43, 234.36)	(-0.61, 1.74)	(-198.05, 20.23)	(-68.31, 7.03)
Topography: Flat	105.70	1.20	-96.20	-41.13
	(-115.76, 327.17)	(-0.55, 2.96)	(-250.21, 57.80)	(-92.98, 10.72)
Topography: Sloped	63.68	0.63	-73.63	-24.92
	(-49.68, 177.04)	(-0.15, 1.40)	(-165.84, 18.59)	(-36.29, -13.54)
Soil-Apathetic Class	85.86	0.58	-69.16	-31.10
	(-63.77, 235.49)	(0.00, 1.16)	(-157.11, 18.8)	(-55.91, -6.3)
Soil-Conscious Class	99.00	1.07	-107.25	-45.02
	(-76.43, 274.43)	(-0.45, 2.60)	(-242.34, 27.84)	(-78.97, -11.08)
Moderate Class	64.94	0.53	-71.33	-22.48
	(-41.90, 171.77)	(-0.45, 1.50)	(-159.84, 17.19)	(-36.60, -8.36)
Uninformed Class	45.63	0.77	-74.64	-23.90
	(-88.4, 179.66)	(-0.62, 2.17)	(-177.12, 27.84)	(-44.26, -3.54)

Note: 95% confidence intervals in parentheses.
[a] *Willingness-to-pay to move to low compaction*

B.2 Survey Refinement

The initial survey instrument was developed in late 2018 and early 2019 using information from focus groups and cognitive interviews with local farmers and students. Pilot physical surveys were distributed on June 18^{th} and June 20^{th}, 2019, and the online survey was distributed via *Qualtrics*[1] beginning on August 9^{th}, 2019. Across both surveys, we received 42 usable responses (14 from the physical survey, 28 from the online survey). Select summary statistics are presented in Table B.3. The average respondent farmer from the pilot sample is approximately 50 years old, and has approximately 25 years of experience in farming. Most farmers manage more rented acres than owned acres.

Table B.3: Pilot Survey Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	42	50.4	16.3	23	36	62.5	84
Years Farming	42	24.9	17.1	4	10	41.5	57
Acres Owned	39	957.1	1,826.0	0.0	130.0	670.0	7,995.0
Acres Rented	35	2,238.8	2,394.8	0.0	315.0	3,925.0	8,000.0

Farmers in the pilot sample engage in a variety of tillage practices. Most (approximately 43%) are engaged in neither strip- nor no-till practices. Approximately 23% of respondents were engaged in strip-till only, and the same proportion was engaged in no-till only. The remaining 11% have used both strip- and no-till on their operation. Using the pilot survey responses, we estimate a basic conditional logit model. Specifically, we assume the indirect utility V_{ijt} associated with respondent i choosing field j at time t is defined as:

$$V_{ijt} = \alpha_j + X'_{jt}\beta, \tag{B.1}$$

[1] https://www.qualtrics.com

where α_j is a dummy variable equal to one if alternative j is Field A or Field B, and zero if alternative j is the opt-out option. The matrix X_{jt} contains information for the four chosen attributes of the choice experiment: water infiltration, organic matter, compaction, and rental rate. Compaction is the only categorical variable, with three levels: high, medium, and low. During the estimation process, the *low* compaction level is set as the base (omitted) category. Results of the pilot survey conditional logit are presented in Table B.4. The sign of each estimated coefficient is as expected; as water infiltration and organic matter increase for a field, the probability of selecting that field increases, *ceteris paribus*. High and medium compacted fields are less likely to be chosen compared to to low compacted field, *ceteris paribus*. Finally, as the rental rate of a field increases, the probability of selected that field decreases.

Table B.4: Pilot Survey Conditional Logit Estimation Results

	(1)
ASC	0.308
	(0.489)
Water Infiltration	8.304***
	(1.470)
Organic Matter	0.398***
	(0.098)
High Compaction[a]	−1.294***
	(0.239)
Medium Compaction[a]	−0.280*
	(0.165)
Price	−0.012***
	(0.005)
AIC	609.4
Observations	1,084
Log Likelihood	−298.688
Note:	*p<0.1; **p<0.05; ***p<0.01

[a] *Omitted base category: low compaction*

After estimation of the conditional logit, take several steps to improve the final survey design, First, we suspected our pivot on the price variable was too small. In the new design, we increase the size of the pivot from \$5 to \$10. Second, we use the estimated coefficients in the pilot study as priors, and develop a D-optimal survey design. We use the *dcreate* package in Stata 14.1 (Hole, 2015) which maximizes the D-efficiency of the survey design. The D-efficiency of a survey design is:

$$D = \left[\left| \left[\sum_{n=1}^{N} \sum_{j=1}^{J} z'_{nj} P_{nj} z_{nj} \right]^{-1} \right|^{1/K} \right]^{-1}, \tag{B.2}$$

where n refers to a respondent, j refers to an alternative, K is the number of explanatory variables, and

$$P_{nj} = \frac{e^{x'_{nj}\beta}}{\sum_{j \in J} e^{x'_{nj}\beta}}, \tag{B.3}$$

which, from the conditional logit, is the probability that respondent n chooses alternative j, and

$$z_{nj} = x_{nj} - \sum_{j \in J} x_{nj} P_{nj}. \tag{B.4}$$

Calculating the D-efficiency of a survey design requires some prior knowledge of β, and is essentially how efficiently a survey design can identify the prior vector of coefficients. By substituting our pilot study $\hat{\beta}$ for β in equation B.2, we find a matrix x_{nj} of size $(N \times J)$ by K of alternative levels combinations that maximizes the D-efficiency measure.

In maximizing the D-efficiency of the survey design, we also iterate over designs until we avoid conditions 1 and 2 from above. Further, we iterate until we find a design that has one dominated alternative to check for preference consistency, as described above. We thus create a D-efficient design that is not compromised by unrealistic alternative combinations and allows us to filter out respondents that are likely using some heuristic decision rule rather than making true utility-maximizing decisions.

Finally, we perform power calculations to determine our target sample size. Specifically, we find the minimum sample size N such that:

$$N > \left((z_{1-\beta} + z_{1-\alpha}) \sqrt{\sum_{\gamma k}} / \delta \right)^2 \tag{B.5}$$

where $\sum_{\gamma k}$ is the kth element of the diagonal of the variance-covariance matrix σ_γ of prior estimates δ. The statistical power can be set as $1 - \beta$, the significance level can be set at α, and then the $z-$score can be calculated from the normal distribution. This approach to finding the required sample size is specific to discrete choice experiments, and is outlined in de Bekker-Grob et al. (2015). This approach finds the minimum required sample size to identify each coefficient. Thus, for a range of β and α, we find the minimum sample size required for each coefficient. Results when $\alpha = 0.05$ and $1 - \beta = 0.8$ are shown in Table B.5. The limiting variables that may hinder identification are the generic ASC and medium compaction. While the ASC may be useful, it is not integral to our research question. The primary coefficients of interest - water infiltration, organic matter, and rental rate - need at most 18 individuals for identification. Across all the subsamples in our analysis, we have more than 18 respondents and we are therefore confident of our estimated coefficients.

Table B.5: Required Sample Sizes To Identify Effects

α	$1-\beta$	ASC	Water Infiltration	Organic Matter	Compaction: High	Compaction: Medium	Rental Rate
0.05	0.8	91.55	8.13	17.82	9.30	115.08	10.66

B.3 Survey Instrument

The final survey instrument deployed by the National Agricultural Statistics Service in 2020 to farmers in the Brazos River Watershed is provided below.

Your Farm. Your Soil.

OMB No.0535-0264
Approval Expires: 4/30/2022
Project Code: 779
Survey ID: 1980
Version 48

United States Department of Agriculture

NATIONAL AGRICULTURAL STATISTICS SERVICE

TEXAS A&M UNIVERSITY

USDA/NASS - Texas
Southern Plains Region
PO Box 70
Austin, TX 78767-0070
Phone: 1-800-626-3142
Fax: 1-855-270-2725
Email: nassrfospr@usda.gov

Please make corrections to name, address, and ZIP Code, if necessary.

Why am I being asked to participate in this survey?

Soil health is critical to farm profitability and is a national concern. Yet, there is great uncertainty about farmers' perceptions regarding soil health. How do you assess the health of the soil on your farm? **How important is it to you?** What do you do to manage for soil health?

The purpose of this survey is to provide data for the first time about how farmers in your area think about soil health. The results of this survey will help researchers understand farmers' thoughts on soil health, which in turn will be used to provide guidance to policy makers on the best ways to achieve state and national soil-health goals. Your participation is important and greatly appreciated.

If you have specific questions regarding the content of this survey, please contact Richard Woodward, Professor, Dept. of Agricultural Economics, Texas A&M University, 979-845-5864, r-woodward@tamu.edu

Section 1: You and your operation		
Q1	How many years have you been farming?	101 ____________ years
Q2	How old are you?	102 ____________ years
Q3	Five years from now, which of the following do you think will be most likely?	103 1 ☐ I will still be operating the farm. 2 ☐ The farm will be operated by one or more relatives (children or other relatives). 3 ☐ The farm will be operated by a non-related farmer. 4 ☐ The farm will be converted into non-farm use. 5 ☐ Do not know.
Q4	When you eventually stop farming, which of the following do you think will be most likely?	104 1 ☐ The farm will be operated by one or more relatives (children or other relatives). 2 ☐ The farm will be operated by a non-related farmer. 3 ☐ The farm will be converted into non-farm use. 4 ☐ Do not know.
Q5	Did/do your parents farm? *(If No, skip to Q6)*	105 1 ☐ Yes 3 ☐ No
	a. Are they still farming?	106 1 ☐ Yes 3 ☐ No, stop farming 6 ☐ No, deceased
	b. Do you currently work with them?	107 1 ☐ Yes 3 ☐ No
	c. Have you ever worked with them?	108 1 ☐ Yes 3 ☐ No
Q6	Roughly, what share of your household income comes from farming?	109 1 ☐ 100% 2 ☐ 75% 3 ☐ 50% 4 ☐ 25% or less
Q7	Roughly, what percent of your working time is dedicated to farming?	110 1 ☐ 100% 2 ☐ 75% 3 ☐ 50% 4 ☐ 25% or less
Q8	Total acreage under management	111 _____ acres owned 112 _____ acres rented
Q9	Total acreage under management	113 _____ acres in row crops 114 _____ acres in pasture

Section 1: You and your operation (continued)				
Q10 In 2019, how many acres did you plant in the following row crops?				
a. Corn	115 ________ acres			
b. Soybean	116 ________ acres			
c. Wheat	117 ________ acres			
d. Cotton	118 ________ acres			
e. Grain Sorghum	119 ________ acres			
f. Other	120 ________ acres			
Q11 In 2019, did you use strip-till on any of the row crops that you manage?	121	1 ☐ Yes	3 ☐ No	
Q12 Do you intend to increase or decrease your use of strip-till in the future?	122	1 ☐ Increase	2 ☐ Decrease	3 ☐ No Change
Q13 In 2019, did you use no-till on any of the row crops that you manage?	123	1 ☐ Yes	3 ☐ No	
Q14 Do you intend to increase or decrease your use of no-till in the future?	124	1 ☐ Increase	2 ☐ Decrease	3 ☐ No Change
Q15 To what extent do you rely on a consultant, like an agronomist or entomologist, to help you make farm management decisions?	125	1 ☐ Extensively 3 ☐ Very Little	2 ☐ Somewhat 4 ☐ None; I do not use a consultant	

Section 2: Details on Base Field

We would like you to give details on 1 row-crop field that you manage. By "field," we mean an area that you manage as one piece in terms of tillage, planting, and harvesting. Please choose a field that you know well.

***If you do not use no-till or strip-till**, please choose a field with which you would be comfortable experimenting with alternative management practices.*

***If you do use no-till or strip-till**, please choose a field on which you might consider switching to conventional tillage.*

Give this field a name so it will be easy to remember (for example, "Johnson")

Q16	What county is this field located in?	201 _________________ county
Q17	How many acres are in this field?	202 _________________ acres
Q18	How many years have you been managing this field?	203 1 ☐ 0 – 5 years; 2 ☐ 6 – 10 years; 3 ☐ 11 – 20 years; 4 ☐ 21 + years
Q19	Is this field owned or rented? *(If owned, skip to Q20)*	204 1 ☐ Rented; 2 ☐ Owned
	a. If rented, is the contract cash or shares?	205 1 ☐ Cash; 2 ☐ Shares
	b. If rented, how likely is it that you will be able to renew the lease for the next five years?	206 1 ☐ Very likely; 2 ☐ Unlikely; 3 ☐ Likely; 4 ☐ Unknown
Q20	All crop(s) planted in 2019?	207 1 ☐ Corn; 2 ☐ Soybeans; 3 ☐ Wheat; 4 ☐ Cotton; 5 ☐ Sorghum; 6 ☐ Peanuts; 7 ☐ Rice; 8 ☐ Oats; 9 ☐ Other (Specify: ____________________)
Q21	Which tillage practice did you predominantly use on this field in 2019?	208 1 ☐ Conventional-till; 2 ☐ No-till; 3 ☐ Strip-till
Q22	Of the last 10 years, how many years have you used no-till or strip-till on this field?	209 _________years

4

Section 2: Details on Base Field (continued)		
Q23 Did you use cover crops on this field in 2019?	210	1 ☐ Yes 3 ☐ No
Q24 Did you use manure on this field in 2019?	211	1 ☐ Yes 3 ☐ No
Q25 What type of irrigation was in the field in 2019?	212	1 ☐ None, Dryland 2 ☐ Center Pivot or Linear 3 ☐ Drip Tape 4 ☐ Furrow 5 ☐ Other (Specify: ________________)
Q26 Were there terraces on the field in 2019?	213	1 ☐ Yes 3 ☐ No
Q27 What is the general topography of the field?	214	1 ☐ Nearly level (Less than 1%) 2 ☐ Gently sloping (1-3%) 3 ☐ Moderately sloping (3-5%) 4 ☐ Strongly sloping (5-8%) 5 ☐ Steep (8-12%)
Q28 Which land type best describes this field?	215	1 ☐ Floodplain/bottomland 2 ☐ Hilly/upland 3 ☐ Neither
Q29 Approximately what percentage of this field is prone to flood for more than a day? (0% to 100%)	216	_________%
Q30 To what extent do you feel that suburban housing near the field affects the choices you make on that field?	217	1 ☐ No effect 2 ☐ A slight effect 3 ☐ A significant effect
Q31 To what extent do you feel that complaints from non-farming residents near the field affect the choices you make on that field?	218	1 ☐ No effect 2 ☐ A slight effect 3 ☐ A significant effect

Section 3: Soil health of your Base Field

Q32 *The table below lists seven <u>changes</u> in soil health characteristics that some farmers desire. For each of these, please check a box to indicate **<u>how important</u>** <u>this change is</u> to you for your **Base Field**.*

		Very Important	Fairly Important	Important	Slightly Important	Not Important at All	Don't Know
a. Increasing water infiltration	301	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
b. Increasing organic matter	302	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
c. Decreasing runoff	303	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
d. Decreasing erosion	304	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
e. Decreasing bulk density	305	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
f. Decreasing compaction	306	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
g. Increasing drainage	307	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐

Q33 *For each of the following indicators of soil health listed, please check a box to indicate you believe that **<u>using no-till or strip-till on your Base Field</u>** would increase or decrease the following soil characteristics?*

		Greatly Increase	Increase	Neither	Decrease	Greatly Decrease	Don't Know
a. Water infiltration	308	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
b. Organic matter	309	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
c. Runoff	310	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
d. Erosion	311	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
e. Bulk density	312	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
f. Compaction	313	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
g. Drainage	314	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐

Section 4: Choices

In this section we ask 9 questions that are all quite similar. Together with responses from all the other respondents, your answers will help us understand how farmers feel about soil health, which will help policy makers develop appropriate policies.
Please answer all 9 questions by checking the box at the bottom of the column you choose.

*Suppose you are looking to expand your operation by renting an additional field of land. There are two fields on the market. Both fields are **identical to your base field** except for:*

- *Water infiltration*
- *Organic matter*
- *Compaction*
- *Rental rate*

*For both fields, the cash rental agreement would be **valid for at least 5 years**. In each of the 9 choice tables, identify the field you would choose to rent.*

Before beginning, **please indicate your estimate of the typical cash rental rate for a field *like your base field***:

400 ____________ ($/acre).

Refer to this as the "typical price"

Choice 1: Please identify the option you would choose.

	Field A	Field B	
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 3 hours	1 inch of standing water absorbs in 5 hours	Neither A nor B
Organic matter (%)	1%	2.5%	
Compaction	Restricts root growth partially	Restricts root growth substantially	
Cash rental rate ($/acre per year)	$10/acre more expensive than typical price	$10/acre less expensive than typical price	
I choose 401	1 ☐	2 ☐	3 ☐

Choice 2: Please identify the option you would choose.			
	Field A	Field B	Neither A nor B
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 3 hours	1 inch of standing water absorbs in 5 hours	
Organic matter (%)	2.5%	0.5%	
Compaction	Restricts root growth partially	Does not restrict root growth	
Cash rental rate ($/acre per year)	$10/acre more expensive than typical price	$10/acre less expensive than typical price	
I choose 402	1 ☐	2 ☐	3 ☐

Choice 3: Please identify the option you would choose.			
	Field A	Field B	
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 3 hours	1 inch of standing water absorbs in 10 hours	Neither A nor B
Organic matter (%)	0.5%	2.5%	
Compaction	Does not restrict root growth	Restricts root growth substantially	
Cash rental rate ($/acre per year)	$10/acre more expensive than typical price	$10/acre less expensive than typical price	
I choose 403	1 ☐	2 ☐	3 ☐

Choice 4: Please identify the option you would choose.			
	Field A	Field B	Neither A nor B
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 10 hours	1 inch of standing water absorbs in 10 hours	
Organic matter (%)	0.5%	0.5%	
Compaction	Does not restrict root growth	Restricts root growth partially	
Cash rental rate ($/acre per year)	$10/acre more expensive than typical price	$10/acre less expensive than typical price	
I choose 404	1 ☐	2 ☐	3 ☐

Choice 5: Please identify the option you would choose.			
	Field A	Field B	Neither A nor B
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 3 hours	1 inch of standing water absorbs in 10 hours	
Organic matter (%)	1%	2.5%	
Compaction	Restricts root growth partially	Restricts root growth substantially	
Cash rental rate ($/acre per year)	$10/acre less expensive than typical price	$10/acre more expensive than typical price	
I choose 405	1 ☐	2 ☐	3 ☐

Choice 6: Please identify the option you would choose.			
	Field A	Field B	Neither A nor B
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 3 hours	1 inch of standing water absorbs in 5 hours	
Organic matter (%)	2.5%	1%	
Compaction	Restricts root growth partially	Restricts root growth substantially	
Cash rental rate ($/acre per year)	$10/acre less expensive than typical price	$10/acre more expensive than typical price	
I choose 406	1 ☐	2 ☐	3 ☐

Choice 7: Please identify the option you would choose.			
	Field A	Field B	Neither A nor B
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 10 hours	1 inch of standing water absorbs in 10 hours	
Organic matter (%)	0.5%	1%	
Compaction	Restricts root growth substantially	Restricts root growth partially	
Cash rental rate ($/acre per year)	$10/acre less expensive than typical price	$10/acre more expensive than typical price	
I choose 407	1 ☐	2 ☐	3 ☐

Choice 8: Please identify the option you would choose.			
	Field A	Field B	
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 3 hours	1 inch of standing water absorbs in 5 hours	Neither A nor B
Organic matter (%)	2.5%	1%	
Compaction	Does not restrict root growth	Restricts root growth substantially	
Cash rental rate ($/acre per year)	$10/acre more expensive than typical price	$10/acre less expensive than typical price	
I choose 408	1 ☐	2 ☐	3 ☐

Choice 9: Please identify the option you would choose.			
	Field A	Field B	
Water infiltration (infiltration into deeply wetted soil)	1 inch of standing water absorbs in 10 hours	1 inch of standing water absorbs in 5 hours	Neither A nor B
Organic matter (%)	0.5%	1%	
Compaction	Does not restrict root growth	Restricts root growth substantially	
Cash rental rate ($/acre per year)	$10/acre less expensive than typical price	$10/acre more expensive than typical price	
I choose 409	1 ☐	2 ☐	3 ☐

APPENDIX C

APPENDIX FOR CHAPTER 3

Table C.1: Two Way Fixed Effects: Intra-Year Effects, Dynamic, Krigged

	Kharif[a]	Rabi[a]	Pre-Monsoon[a]
	(1)	(2)	(3)
$\hat{\delta}_{2006}$	−0.091	0.035	−0.053
	(0.076)	(0.088)	(0.081)
$\hat{\delta}_{2007}$	−0.040	0.100	−0.113
	(0.072)	(0.091)	(0.090)
$\hat{\delta}_{2008}$	−0.182**	−0.067	−0.058
	(0.085)	(0.081)	(0.099)
$\hat{\delta}_{2009}$	−0.096	−0.018	0.026
	(0.070)	(0.085)	(0.085)
$\hat{\delta}_{2010}$	−0.039	0.058	0.047
	(0.106)	(0.070)	(0.090)
$\hat{\delta}_{2011}$	−0.098	−0.102	−0.140
	(0.078)	(0.087)	(0.099)
$\hat{\delta}_{2013}$	0.009	−0.068	0.050
	(0.087)	(0.114)	(0.102)
$\hat{\delta}_{2014}$	−0.212**	−0.068	−0.012
	(0.107)	(0.066)	(0.079)
$\hat{\delta}_{2015}$	−0.128	0.152	−0.055
	(0.098)	(0.204)	(0.072)
Observations	3,078	3,078	3,078
R^2	0.003	0.002	0.002
F Statistic	1.079	0.642	0.579

Note: *p<0.1; **p<0.05; ***p<0.01

[a] *Difference between the monsoon season and the labeled season*