Essays on the economics of irrigation efficiency and carbon markets

by

Micah Vollmer Cameron Harp

B.S., Cornell University, 2015
M.S., Arizona State University, 2016

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Abstract

Essay 1 - Efficiency and water use: Dynamic effects of irrigation technology adoption.

With the United States preparing to make a historic investment in drought mitigation, clarifying the impact of irrigation efficiency improvements on water resources is critically important. This paper uses three transitions in irrigation technology to investigate whether rebound effects cause such efficiency improvements to increase resource extraction, a phenomenon known as Jevon's paradox. We demonstrate how staggered adoption of irrigation technologies and dynamic treatment effects can cause improperly specified two-way fixed effects (TWFE) estimates to falsely indicate that technology adoption increases withdrawals. Using estimators appropriate for these circumstances, we find no significant evidence of rebound effects causing Jevon's paradox. The significant, dynamic effects we find explain this discrepancy and, perhaps more importantly, reveal irrigators' process of adaptation to each new technology at the intensive and extensive margins.

Essay 2 – Where will the United States farm carbon?

Geographic variation in the profitability and carbon sequestration potential of conservation agriculture practices will determine how cost-effective agricultural carbon credit programs are across the United States. As data on agricultural carbon credit program enrollment are unavailable, we use county-level data on the rate of adoption of conservation agriculture practices as a proxy for the profitability of the practices. We then map the predicted rates of adoption with spatially explicit carbon sequestration data for two practices and predict how the cost efficiency of agricultural carbon credit programs will vary spatially in the United States.

Essay 3 – Compensation mechanism, program scale, and the efficiency of voluntary carbon offset programs.

In the present voluntary market for agricultural carbon credits, the fine scale nature of participation and evaluation can inhibit program administrators' efforts to predict producers' counterfactual behavior in the absence of a program and mitigate adverse selection. In addition, the uncertainty in estimates of soil carbon sequestration, the outcome used to issue carbon credits, increases rapidly as the scale of analysis decreases. Scaling up carbon farming programs to a jurisdiction level could address these issues by reducing uncertainty in both estimates, increasing the efficiency of achieving societal carbon sequestration goals. In such a jurisdictional approach to carbon farming programs, the program provides payments to jurisdictions based on the aggregate outcome of all producers within a jurisdiction. In this paper, we use a simulation approach to investigate the success of carbon farming programs with two different compensation mechanisms as a function of the program scale and design. When operating at the individual scale, we find programs using a per unit payment-for-sequestration approach are more efficient than those using a constant payment-per-practice mechanism only if the program administrator can estimate producers' carbon sequestration potential with sufficient accuracy. As the uncertainty in carbon sequestration rises, the payment-for-practice program becomes more efficient in terms of the average cost of an additional unit sequestered. Finally, increasing the scale of the program improves the efficiency of the payment-for-practice program to a greater degree such that it can be more efficient than the alternative, even if the error in estimating carbon sequestration is low.

Essays on the economics of irrigation efficiency and carbon markets

by

Micah Vollmer Cameron Harp

B.S., Cornell University, 2015
M.S., Arizona State University, 2016

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Major Professor
Nathan P. Hendricks

Copyright

Abstract

Essay 1 - Efficiency and water use: Dynamic effects of irrigation technology adoption.

With the United States preparing to make a historic investment in drought mitigation, clarifying the impact of irrigation efficiency improvements on water resources is critically important. This paper uses three transitions in irrigation technology to investigate whether rebound effects cause such efficiency improvements to increase resource extraction, a phenomenon known as Jevon's paradox. We demonstrate how staggered adoption of irrigation technologies and dynamic treatment effects can cause improperly specified two-way fixed effects (TWFE) estimates to falsely indicate that technology adoption increases withdrawals. Using estimators appropriate for these circumstances, we find no significant evidence of rebound effects causing Jevon's paradox. The significant, dynamic effects we find explain this discrepancy and, perhaps more importantly, reveal irrigators' process of adaptation to each new technology at the intensive and extensive margins.

Essay 2 – Where will the United States farm carbon?

Geographic variation in the profitability and carbon sequestration potential of conservation agriculture practices will determine how cost-effective agricultural carbon credit programs are across the United States. As data on agricultural carbon credit program enrollment are unavailable, we use county-level data on adoption of conservation agriculture practices to predict a proxy for profitability, the rate of adoption. We then map the estimated rates of adoption with spatially explicit carbon sequestration data for two practices and predict how the cost efficiency of agricultural carbon credit programs will vary spatially in the United States.

Essay 3 – Compensation mechanism, program scale, and the efficiency of voluntary carbon offset programs.

In the present voluntary market for agricultural carbon credits, the fine scale nature of participation and evaluation can inhibit program administrators' efforts to predict producers' counterfactual behavior in the absence of a program and mitigate adverse selection. In addition, the uncertainty in estimates of soil carbon sequestration, the outcome used to issue carbon credits, increases rapidly as the scale of analysis decreases. Scaling up carbon farming programs to a jurisdiction level could address these issues by reducing uncertainty in both estimates, increasing the efficiency of achieving societal carbon sequestration goals. In such a jurisdictional approach to carbon farming programs, the program provides payments to jurisdictions based on the aggregate outcome of all producers within a jurisdiction. In this paper, we use a simulation approach to investigate the success of carbon farming programs with two different compensation mechanisms as a function of the program scale and design. When operating at the individual scale, we find programs using a per unit payment-for-sequestration approach are more efficient than those using a constant payment-per-practice mechanism only if the program administrator can estimate producers' carbon sequestration potential with sufficient accuracy. As the uncertainty in carbon sequestration rises, the payment-for-practice program becomes more efficient in terms of the average cost of an additional unit sequestered. Finally, increasing the scale of the program improves the efficiency of the payment-for-practice program to a greater degree such that it can be more efficient than the alternative, even if the error in estimating carbon sequestration is low.

Table of Contents

List of Figures

List of Tables

Acknowledgements

To begin, I want to thank Dr. Nathan P. Hendricks for investing in me throughout my doctoral education and dissertation research. When I was confused or unsure, you guided my investigation instead of directing my course. When I was incorrect, you challenged me and taught me to carefully interrogate my understanding. You constantly demonstrate how an academic should conduct themselves both in and outside the academy. I know I am a better student, researcher, and person because of your guidance.

I would also like to thank my committee members Drs. Stacy Hutchinson, Gabe Sampson, Jesse Tack, and Jisang Yu for their contributions to my work. Your comments, questions, and insights pushed me to broaden and clarify my thinking. In addition, I would like to acknowledge the support I received as part of the National Science Foundation Research Traineeship (NRT) program funded by NSF Award No. 1828571, by the USDA National Institute of Food and Agriculture, Agricultural and Food Research Initiative Competitive Program, Agriculture Economics and Rural Communities, grant number 2017-67023-26276, and from the Environmental Defense Fund grant concerning jurisdictional agricultural carbon credit programs.

Finally, I want to express my thanks to the people who have loved and supported me throughout my life. Jeff and Ellen, thank you for your love of learning and for passing it on to me. Most importantly, I would like to conclude by acknowledging the continual love and support of my incredible wife, Kelly. Being married to you is the greatest blessing I will ever have. Your passion and drive continually inspire me, thank you.

Chapter 1 - Efficiency and water use: Dynamic effects of irrigation technology adoption

1.1 Introduction

As a historic drought stresses the water resources of the western United States, agricultural producers and policy makers alike are under pressure to find solutions which ensure food production with an increasingly scarce resource (Rosa et al. 2018; Droppers et al. 2021). In response to this crisis, the Inflation Reduction Act of 2022 includes four billion dollars to facilitate drought mitigation in states west of the Mississippi and specifies the following as one of three approved activities: “Voluntary system conservation projects that achieve verifiable reductions in use of or demand for water supplies or provide environmental benefits in the Lower Basin or Upper Basin of the Colorado River” (117th Congress 2022). Public reporting concerning this allocation, despite the unspecified means of achieving reductions in the previous statement, suggests producers in priority basins could be paid to install more efficient irrigation technologies (Wilson 2022). While improving irrigation efficiency sounds congruent with the goal of continuing production with scarcer resources, the literature suggests adoption of increasingly efficient irrigation technologies may be counterproductive and result in the degradation of water resources.

One reason for adverse consequences of technology adoption is based on the hydrologic water balance in a system. While the technology can affect water withdrawals, it may also affect return flows, which highlights the important distinction between water use and consumption (Ward and Pulido-Velazquez 2008; Huffaker 2008). The second reason for adverse consequences of technology adoption is the potential for Jevon's paradox, where changes in behavior following technology adoption could cause an increase in water use. While the water balance effect is a critical component of water management—especially in surface water contexts—we focus our analysis on the potential for rebound effects and Jevon's paradox in water withdrawals.

Rebound effects are generated when resource users adapt to improvements in resource use efficiency such that the resource savings produced by the increase in efficiency are partially or completely offset (Paul et al. 2019). Jevons' paradox, or a "backfire," occurs if the rebound effect is large enough to create a net increase in resource consumption following an efficiency improvement (Greening et al. 2000; Alcott 2005). William Jevons first proposed this phenomenon in his 1865 book on coal where he explicitly considered general equilibrium effects of increasing energy efficiency (Jevons 1865).

The more recent literature on energy use is chiefly concerned with these general equilibrium effects in macroeconomic contexts or the microeconomics of energy efficiency within a consumer framework (Khazzoom 1980; Saunders 1992; Brookes 2000). Our work differs due to our focus on a production environment, but it is similar to previous microeconomic investigations in that we concern ourselves with "direct" rebound effects—the increase in resource use induced by reductions in the cost of use due to efficiency improvements (Chan and Gillingham 2015). While empirical evidence suggests rebound effects in utility maximization contexts are rarely sufficient to create such a paradox (Sorrell et al. 2009; Gillingham et al. 2013), a review of empirical literature specific to agricultural production contexts finds significant evidence of strong rebound effects or Jevon's paradox and highlights irrigation technologies as one of most common culprits (Paul et al. 2019).

In their global review of 230 case studies involving adoption of a more efficient irrigation technology, Pérez-Blanco et al. (2020) find water withdrawals increased in 9.1 percent of studies with available data and another 10.8 percent had ambiguous evidence. While they do not speculate on the reasons behind the cases with increased withdrawals, Pérez-Blanco et al. (2020) suggest the increase in water consumption they find in the majority of cases is due to a lack of restrictions on water use. In an earlier review, Berbel et al. (2015) consider empirical work largely drawn from Spain and conclude adoption of a more efficient irrigation technology can increase water use if irrigators have the option to expand irrigated acreage or adopt more water intensive crops. Perry and Steduto (2017) echo these conclusions and suggest increases in water consumption precede future increases in water use because of the rise in profitability of the resource following adoption of the more efficient technology. In all three review papers, irrigators' unconstrainted profit maximization decisions over the extent of irrigated acreage, intensity of irrigation for current crops on a per acre basis, or adoption of more water intensive crops are the driving force behind a backfire after a more efficient technology is adopted.

Kansas serves as a natural testbed for these theories due to its rich data on irrigation and its producers' early adoption of pressurized irrigation systems relative to other states. In 2018 for example, just 3 percent of the 2.4 million irrigated acres in Kansas used a gravity-fed system, but in the same year 45 percent of the 2.5 million irrigated acres in the neighboring state of Colorado used a gravity-fed system (United States Department of Agriculture 2020). Given the widespread transition from gravity-fed to pressurized systems occurring in the United States (Hrozencik 2022), the well-documented behavior of irrigators in Kansas provides valuable insights to inform other regions currently undergoing a similar transition. As a testament to this, two influential studies have measured rebound effects for irrigation water use in Kansas. Pfeiffer and Lin (2014) find that the adoption of Low Energy Precision Application (LEPA) irrigation, a more efficient center pivot system, increased water use by farmers compared to a traditional center pivot system. Their results indicate that this Jevons' paradox primarily occurred because farmers adopted more water intensive crops and thus applied more water per irrigated acre, rather than an expansion in irrigated acres. This same conversion in irrigation technologies is studied in Li and Zhao (2018), where the authors find significant heterogeneity in rebound effects depending on the size of irrigators water rights. Specifically, irrigators with larger water rights are more likely to increase water use following LEPA adoption.

In this paper, we estimate the effect of adopting a more efficient irrigation technology for three different types of technology transitions in Kansas using a Difference-in-Differences identification strategy and an estimation approach amenable to settings with staggered adoption and heterogeneous treatment effects. For each of the three transitions, we find significant evidence of irrigators adjusting groundwater use following adoption due to adjustments made at the intensive and extensive margins over several years. While the dynamic treatment effects for the adoption of LEPA and endgun removal suggest irrigators make minute changes following these conversions, the magnitude of the immediate reduction in groundwater withdrawals we observe when irrigators switch from flood to center pivot irrigation indicates this conversion could reduce near-term strain on water resources. We anticipate these results will be especially pertinent to groundwater reliant regions currently undergoing the transition from gravity-fed to pressurized irrigation systems. In addition to revealing irrigators' adaptation behavior, the significant dynamic treatment effects we find provide an explanation for the difference in average treatment effect results between the two estimators we use which accommodate staggered adoption and heterogeneous treatment effects and the alternative two-way fixed effects specification (TWFE) we include for comparison (Goodman-Bacon 2021). As we demonstrate in a simulation context, if the effect of adopting the new technology grows stronger over time such that contemporaneous reductions in groundwater use by later adopters are smaller than those of early adopters, faulty comparisons of these groups by TWFE can create the illusion of later adopters increasing water use in comparison.

Our work contributes to the study of rebound effects, Jevon's paradox, and technology adoption in two significant dimensions. First, we demonstrate how to address two significant empirical challenges, the staggered nature of irrigation technology adoption and heterogenous treatment effects, in an applied context and recover unbiased estimates of the effect of adopting a more efficient technology. While these concerns are motivated by our empirical context, we anticipate future studies of technology adoption will benefit from our work given that staggered adoption is typical of technology diffusion processes (Griliches 1957; 1958; Feder et al. 1985; Sunding and Zilberman 2001). Second, we estimate dynamic treatment effects to understand how irrigators adapt to the new technology over time and to demonstrate how they function as an often unaccounted for source of treatment effect heterogeneity (Cunningham 2021). The evidence of dynamic treatment effects we find indicates irrigators adjust to the new technology over multiple growing seasons as is the case in other agricultural contexts (Conley and Udry 2010). Finally, for policy makers considering technology improvements as a water conservation tool, our results suggest program evaluations in similar contexts need to incorporate longer post-treatment periods and estimators which accommodate adaptation.

1.2 Staggered adoption and dynamic treatment effects

In this section, we illustrate the complexities of our empirical setting created by the interaction between staggered technology adoption and dynamic treatment effects. In doing so, we illustrate how the results from two-way fixed effects estimation could suggest evidence of Jevon's paradox despite there being a net reduction in resource use after a new technology is adopted. We use a hypothetical staggered adoption setting to demonstrate, both graphically and numerically, the source of the bias in two-way fixed effects. While previous studies have described the bias of two-way fixed effects in detail (Borusyak 2018; de Chaisemartin and D'Haultfoeuille 2020; Sun and Abraham 2021; Goodman-Bacon 2021), our purpose in this section is to illustrate the bias in the context of irrigation technology adoption and water use.

Consider a collection of irrigators with an opportunity to invest in a more efficient technology at any period, t, before the last period in the panel dataset, T. For the following demonstration, let $T = 3$. We consider three groups (or cohorts) of irrigators indexed by the subscript g which indicates the first year in which they adopted the new technology. Additionally, let the indicator variable, $G_{i,g}$, indicate the year in which a group is first treated by taking a value of one if irrigator i is first treated in period g. The first group of irrigators (i.e., early adopters), adopt the more efficient technology at the beginning of the second season such that $g = 2$ and $G_{i,2} = 1$ for irrigators in this group. The second group (i.e., late adopters), adopt the more efficient technology at the beginning of the third season meaning $g = 3$ and $G_{i,3} = 1$. The last group of irrigators do not adopt the new technology in any of the observed years (i.e., never adopters), so we set their group index to the period after the end of the panel dataset such that $g = 4$ and $G_{i,4} = 1$.

The outcome of interest, $Y_{i,g,t}$, is the acre-feet of groundwater withdrawn by irrigator i in group g for year t, and the treatment effect we want to recover is the cumulative change in groundwater extraction that occurred by time T due to irrigators adopting the more efficient technology. To express this average treatment effect (ATE), we represent irrigation technology adoption status using a binary variable, $D_{i,g,t}$, that takes a value of 1 for every period in which the respective irrigator has adopted the newer, more efficient irrigation technology. Then, as each observation within a group has an identical treatment sequence, meaning $D_{i,g,t} = D_{g,t}$ within group g, we represent the treatment path for all groups $g \in 2,3,4$ as $D_g = \{D_{g,1}, D_{g,2}, D_{g,3}\}$. As an example, the treatment sequence for the early adopters is written: $D_2 = \{0, 1, 1\}$. We define $Y_{i,g,t}(D_g)$ as the potential outcome for irrigator i in group g at time t if it experienced treatment trajectory D_g, a modification to the potential outcomes framework of Callaway and Sant'Anna (2020). Lastly, we define N_g as the number of irrigators in group g and N_s as the total number of irrigators who adopt the more efficient technology before the end of the panel. We can then express the average treatment effect in year T as the weighted average of the treatment effects for early and late adopters:

$$ATE = \frac{N_2}{N_s} E\left[Y_{i,2,3}(D_2) - Y_{i,2,3}(D_4)\right] + \frac{N_3}{N_s} E\left[Y_{i,3,3}(D_3) - Y_{i,3,3}(D_4)\right]. \quad \text{(Eq. 1.1)}$$

However, we cannot compute the ATE in Eq. 1.1 because we do not observe the counterfactual outcome evolutions $Y_{i,2,3}(D_4)$ and $Y_{i,3,3}(D_4)$ for the early and late adopters respectively. A difference-in-differences (DD) identification strategy effectively imputes these counterfactual outcomes, where adoption of the technology is considered as the "treatment." The crucial parallel trends assumption requires irrigators who adopt the new technology to have the same change in the dependent variable as those that do not adopt in the counterfactual scenario where the adopters do not actually adopt. Using this assumption, we can rewrite Equation 1.1 as the following difference-in-differences estimand using observable outcomes:

$$ATE^{DD} = \frac{N_2}{N_s} E\left[\left(Y_{i,2,3}(D_2) - Y_{i,2,1}(D_2)\right) - \left(Y_{i,4,3}(D_4) - Y_{i,4,1}(D_4)\right)\right] + \frac{N_3}{N_s} E\left[\left(Y_{i,3,3}(D_3) - Y_{i,3,1}(D_3)\right) - \left(Y_{i,4,3}(D_4) - Y_{i,4,1}(D_4)\right)\right]. \quad \text{(Eq. 1.2)}$$

With the effect of interest defined and identification strategy selected, we now specify the groundwater extraction behavior for the three groups and demonstrate how two-way fixed effects estimation fails to recover the ATE. At time $t = 1$, irrigators in the early adopter, late adopter, and control groups apply 160, 165, and 170 acre-feet of groundwater each year respectively. Both groups of adopters decrease their groundwater withdrawals by two acre-feet in their first season using the more efficient technology, and the early adopters reduce their withdrawals by an additional 18 acre-feet in their second season post-adoption. The cumulative reduction of 20 acre-feet worth of groundwater represents a 12.5% decrease in withdrawals for the early adopters. Lastly, for simplicity, we assume irrigators in the early and late adopter groups would have maintained their respective initial levels of groundwater use in the counterfactual scenario with no adoption of the more efficient technology. Figure 1.1 displays the described scenario graphically.

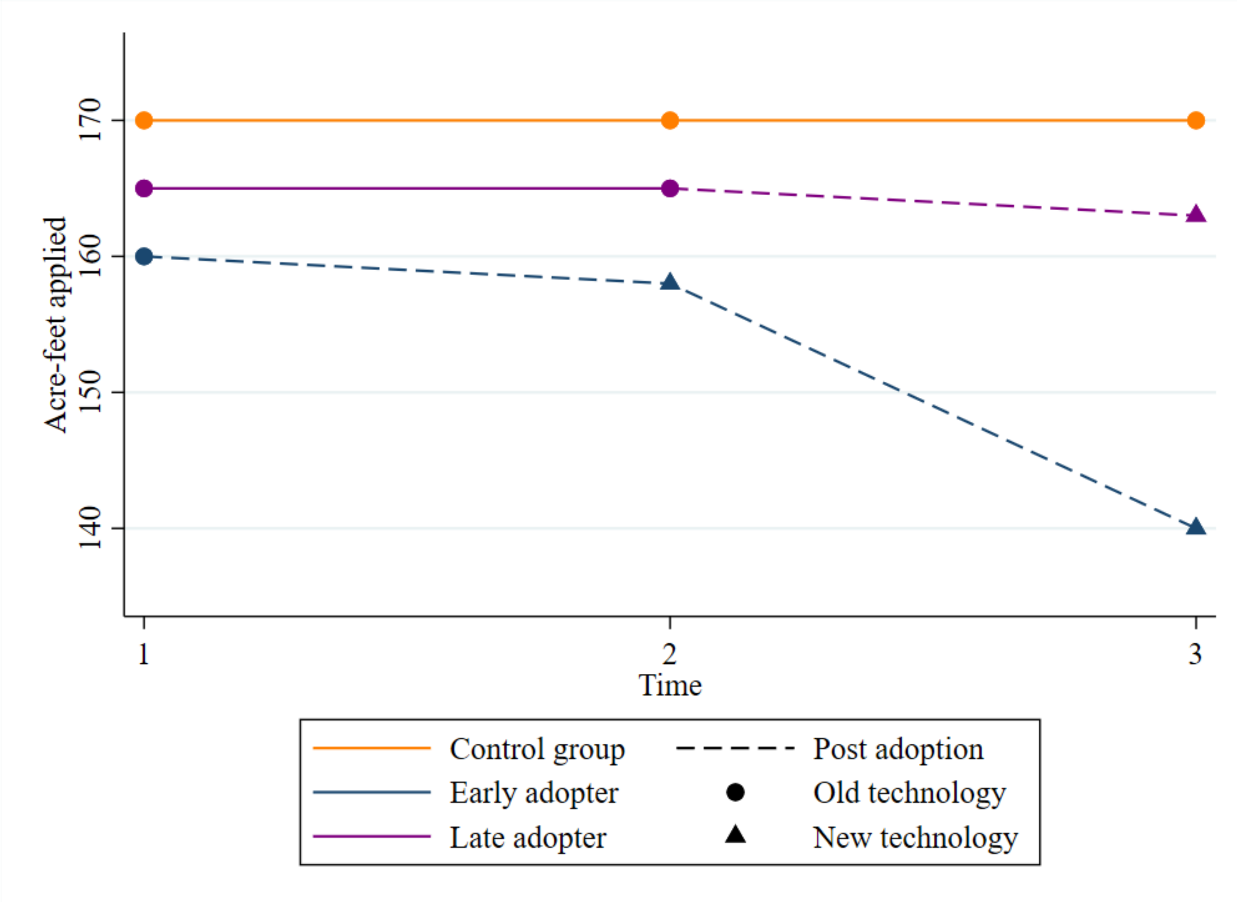

Figure 1.1. Hypothetical irrigator water use behavior with two treatment cohorts. The early group adopts the new technology in the season preceding time t=2 and the late group adopts the technology in the season before time t=3.

As each unique combination of i and g comprises a panel unit, we can produce the following linear unobserved effects panel data model with time and unit fixed effects:

$$Y_{i,g,t} = \delta_{i,g} + \gamma_t + \beta^{DD} D_{i,g,t} + u_{i,g,t}. \qquad \text{(Eq. 1.3)}$$

Estimating Eq. 1.3 using fixed-effects estimation produces a biased estimate of the ATE because of the staggered nature of adoption, even though the parallel trends assumption holds. As demonstrated by the Difference-in-Difference Decomposition Theorem of Goodman-Bacon (2021), the two-way fixed effects estimator for $\hat{\beta}^{DD}$ is a weighted average of simpler two-by-two difference-in-differences (DD) estimators including two erroneous comparisons between early and later adopters of the new technology. Figure 1.2 displays a Goodman-Bacon style decomposition of the four difference-in-differences estimates used by the two-way fixed effects estimator. In Figure 1.2, panel (a) is a DD using never-adopters as the control for late adopters, panel (b) uses never adopters as the control for early adopters, panel (c) uses late adopters as a control for early adopters before late adopters adopt, and panel (d) uses early adopters as a control for late adopters when late adopters adopt.

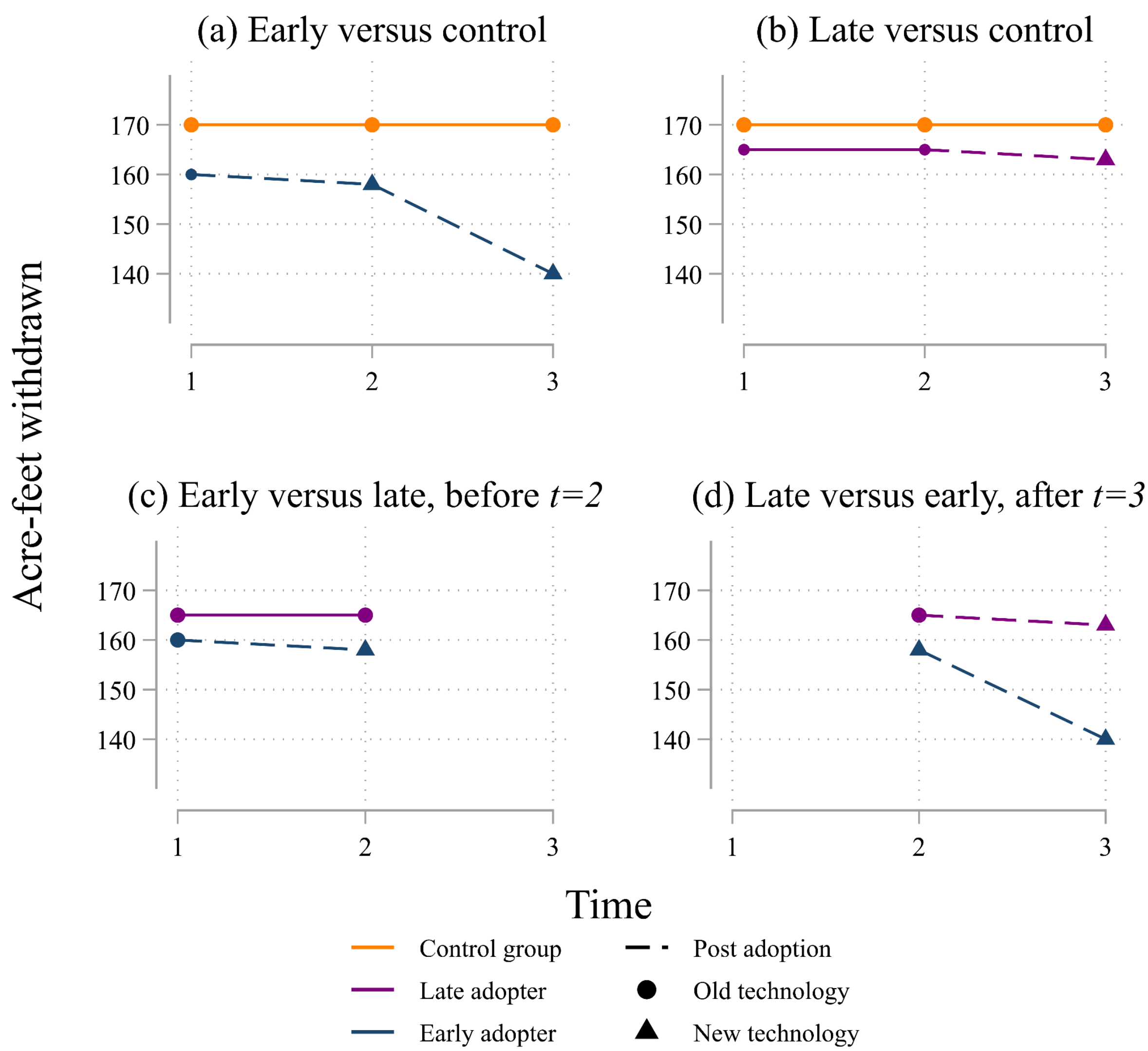

Figure 1.2. Decomposition of the two-way fixed effects estimator applied to the staggered adoption example displayed in Figure 1.1. Panel (d) displays the problematic use of early adopters as the control group for late adopters.

The treatment effect estimated from the DD scenario displayed in panels (a), (b), and (c) of Figure 1.2 recover the true average treatment effect (ATE) for each of the treated cohorts at their respective times. However, the treatment effect estimated by the DD in panel (d) is biased because it uses the faulty counterfactual outcome produced using the early adopters' outcome trajectory. In fact, panel (d) gives the illusion of a positive treatment effect because the late adopter (i.e., the treated) decreases water use less than the early adopter (i.e., the control). The weights for each DD in Figure 1.2 used to create the two-way fixed effect estimate depend on the variance of the fixed-effects-adjusted independent variable and the subsample size in each comparison. As such, the size of the early and late adopter groups and the variation in treatment effects amongst them will determine how heavily the problematic DD in panel (d) of Figure 1.2 is weighted (Słoczynski 2020; Goodman-Bacon 2021). If weighted heavily enough, the incorrectly signed DD estimate in this example would suggest Jevon's paradox occurs for this technology adoption, despite the exact opposite being true.

To numerically demonstrate the bias introduced by improperly applying TWFE estimation, we conduct a Monte Carlo simulation of the scenario outlined above and depicted in Figure 1.1. We conduct one thousand replications of the simulation where each observation includes a random noise term with mean zero and a standard deviation of one. There are a total of 300 irrigators in the panel dataset we simulate with 50 early adopters, 200 later adopters, and 50 never-adopters. Using the estimating equation in Equation 1.3, the resulting TWFE coefficient for $\hat{\beta}^{DD}$ indicates a one acre-foot *increase* in groundwater withdrawals with a standard deviation of 0.14 across 1,000 replications. However, using Equation 1.2, the true average treatment effect in year T given these group sizes is $\frac{50}{250}[140 - 160] + \frac{200}{250}[163 - 165] = -5.6$ acre-feet. In the next two sections, we address the differences between this simplified case and our empirical context and then outline each of the three estimators we use to estimate the effect of adopting a more efficient irrigation system on groundwater extraction.

1.3 Data

1.3.1 WIMAS

Historical water use data for Kansas irrigators are available through the Water Information Management and Analysis System (WIMAS), a joint effort by the Kansas Department of Agriculture, Division of Water Resources and Kansas Geological Survey (Wilson et al. 2005). Water right records within WIMAS contain annual reported water withdrawals, irrigated acreage, the point of diversion where water is withdrawn from, and the irrigation technology used, along with other data. The depth applied is constructed (i.e., total withdrawals per acre) by dividing total withdrawals by acres irrigated. As WIMAS is comprised of data reported by individual irrigators, there are oddities and extreme values within some records. To prevent these values from distorting our results, we remove observations that have total withdrawals, acres irrigated, or depth applied greater than the 99th percentile. Additionally, we remove observations with a reported quantity for withdrawals but zero reported irrigated acres or vice versa. Most of the time irrigation technology is not reported when withdrawals or acres irrigated are equal to zero. There are a few times when the irrigation technology is reported but, in these cases, we replace the technology as missing to be consistent.

Due to the structure of water rights in Kansas, multiple water rights can be associated with a single well, or vice versa. So, for this study, we aggregate water use to a unit of observation called the water right group (Rosenberg 2020). Specifically, we begin by matching water rights and their associated points of diversion to their place of use—the specific tract where the water right is authorized to be applied. Then, we find overlaps between places of use based on shared points of diversion or shared water rights to define the water right groups. As such, water right groups are networks of water rights sharing common points of diversion or places of use. Among the water right groups, 44% contain a single point of diversion and a single water right, while the rest contain overlaps. Using water right groups as the unit of aggregation avoids the concern of water use being shuffled between multiple water rights with the same place of use and thus allows us to accurately track changes in water withdrawals through time, even if the related water rights or points of diversion change for administrative reasons. Summary statistics for the dataset containing the three technologies we examine are presented in Table 1.1. Figure 1.3 illustrates the variation of the three dependent variables and the adoption of the three technologies over time.

Table 1.1. Summary statistics for water right groups with flood, center pivot, or LEPA irrigation systems across all years from 1991 to 2019.

Variable	N	Mean	Std. Dev.	Min.	Max.
Total withdrawals (AF)	181,052	208.06	190.18	0.01	1,431.00
Flood irrigation	34,476	178.69	204.86	0.01	1,426.96
Center pivot irrigation	36,232	201.60	181.04	0.02	1,427.24
LEPA irrigation	110,344	219.36	187.25	0.01	1,431.00
Irrigated acres	181,052	185.01	138.02	1.00	910.00
Flood irrigation	34,476	152.98	142.90	1.00	910.00
Center pivot	36,232	179.42	123.76	4.00	910.00
LEPA irrigation	110,344	196.85	139.20	1.00	910.00
Depth applied (ft)	181,052	1.11	0.49	0.00	2.47
Flood irrigation	34,476	1.12	0.56	0.00	2.47
Center pivot irrigation	36,232	1.10	0.49	0.00	2.47
LEPA irrigation	110,344	1.11	0.46	0.00	2.47
Water right group variables					
Authorized quantity (AF)	181,052	379.57	358.25	0.00	7,698.30
Authorized acres	181,052	253.18	229.28	0.00	4,093.60
Soil variables					
Sand content (%)	181,052	36.17	28.13	5.23	97.35
Silt content (%)	181,052	41.63	21.46	0.60	71.11
Available water capacity (cm/cm)	181,052	0.17	0.04	0.05	0.22
Pre-development aquifer variables					
Specific yield	181,052	16.75	3.86	0.00	25.00
Depth to water	181,052	65.43	56.12	0.00	260.19
Saturated thickness	181,052	153.64	113.92	-0.16	610.69
Hydraulic conductivity	181,052	82.09	26.03	0.00	135.00
Weather variables					
Preseason precipitation (mm)	181,052	30.45	15.64	2.08	96.50
Growing season precipitation (mm)	181,052	78.07	26.31	13.30	196.47
Preseason evapotranspiration (mm)	181,052	2.27	0.23	1.52	2.93
Growing season evapotranspiration (mm)	181,052	5.14	0.31	4.34	6.18

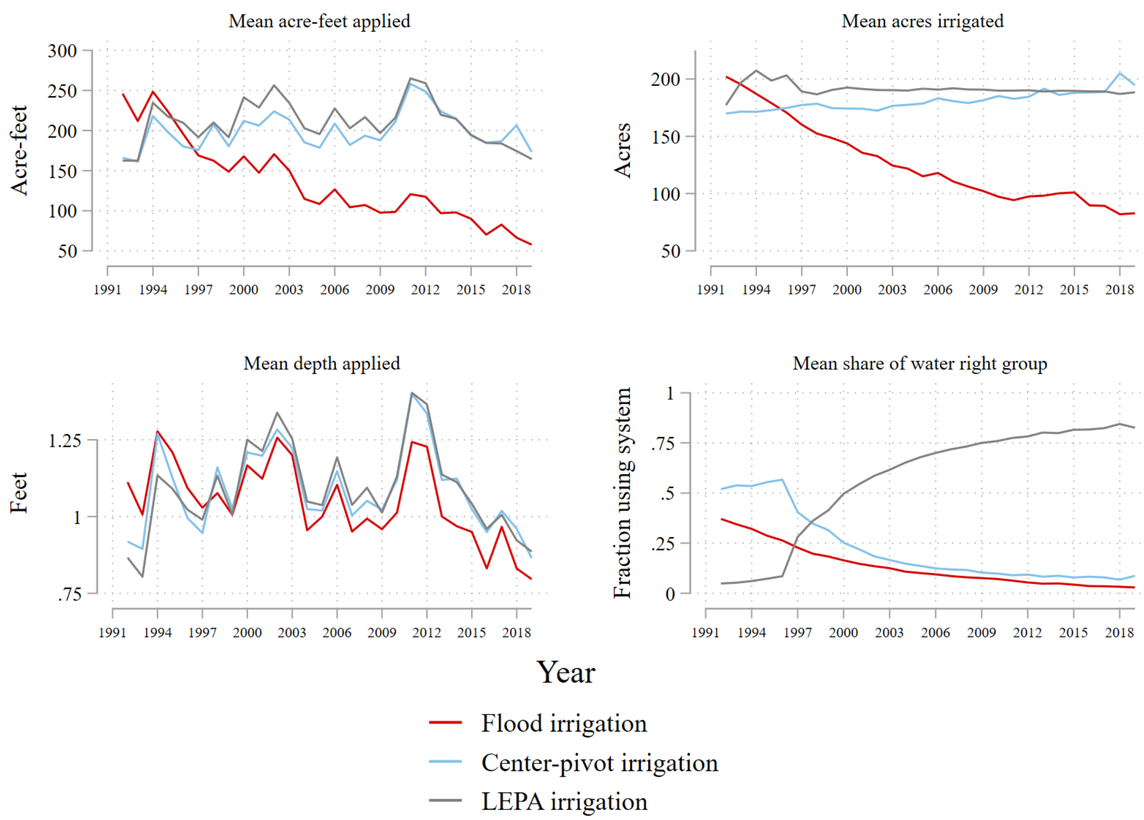

Figure 1.3. Average total withdrawals, acres irrigated, depth applied, and mean share of water right groups by irrigation system over time.

1.3.2 Soil, weather, and aquifer characteristics

We draw monthly data for cumulative precipitation, average maximum temperature, and average minimum temperature from the Parameter-elevation Regressions on Independent Slopes Model repository maintained by Oregon State University (PRISM Climate Group 2014b). We then calculate the mean value of each variable for every Public Land Survey System (PLSS) section (roughly 1 square mile) in Kansas and match to water right groups by place of use. Evapotranspiration data are generated with water right group specific data on average temperature, latitude, and elevation using the Penman-Monteith equation as outlined in Allen, Pereira, and Smith (1998). The elevation data are drawn from the Soil Survey Geographic

(SSURGO) Database along with other time invariant soil characteristics including the water holding capacity, soil texture, and hydraulic class (Soil Survey Staff 2022). Soil characteristics reflect the dominant soil type at the section level and are matched to water right group's respective places of use. Time invariant hydrologic characteristics of the High Plains Aquifer at the section level are obtained from the Kansas Geological Survey and include the predevelopment values for saturated thickness, depth to groundwater, specific yield, and hydraulic conductivity. Summary statistics for both the time varying and invariant covariates are displayed in the bottom section of Table 1.1.

1.3.3 Creation of technology adoption sub-samples

To isolate the effect of each respective technology adoption, we create three sub-samples of our panel dataset. The first contains irrigators who use flood irrigation, center pivot irrigation, or center pivot with LEPA irrigation and allows us to isolate the effect of transitioning from flood irrigation to a center pivot system. For the second, we only include irrigators using a center pivot system, with or without LEPA, so we can estimate the effect of adopting LEPA. The last also contains irrigators using a center pivot system with or without LEPA because endguns can be used in either case. The difference-in-differences identification strategy and estimators we employ, detailed in full in the next section, require some additional modifications to these sub-samples to ensure we can construct appropriate counterfactual outcomes using the behavior of irrigators who do not adopt the new technology.

To estimate the effect of transitioning from flood to center pivot irrigation, we limit the respective sub-sample to the years from 1991 to 2012. By 2012, ninety percent of the irrigators who we observe transitioning between flood and center pivot irrigation have already made the switch. Furthermore, as shown in Figure 1.4, there are very few irrigators remaining using flood irrigation after this point. As such, a larger portion of the irrigators who remain using flood irrigation may be systematically different from those will adopt center pivot irrigation at some point in the future, violating the conditions necessary for our identification strategy. For the conversion from traditional center pivot to LEPA, we use the entire sample period 1991-2019 because at its lowest the fraction of irrigated acreage using a center pivot system is above nine percent. For the endgun removal sub-sample, we use data from 2006 until 2019 because the earliest cohort removes their endguns in 2010.

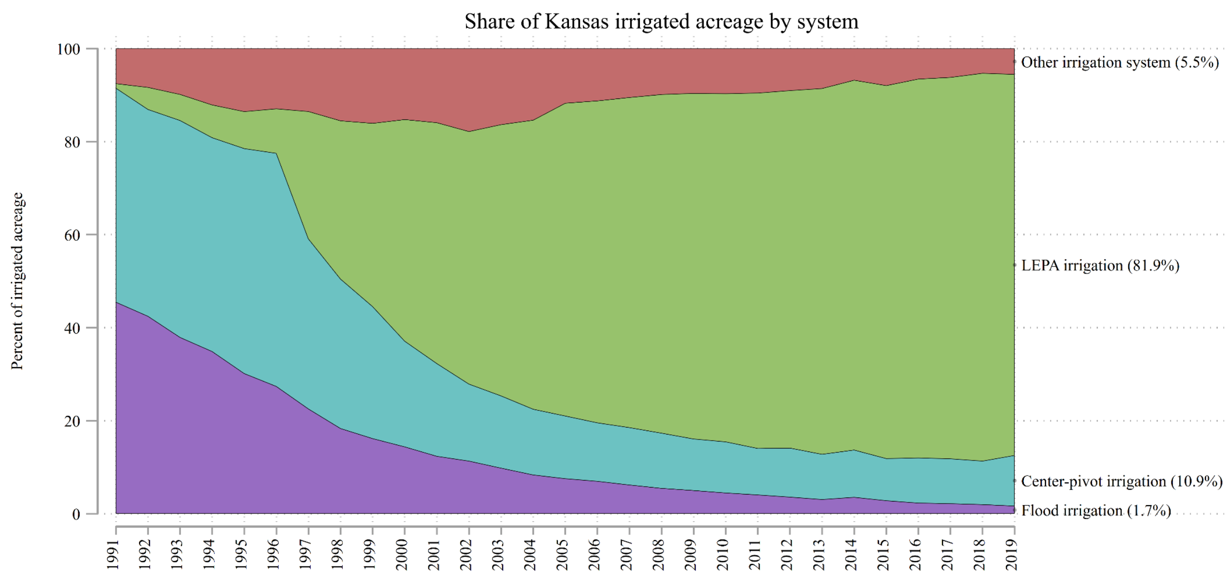

Figure 1.4. Percent of Kansas' total irrigated acreage using irrigation technologies over time.

To preclude the use of faulty controls, we remove water right groups that have already adopted the technology before the first year of the sub-sample. For the same reason, we do not include water right groups who changed technologies during the sub-sample time period if we do not observe the change occurring between consecutive years. In these instances, we either cannot estimate the effect of a change in technology without observing a transition, or we are unable to determine when the transition occurred. As our estimators described in the next section rely on discrete treatment categories, we remove any water right groups using multiple technologies in a single year. Finally, we remove water right groups that report adopting the new technology but then revert to the older technology at a later point in time.[1]

Given the small number of water right groups who recorded removing endguns in the third sub-sample (67 in total), we further limit the control pool for this sub-sample in two ways. First, we only keep water right groups in the fifth groundwater management district (GMD) where the endgun removal program occurred who have observations in all years from 2006 through 2019. Next, we use Mahalanobis distance matching based on water right groups' first recorded total withdrawals in the pre-treatment period, predevelopment aquifer characteristics, and soil variables from Table 1.1 (King and Nielsen 2019; Leuven and Sianesi 2003). The time periods and balance of adopting versus non-adopting water right groups in the resulting sub-samples are contained in Table 1.2.

[1] For example, some groups report flood irrigation after converting to center pivot. Converting a technology involves large fixed costs so it highly unlikely that a farmer would actually revert back to an older technology. There are a couple of reasons why reverting might get reported. One reason is that it could be a reporting error. A second reason is that different portions of the water right group that have different technologies could have been irrigated in different years. For example, field A has flood technology and field B has center pivot but both fields are in the same water right group. If the farmer alternates between irrigating field and A and B in different years, then it will appear that technology conversion is occurring when there is no true conversion.

Table 1.2. Sub-sample characteristics for water right groups in each technology transition.

Variable	Technology transition panel datasets		
	Flood adopting center pivot or LEPA	Center pivot adopting LEPA	Center pivot or LEPA removing endguns
Years included	1991-2012	1991-2019	2007-2019
Water right groups adopting within sub-sample years	1,532	3,716	67
Water right groups adopting after sub-sample years	95	0	0
Never-adopter water right groups	1,144	273	46 (3,205 in GMD 5 before MDM matching)
Filtering of control pool	None	None	Only those in GMD 5, MDM matching

1.4 Empirical specifications for Difference in Differences models

Each of the three estimators described in this section involve variations of the following linear equation:

$$Y_{i,g,t} = \delta_{i,g} + \gamma_t + \Phi X_{i,g,t} + \Psi Z_{i,g} + \beta^{DD} D_{i,g,t} + u_{i,g,t}, \qquad \text{(Eq. 1.4)}$$

where $Y_{i,g,t}$ represents the outcome of interest (i.e., withdrawals, acres irrigated, or depth applied) and $X_{i,g,t}$ and $Z_{i,g}$ denote time varying and invariant covariates. Covariates varying from year to year, $X_{i,g,t}$, that are included in all estimators are pre-season and growing season total precipitation and evapotranspiration. One estimator also allows the inclusion of time invariant controls $Z_{i,g}$, which include soil characteristics and predevelopment aquifer attributes. By including these covariates, we rely on some form of the conditional parallel trends assumption for each of the three estimators below. The assumption of conditional parallel trends allows for differential trends in evolution of the dependent variables so long as they are explained by the changes in time-varying covariates and the covariates included are unaffected by the treatment (Sant'Anna and Zhao 2020; de Chaisemartin and D'Haultfoeuille 2020).

1.4.1 Two-way fixed effects

We include the two-way fixed effects specification so we can compare the resulting estimates to those of previous studies and the two recent estimators which accommodate staggered adoption. As in the demonstration case, the subscript g indexes cohorts of irrigators who first adopt the new technology at time $t = g$. Within each cohort, the individual irrigators are indexed by the subscript i running from 1 to the total number of irrigators in group g at time t, $N_{g,t}$. The two-way fixed effects results in our study use the within fixed effects estimator that are calculated by applying the within transformation:

$\ddot{Y}_{i,g,t} = \gamma_t + \Phi\ddot{X}_{i,g,t} + \beta^{DD}\ddot{D}_{i,g,t} + \ddot{u}_{i,g,t}$; where $\ddot{x}_{i,g,t} = x_{i,g,t} - \bar{x}_{i,g}$. (Eq. 1.5)

However, as shown above, the TWFE estimate of $\hat{\beta}^{DD}$ will be a biased estimate of the average treatment effect. In addition to addressing unobserved water right group effects, time-demeaning also causes the time invariant covariates contained in $Z_{i,g}$ to drop out.

1.4.2 de Chaisemartin & D'Haultfouille

To address the problematic staggered nature of adoption, we first employ the DID_ℓ estimator from de Chaisemartin and D'Haultfoeuille (2022) to estimate the effect of adopting a new technology over all periods following adoption.[2] As in de Chaisemartin and D'Haultfoeuille

[2] The DID_ℓ estimator expands upon the DID_M estimator of instantaneous treatment effects outlined in de Chaisemartin and d'Haultfoeuille (2020). While the DID_ℓ estimator accommodates treatments that turn on and off over time, we limit our sample so the treatment group includes irrigators who switch into the new technology once and never revert to the older technology.

(2022), we let $N^1_{t,\ell}$, indicate the number of irrigators in the group who first adopted the new technology ℓ years before t, and N^{nt}_t denote the number of irrigators in groups who have not adopted the technology from the beginning of the time series until t. Then, to estimate the average outcome across all post-treatment periods, and the dynamic treatment effects for each post-treatment period, we use the following difference-in-differences estimator from de Chaisemartin and D'Haultfoeuille (2022),

$$\widehat{DID}_{t,\ell} = \frac{1}{N^1_{t,\ell}} \sum_{\forall\, i \in g=t-\ell} \left(Y_{i,g,t} - Y_{i,g,t-\ell-1}\right) - \frac{1}{N^{nt}_t} \sum_{g=t+1} \left[\sum_{i=1}^{N_{g,t}} \left(Y_{i,g,t} - Y_{i,g,t-\ell-1}\right)\right].$$

(Eq. 1.6)

In Equation 1.6, the index of the first summation restricts g such that observations in the treatment group of interest are selected.[3] So the treatment group, those with $g = t - \ell$, will only contain the cohort who were first treated at time $t - \ell$. The second summation index selects the control observations for periods $t - \ell - 1$ to t, observations with $g > t$, so the control group contains all cohorts who are not yet treated at time t. As such, it compares the evolution of the outcome variable from $t - \ell - 1$ to t for those who adopted the new technology ℓ years ago with groups who have yet to adopt by t. Note, $g = T + 1$ for groups who never adopt the technology, so they are included in this second summation along with irrigators who adopt the technology in periods after t. This means the control group is comprised of both "not-yet-treated" and "never-treated" irrigators. Unlike the TWFE specification in Equation 1.5, Equation 1.6 only uses the period immediately prior to adoption, $t - \ell - 1$, as the reference point to calculate the effect of treatment.

[3] Our notation is slightly different than the estimator presented in de Chaisemartin and d'Haultfoeuille (2022) because we do not allow reversibility of treatment.

To account for annual precipitation and evapotranspiration, we also incorporate controls into the DID_ℓ estimator. To do so, an OLS regression is estimated of the outcome evolution within the control group, $(Y_{i,g,t} - Y_{i,g,t-\ell-1} \mid g > t)$, on the evolution of its time-varying covariates during the same period, $(X_{i,g,t} - X_{i,g,t-\ell-1} \mid g > t)$, and time fixed effects as in de Chaisemartin and d'Haultfoeuille (2022). Letting $\hat{\theta}_0$ indicate the estimated coefficients for $(X_{i,g,t} - X_{i,g,t-\ell-1})$ from this regression, the change in outcomes we would expect given the treatment group's evolution of covariates is subtracted from the outcome evolution terms to remove the variation in outcomes due to changes in the covariates:

$$\widehat{DID}_{t,\ell}^{X} = \frac{1}{N_{t,\ell}^{1}} \sum_{\forall\, i \in\, g\, =\, t-\ell} \left(Y_{i,g,t} - Y_{i,g,t-\ell-1} - (X_{i,g,t} - X_{i,g,t-\ell-1})'\hat{\theta}_0\right) - \frac{1}{N_t^{nt}} \sum_{g=t+1} \left[\sum_{i=1}^{N_{g,t}} \left(Y_{i,g,t} - Y_{i,g,t-\ell-1} - (X_{i,g,t} - X_{i,g,t-\ell-1})'\hat{\theta}_0\right)\right]. \qquad \text{(Eq. 1.7)}$$

Notably, the inclusion of covariates in the $DID_{t,\ell}^{X}$ estimator from Equation 1.7 requires the same conditional parallel trends assumption as the two-way fixed effects regressions above (de Chaisemartin and D'Haultfoeuille 2020). Namely, the conditional common trends assumption requires that any differential trends in the outcome variable between groups are explained with a linear model in $X_{g,t} - X_{g,t-\ell-1}$, where the terms $X_{g,t}$ and $X_{g,t-\ell-1}$ contain the time-varying covariates for the reference groups in the period immediately prior to adoption, $t - \ell - 1$, and at time t. The covariates we include when employing this estimator are the average precipitation received and evapotranspiration experienced during the pre-season and growing season months at the locations where water right groups irrigate.

Until this point, we focused on identifying the effect of ℓ years of treatment for the group which first adopted the new technology at time $t - \ell$. Now, we turn our attention to aggregating the individual $DID_{t,\ell}^{X}$ estimates into an average effect for each length of treatment, ℓ, and the overall effect across all treatment durations. First, we construct a weighted average of

$DID_{t,\ell}^{X}$ terms for each value of ℓ to generate event-study style estimates of the effect of using the new technology for ℓ years including the instantaneous treatment effect. To generate the weights, we define N_{ℓ}^{1} as the number of irrigators across all groups who are observed for at least ℓ years after adopting the new technology. Defining $\tilde{T}$ as the last period in which we observe an irrigator who still has not adopted the newer technology, we construct a weighted average of the effect of ℓ years using the new technology regardless of when the adoption decision was first made,

$\widehat{DID}_{\ell}^{X} = \frac{1}{N_{\ell}^{1}} \sum_{t=\ell+2}^{\tilde{T}} N_{t,\ell}^{1} DID_{t,\ell}^{X}.$ (Eq. 1.8)

The weight for each $DID_{+,t,\ell}^{X}$ in Equation 1.8 is the number of irrigators in groups who adopted the technology ℓ years ago relative to period t, $N_{t,\ell}^{1}$, divided by the total number of irrigators who have a recorded outcome ℓ years after adopting the new technology, N_{ℓ}^{1}.

To recover the overall average treatment effect, we then take a weighted average of these $\hat{\delta}_{\ell}^{CD}$, or DID_{ℓ}^{X}, values across all values for ℓ. The weights are the number of irrigators in groups who are observed ℓ years after adoption divided by the total number of irrigators who adopt the technology at any point, $w_{\ell} = N_{\ell}^{1} / \sum_{\ell} N_{\ell}^{1}$. The overall average treatment effect can be expressed as:

$\hat{\delta}^{CD} = \sum_{\forall \ell \geq 0} w_{\ell} DID_{\ell}^{X},$ (Eq. 1.9)

where the CD superscript denotes that it is the de Chaisemartin and D'Haultfoeuille (2022) estimator.

There are four assumptions necessary for the individual $\widehat{DID}_{t,\ell}^{X}$ estimators to recover the effect of adopting the technology ℓ years before period t, and for the resulting aggregation, $\hat{\delta}^{CD}$, to be an unbiased estimator of the average treatment effect of adopting the new technology. First, the treatment design must be "sharp" so the sequence of treatment across all irrigators in a group is the same. We express this in the example earlier by stating $D_{i,g,t} = D_{g,t}$. The second assumption, non-pathological design, requires there to be at least one cohort of irrigators who adopt the new technology at a time when another cohort continues to use the old technology. The first two assumptions are satisfied with our data.

The third assumption, that of no anticipation, requires there be no dependence between a group's untreated outcome and its future treatments. In our context, this means that we assume irrigators do not adjust their water use with the old technology in anticipation of adopting the more efficient technology. The fourth assumption requires the trajectories of the counterfactual, or never adopter, outcomes remain the same across all cohorts after accounting for the impact of covariates with a linear model in $X_{i,g,t} - X_{i,g,t-\ell-1}$. Using $Y_{i,g,t}(D_{T+1})$ to indicate the counterfactual outcome for groups with $g \leq T$, this assumption requires there to be a vector θ_0 so that within each group and time period cell, (g, t), the following does not change across groups for $t \geq 2$:

$$E\left[Y_{i,g,t}(D_{T+1}) - Y_{i,g,t-1}(D_{T+1}) - (X_{i,g,t} - X_{i,g,t-\ell-1})'\hat{\theta}_0 | D_{i,g,t}, X_{i,g,t}\right]. \quad \text{(Eq. 1.10)}$$

Equation 1.10 expresses a specific version of conditional parallel trends assumption, where groups can have differences in their counterfactual behavior if it is explained by the change in their covariates over time. This expression also requires the adoption behavior and counterfactual outcomes for different groups of irrigators to be independent, and that any random shocks affecting a group's never-treated outcome be mean independent of the group's treatment sequence. In our context, this allows water right group's counterfactual outcomes and treatment trajectories to differ so long as it is explained by the change in weather between time periods. However, we expect differential trends in water right groups' behavior due to heterogeneity in their time invariant characteristics, such as total irrigated acreage, to violate this assumption. As such, we create five sets of water right groups by quintile of groundwater withdrawn in the first year water right groups appear in each sub-sample dataset. We then use the non-parametric matching feature of de Chaisemartin et al.'s *did_multiplegt* STATA package to only compare outcomes within these quintiles when estimating $\hat{\delta}^{CD}$ for each of three technology changes (de Chaisemartin et al. 2019).

1.4.3 Callaway and Sant'Anna

The approach outlined in Callaway and Sant'Anna (2020) is similar to that of de Chaisemartin and D'Haultfoeuille (2022) in its focus on what Callaway and Sant'Anna (2020) call "group-time average treatment effects." For our context, each group-time average treatment effect, $ATT(g,t)$, represents the average time-specific effect of adopting the more efficient irrigation technology on a cohort of individuals who adopted the technology at the same time. A group-time average treatment effect is akin to the result produced by estimating Equation 1.9 for a single value of g and ℓ. In fact, in the absence of covariates and when the not-yet-treated groups are used as controls, the estimator of de Chaisemartin and D'Haultfoeuille (2022) and outcome regression version of Callaway and Sant'Anna's (2020) estimator are numerically equivalent. While we use the same not-yet-treated groups as controls for both estimators, indicated by the superscript ny, our inclusion of covariates and use of the doubly robust version of Callaway and Sant'Anna's (2020) estimator necessitates a discussion of the distinctions between these estimators and their relative advantages.

As shown above, the $DID_{t,\ell}^{X}$ estimator in Equation 1.7 accounts for covariates using linear regression and accommodates the inclusion of time-varying covariates in the post-adoption period. In contrast, Callaway and Sant'Anna (2020) use a non-parametric approach to address differential trends attributable to covariates and do not include post-treatment covariates. Furthermore, one of the advantages of Callaway and Sant'Anna's (2020) approach is that it can be estimated using a doubly robust estimator, indicated by the subscript dr, meaning the average treatment effect will be recovered if either the treatment effect evolution or propensity score models are properly specified (Sant'Anna and Zhao 2020). Constructing the $ATT_{dr}^{ny}(g,t)$ estimator is a two-step process beginning with estimating the population outcome regression, $m_{i,g,t}^{ny}(Z_{i,g})$, and the propensity score, $p_{i,g,t}(Z_{i,g})$.

Similar to the $DID_{t,\ell}^{X}$ estimator, the purpose of the outcome regression component, $m_{i,g,t}^{ny}(Z_{i,g})$, is to remove the portion of the change in observed outcomes due to covariate driven differential trends. For a given value of g and t, the change in outcomes for the control group from $t = g - 1$ to $t = t$, or $\left(Y_{i,g,t} - Y_{i,g,g-1} \mid g > t\right)$, is regressed on the time *invariant* covariates for the control group, $Z_{i,g}$. Letting $\hat{\beta}_{i,g,t}^{ny}$ be the vector of coefficients resulting from this regression, then $\widehat{m}_{i,g,t}^{ny}\left(Z_{i,g}; \hat{\beta}_{i,g,t}^{ny}\right)$ is the predicted change in outcomes from time $t = g - 1$ to $t = t$ for treated irrigators in group g due to covariate-driven differential trends. Next, the propensity score component to the estimator, $p_{i,g,t}(Z_{i,g})$, is defined as the probability of irrigator i being in the group first treated at time $t = g$ rather than not-yet-treated at time t conditional on the time invariant covariates, $Z_{i,g}$. Letting N indicate the total sample size and defining $G_{i,g}$ as an indicator variable equal to one if irrigator i is first treated in period g, the parametric estimators of the outcome regression and propensity score functions, $\hat{p}_{i,g,t}\left(Z_{i,g}; \hat{\pi}_{i,g,t}\right)$ and $\widehat{m}_{i,g,t}^{ny}(Z_{i,g}; \hat{\beta}_{i,g,t}^{ny})$, are then used to calculate the group-time average treatment effects for all groups as follows:

$$\widehat{ATT}_{dr}^{ny}(g,t) = \frac{1}{N} \sum_{\forall i} \left[\left(\widehat{w}_g^{treat} - \widehat{w}_g^{comp,ny} \right) \left(Y_{i,g,t} - Y_{i,g,g-1} - \widehat{m}_{g,t}^{ny}\left(Z_{i,g}; \hat{\beta}_{g,t}^{ny}\right) \right) \right] \quad \text{(Eq. 1.11)}$$

where

$$\widehat{w}_g^{treat} = \frac{G_{i,g}}{\frac{1}{N}\sum_{\forall i}\left(G_{i,g}\right)}; \; \widehat{w}_g^{comp,ny} = \frac{\frac{\hat{p}_{i,g,t}\left(Z_{i,g};\hat{\pi}_{i,g,t}\right)\left(1-D_{i,g,t}\right)\left(1-G_{i,g}\right)}{1-\hat{p}_{i,g,t}\left(Z_{i,g};\hat{\pi}_{i,g,t}\right)}}{\frac{1}{N}\sum_{\forall i}\left(\frac{\hat{p}_{i,g,t}\left(Z_{i,g};\hat{\pi}_{i,g,t}\right)\left(1-D_{i,g,t}\right)\left(1-G_{i,g}\right)}{1-\hat{p}_{i,g,t}\left(Z_{i,g};\hat{\pi}_{i,g,t}\right)}\right)}.$$

In Equation 1.11, the doubly robust specification of Callaway and Sant'Anna's (2020) estimator clearly displays the two methods of addressing covariate-driven non-parallel trends. The first method is demonstrated in the inverse probability weighting component that subtracts $\widehat{w}_g^{comp,ny}$ containing $\hat{p}_{i,g,t}\left(Z_{i,g};\hat{\pi}_{i,g,t}\right)$ terms from the population weight. As demonstrated in Abadie (2005), the inverse probability weighting approach to estimating the treatment effects in a DID setting compares the outcomes of treatment and control groups using the probability of treatment conditional on covariates. The second method, subtracting $\widehat{m}_{i,g,t}^{ny}(Z_{i,g};\hat{\beta}_{i,g,t}^{ny})$ from the outcome evolution, uses the outcome evolution for control group observations with matching covariates to address differential trends in the outcomes (Heckman et al. 1997). For our estimation, we use the default option for Callaway and Sant'Anna's *att_gt* R package which employs the inverse probability tilting estimator of (Graham et al. 2012) to estimate the propensity scores and weighted-least squares to estimate the outcome regressions. The covariates we include in $Z_{i,g}$ are the pre-development saturated thickness and hydraulic conductivity of the aquifer at each water right group's wells, the available water capacity and soil texture where the water right group irrigates, and the acre-feet the water right group applied in the first year they appear in each sub-sample dataset.

As with the individual $DID_{t,\ell}^{X}$ estimates, the individual $\widehat{ATT}_{dr}^{ny}(g,t)$ estimates must be aggregated to generate an estimate of the average treatment effect, $\hat{\delta}^{CS}$. For a given value of g, we first create the following average treatment effect for each cohort, $\hat{\delta}^{CS}(g)$, by averaging the treatment effects across all values of $t \geq g$ and dividing by the maximum number of periods an irrigator in group g could be treated for:

$\hat{\delta}^{CS}(g) = \frac{1}{T-g+1} \sum_{t=g} \widehat{ATT}^{ny}_{dr}(g,t).$ (Eq. 1.12)

Then, we take a weighted average of the $\hat{\delta}^{CS}(g)$ values across all cohorts to produce the final estimate of the ATE. Letting N_g indicate the number of observations across all irrigators in cohort g, the weights for this final aggregation are the probabilities of belonging to a given cohort conditional on being in a treated cohort,

$\hat{\delta}^{CS} = \frac{1}{\sum_{g=1}^{T} N_g} \sum_g N_g * \hat{\delta}^{CS}(g).$ (Eq. 1.13)

We employ one other aggregation so we can plot event-study style estimates of dynamic treatment effects analogous to the DID^X_ℓ estimates outlined earlier. For the following equation, we use the same definition for ℓ as earlier, the number of years since an irrigator adopted the new technology. Additionally, we use $N_{g,\ell}$ to indicate the number of irrigators in cohort g who are observed for ℓ periods after the is adopted. The following equation for $\theta^{CS}(\ell)$ estimates the effect of ℓ years using the new technology:

$\hat{\theta}^{CS}(\ell) = \frac{1}{\sum_{g=1}^{T-\ell} N_{g,\ell}} \sum_{g \in G} N_{g,\ell} * \widehat{ATT}^{ny}_{dr}(g, g+\ell)$ (Eq. 1.14)

There are five necessary assumptions for the individual $\widehat{ATT}^{ny}_{dr}(g,t)$ estimates and the aggregations we use to identify their respective effects. The first two assumptions concern the design of the pseudo experiment, imposing restrictions on the treatment variable and sampling process. First, treatment must be irreversible, meaning no irrigator can adopt the new technology and then switch back to the earlier technology in a later time period. Second, the panel dataset must be representative of the overall population.

The next three assumptions, the identifying assumptions, are concerned with establishing parallel trends between adopters and never adopters. The third assumption requires the temporal extent of behavior in anticipation of adopting the new technology to be pre-specified. So, irrigators can adjust water use behavior in anticipation of adopting the technology in the future, but we would need to specify how many years prior to adoption this anticipation behavior occurs. To remain consistent with the DID_ℓ estimator, we assume there is no anticipation behavior.

The fourth identifying assumption, the conditional parallel trends assumption, differs slightly from that of the DID_ℓ estimator. For the $\widehat{ATT}_{dr}^{ny}(g,t)$ estimator, the conditional parallel trends assumption requires the expected counterfactual outcome evolution for adopters to equal the expected outcome evolution of irrigators who have not yet adopted by the end of the period in question, where both outcome evolutions are conditioned on pre-treatment covariates.

The fifth assumption, known as the "overlap" assumption, is required to take advantage of the doubly robust version of Callaway and Sant'Anna's (2020) estimator and is commonly employed in difference-in-differences literature (Sant'Anna and Zhao 2020). In essence, the overlap assumption requires there to be adopters and not yet adopters with similar propensity scores so that the control group's outcome evolution is truly reflective of the adopters' counterfactual outcome. As outlined above, the $\hat{\delta}^{CS}$ estimator matches water right groups based on pre-treatment covariates, and then uses these matched controls to generate counterfactual outcomes. In contrast, the $\hat{\delta}^{CD}$ estimator uses the evolution of the outcome variable amongst all control observations in the same quintile of initial withdrawals conditional on the evolution of time-varying covariates. As such, the $\hat{\delta}^{CD}$ estimator may end up using water right groups who are dissimilar to the treated groups as controls because they share similar time-varying covariate behavior as the treated units. So, while the $\hat{\delta}^{CS}$ estimator requires an additional assumption, it is our preferred specification because of its doubly robust properties and its method of constructing counterfactual outcomes.

1.5 Results and Discussion

First, we show overall average treatment effects for each of the estimators on total withdrawals. We also decompose the effect on total withdrawals into the effect on acres irrigated (i.e., extensive margin) and depth applied (i.e., intensive margin). Then we address the differences in results between TWFE and more recent DID methods by showing the estimated dynamic treatment effects and discuss how the necessary assumptions for the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimators contribute to any disparities. In addition to illustrating a source of bias in the TWFE results, we also examine the dynamic treatment effects because they provide insight into how irrigators adapt over time to a technology transition.

1.5.1 Average treatment on the treated estimates

The overall average treatment effect estimates for each of the three transitions are plotted in Figure 1.5. Each panel in Figure 1.5 shows results across the three different estimators for a specific technology conversion and across all dependent variables. The panels of Figure 1.5 are oriented such that each row shows the results for one of the three different types of irrigation technology conversions and the columns show results for different dependent variables. Note that the impact on total withdrawals does not equal the sum of the impact on acres irrigated and depth applied.

Before we examine each transition's impact on the three dependent variables individually, it is worth noting the difference in standard errors between the three estimators.

While they are applied to the same sub-samples of the data and clustered at the same level, the water right group, the standard errors for the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimates are generally larger than those of the TWFE estimator. This is the consequence of these estimators' accommodating cohort specific treatment trajectories and limiting the number of pre-treatment periods used as reference periods. In contrast, the TWFE specification in Equation 1.5 pools all treated observations across cohorts and uses all pre-treatment periods.

As depicted in the top row of panels in Figure 1.5, all three estimators indicate that the change from flood to center pivot leads to an increase in irrigated acres and a decrease in depth applied. While our estimates are slightly smaller in magnitude, the significant decrease in depth-applied due to irrigators switching from flood into center pivot irrigation is consistent with prior empirical research (Hendricks and Peterson 2012). When considered in combination with the negative or statistically insignificant impact on total withdrawals, this indicates producers are using the water saved due to center pivot systems' increased efficiency to bring additional acreage into production at a comparatively lower level of irrigation intensity. This is feasible given the relatively large difference in efficiency between flood and center pivot irrigation systems, as well as the capability to install center pivot systems in areas where the topography would make flood irrigation impossible (NRCS 1997).

The TWFE estimator produces statistically significant estimates of Jevon's paradox when water right groups adopt LEPA in center pivot irrigation systems, similar to the results found by Pfeiffer and Lin (2014) and the estimates for irrigators with moderately sized water rights from Li and Zhao (2018). As shown in the second row of panels in Figure 1.5, the TWFE result on total withdrawals is driven by an increase in depth applied rather than a change in irrigated acres. While previous literature attributed some of the increase in withdrawals to an expansion at the extensive margin which is not observed here, the increase in depth-applied suggested by the TWFE results is consistent with their arguments that producers switched to more water-intensive crops (Pfeiffer and Lin 2014; Li and Zhao 2018). However, neither of the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimators support this conclusion. Additionally, the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ produce small and statistically insignificant estimates for the effect of adoption on acre-feet withdrawn.

For the final change in irrigation technology, removal of endguns from center pivot systems, we find a mix of results. In comparison to the other technology transitions, there are relatively few water right groups who make the transition in this sub-sample so the standard errors are comparatively larger. Despite larger standard errors, the $\hat{\delta}^{CS}$ estimates for the effect of removing endguns on acres irrigated and depth-applied are statistically significant. The decrease in irrigated acreage suggested by the $\hat{\delta}^{CS}$ estimator is slightly lower than expected, as endguns typically account for a tenth of the total irrigated acreage for a center pivot system irrigating a full circle (Martin et al. 2017). Aside from the issues posed by our small sample size, differences in the model of endguns used, configuration of the center pivot systems, or how irrigators calculate the coverage of endguns could explain why we do not observe a greater decrease in irrigated acreage. The significant increase in depth-applied indicated by the $\hat{\delta}^{CS}$ estimator suggests irrigators apply similar quantities of groundwater as they did prior to removing endguns but to a smaller area after the transition. Lastly, while the effects are statistically insignificant, all the estimators produce positively signed estimates of the impact of endgun removal on withdrawals. Removing the endgun from a center pivot system reduces the required operating pressure for the system. So, similar to the advantages provided by LEPA systems, removing endguns can increase the efficiency of an irrigation system by reducing air water loss, decreasing canopy loss, and through changes to pump operation (Rogers et al. 2017). If these factors significantly decrease the marginal cost of applying groundwater, it may become profitable for irrigators to re-allocate the groundwater previously applied to the corners toward the interior of the center pivot system.

In summary, we do not find unanimous evidence of Jevon's paradox occurring in any of the three irrigation technology transitions we consider. If there were such evidence, we would expect both the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimates in Figure 1.5 to indicate an increase in total withdrawals after a technology change because of their ability to accommodate staggered adoption and heterogeneous treatment effects. For the conversion from flood to center pivot irrigation and adoption of LEPA, the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimates of the effect of changing technologies on acre-feet withdrawn have signs opposite one another and are statistically insignificant. Finally, while both the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimates for the impact of endgun removal on acre-feet withdrawn are positive, they are statistically insignificant. In the next section, we explore the dynamic effects of each technology adoption to investigate how producers adapt to the new technologies over time and illustrate why the estimates of the average treatment effect in Figure 1.5 differ between TWFE the two estimators which accommodate staggered adoption, $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$.

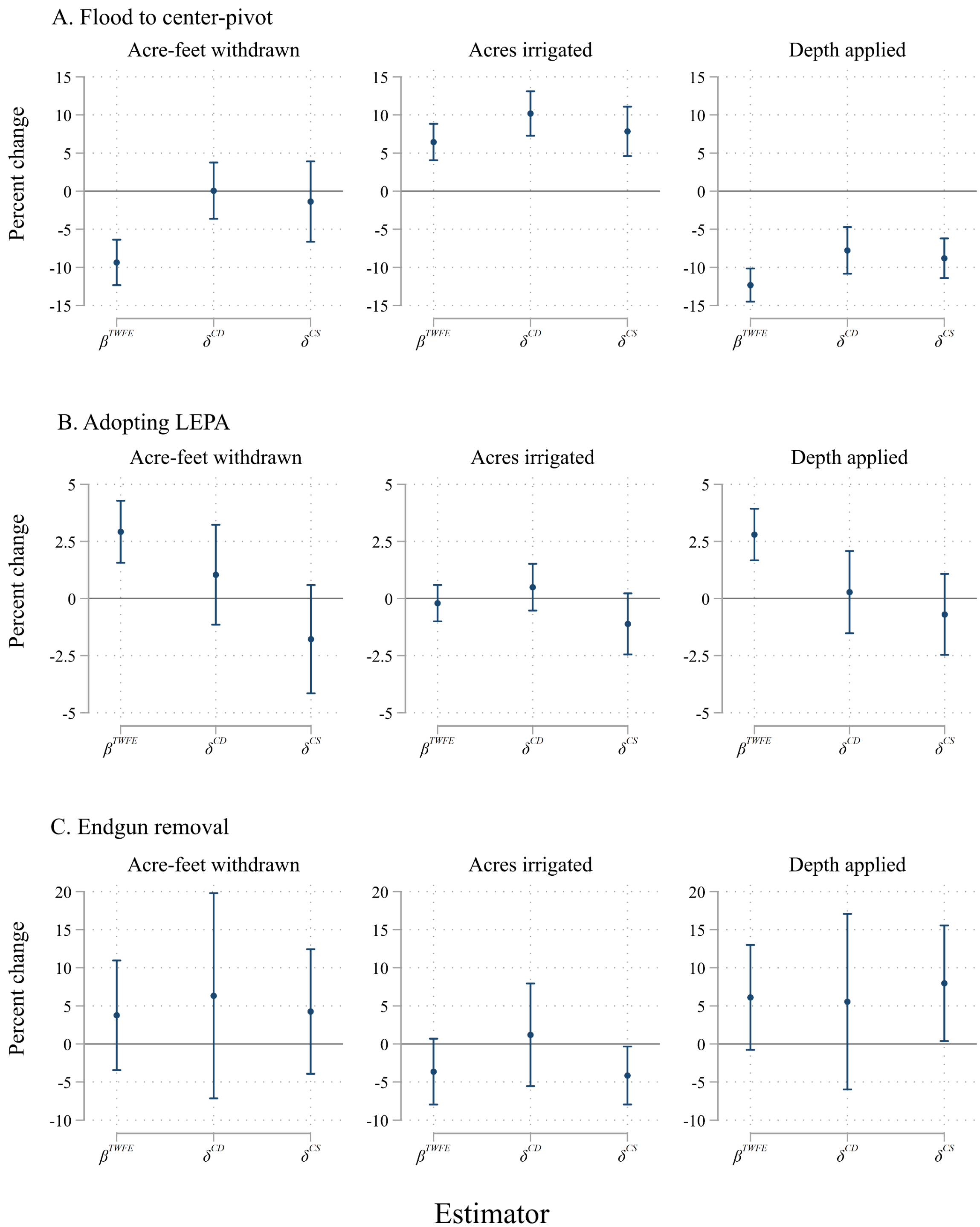

Figure 1.5. Estimated effect of adopting each technology across all three estimators. Treatment effects are expressed as a percent change relative to the sample mean of each dependent variable.

1.5.2 Dynamic treatment effects

The average treatment effects for the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimators displayed in Figure 1.5 are a weighted average of the group-time average treatment effects across all groups. In Figures 1.6, 1.7 and 1.8, we display the effect of changing from flood to center pivot irrigation, incorporating LEPA into a center pivot system, and removing endguns as a function of the time since conversion. By examining how the effect of changing technologies evolves over time, the dynamic treatment effects, we can understand the differences between the ATT estimates in Figure 1.5 and investigate how irrigators adjust to each new irrigation technologies over time. The estimators used to generate the dynamic effect estimates, labelled $\hat{\delta}^{CD}(l)$ and $\hat{\delta}^{CS}(l)$ for brevity, are expressed in Equations 1.8 and 1.14 respectively. As in Figure 1.5, the dynamic effects displayed in Figures 1.6, 1.7, and 1.8 are expressed as a percent change relative to the mean values of the dependent variables in the pertinent sub-sample. For both estimators, the null effect indicated for the period one year before adoption is the product of using the year prior to adoption as the reference year, meaning it does not have any statistical significance.

While slightly different than our theoretical case, the dynamic effects of switching from flood to center pivot irrigation in Figure 1.6 clearly pose an issue for TWFE given the staggered adoption we observe. Both the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimates of the instantaneous effect of the conversion, the effects at time $t = 0$, indicate a roughly 10% reduction in groundwater withdrawals in the first year water right groups switched to center pivot irrigation. While they do not consider dynamic impacts, Peterson and Ding (2005) predict a similarly sized reduction in withdrawals when evaluating the change from flood to center pivot irrigation with a risk-programming model. In addition to the significant instantaneous effect we observe, the $\hat{\delta}^{CS}$ estimator indicates the effects one and two years after adoption are significant reductions in groundwater withdrawals of around 5 percent. However, the effect of converting from flood to center pivot becomes increasingly positive over time, such that at time $t = 9$ the $\hat{\delta}^{CD}$ estimator suggests a 10% increase in groundwater withdrawals relative to the control group. While the $\hat{\delta}^{CS}$ estimates of the dynamic effects also increase over time, none of them indicate a significant increase in withdrawals. Overall, the evolution of groundwater withdrawals depicted by both estimators in Figure 1.6 provides a clear explanation for the difference between the significant reduction in withdrawals suggested by TWFE and the insignificant effects for the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimators in Figure 1.5. If the behavior of an irrigator who adopted center pivot irrigation two years ago is used to create the counterfactual for an irrigator making the change from flood this season, the immediate decrease in withdrawals the new adopter exhibits would be unduly magnified by the slight increase in withdrawals suggested by the faulty counterfactual.

For the adoption of LEPA into center pivot irrigation systems in Figure 1.7, the $\hat{\delta}^{CS}(l)$ estimates of the dynamic treatment effects indicate water right groups slowly decrease their groundwater withdrawals over several years following adoption. The estimated effect ten years after adopting LEPA is a statistically significant 6% reduction in groundwater withdrawals. This gradual reduction over time, in combination with the staggered nature of adoption, is the reason for the significant increase in groundwater withdrawals estimated by the TWFE model. We see a similar pattern in the $\hat{\delta}^{CS}(l)$ estimates for depth-applied, with the effect seven years after adoption being a two percent reduction. The majority of estimates for both estimators, however, are statistically indistinguishable from zero. One exception to this is the effect of adopting LEPA on withdrawals two years prior to adoption estimated using the $\hat{\delta}^{CS}(l)$ estimator. We test for evidence against our parallel trends assumptions, for every other transitions as well, by performing a placebo test which verified the pre-treatment coefficients for both the $\hat{\delta}^{CS}(l)$ and

$\hat{\delta}^{CD}(l)$ estimators are not jointly different than zero. The other reason for highlighting this result is to illustrate the difference in control groups produced by each estimator's method of constructing counterfactual outcomes. Again, because the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CS}(l)$ estimators model both the probability of adoption and the outcome evolution, we believe they provide the best estimates of the average treatment effect and dynamic treatment effects.

Finally, as depicted in Figure 1.8, the removal of endguns from center pivot systems is the one technology transition for which the $\hat{\delta}^{CS}(l)$ estimator displays some evidence of Jevon's paradox occurring. Four years after removing endguns, the $\hat{\delta}^{CS}(l)$ estimate suggests irrigators who removed endguns from their center pivot systems exhibited a 12% increase in withdrawals in comparison to the control group. However, this effect becomes insignificant five years after adoption, so we do not treat it as conclusive evidence of Jevon's paradox. The increasingly positive impact of endgun removal on depth-applied produced by the $\hat{\delta}^{CS}(l)$ estimator in Figure 1.8 explains the significant, positive average treatment effect produced by its analogous $\hat{\delta}^{CS}$ estimator in Figure 1.5. The $\hat{\delta}^{CS}(l)$ estimates indicate irrigators increase depth-applied by 10% three years after endgun removal, and this effect persists for the fourth and fifth years as well. As hypothesized earlier when discussing the average treatment effect, these results indicate that the water irrigators previously applied using endguns is re-allocated to the inner spans of the center pivot system. In contrast to the large instantaneous effect observed when adopting center pivot systems, the dynamic treatment effects due to removing endguns on depth-applied depict a gradual adaptation over several years.

For both the flood to center pivot and center pivot adopting LEPA transitions, the dynamic treatment effects displayed in Figures 1.6 and 1.7 underscore the importance of using estimators that accommodate staggered adoption in this context. In both cases, the differences between the biased TWFE results and those of the $\hat{\delta}^{CS}$ and $\hat{\delta}^{CD}$ estimators for the average treatment effects concerning withdrawals are due to the TWFE specification improperly using early adopters as controls for late adopters. For the adoption of LEPA into center pivot systems, TWFE produces biased estimates of the wrong sign due to the evolution of treatment effect dynamics as described in the simulated scenario earlier. In the case of flood irrigators switching to center pivot irrigation, the combination of an immediate reduction and a long-term slow increase in withdrawals leads to TWFE overestimating the impact of the conversion. In addition to highlighting the shortcomings of an improper TWFE specification when staggered adoption and dynamic treatment effects are present, the dynamic treatment effects in each case also reveal how irrigators adjust to changes in irrigation technology over time. For example, in the conversion of flood to center pivot irrigation, we observe a large change in behavior immediately after the transition in technologies which is then followed by a gradual adjustment at the extensive margin. In comparison, there are few perceptible changes in irrigators' behavior when adopting LEPA into center pivot systems, suggesting any adjustments to the new technology are minute in magnitude.

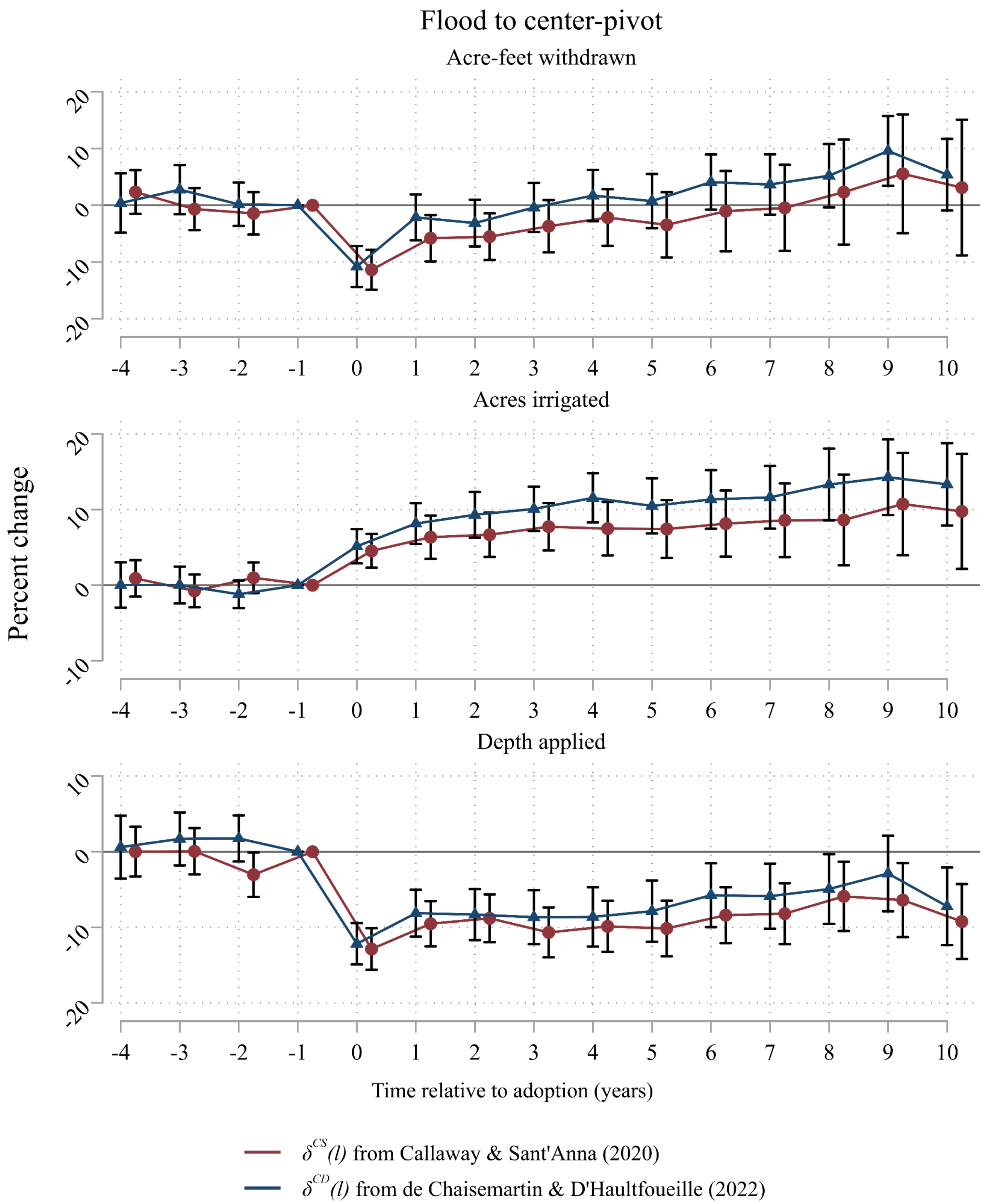

Figure 1.6. Dynamic treatment effects due to converting from flood to center pivot irrigation. Effects are expressed as a percent change relative to the sample mean of each dependent variable.

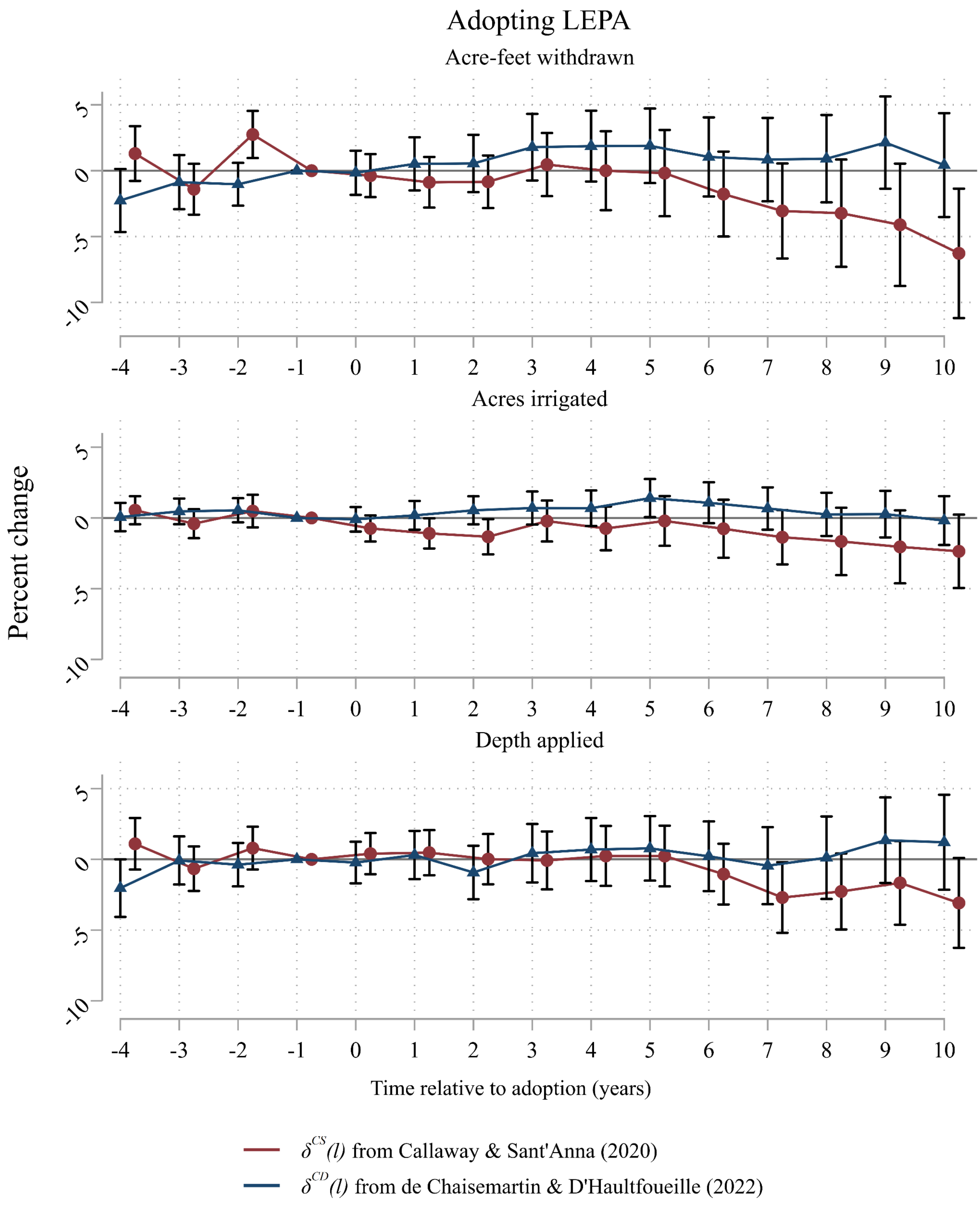

Figure 1.7. Dynamic treatment effects due to adoption of LEPA into center pivots systems. Effects are expressed as a percent change relative to the sample mean of each dependent variable.

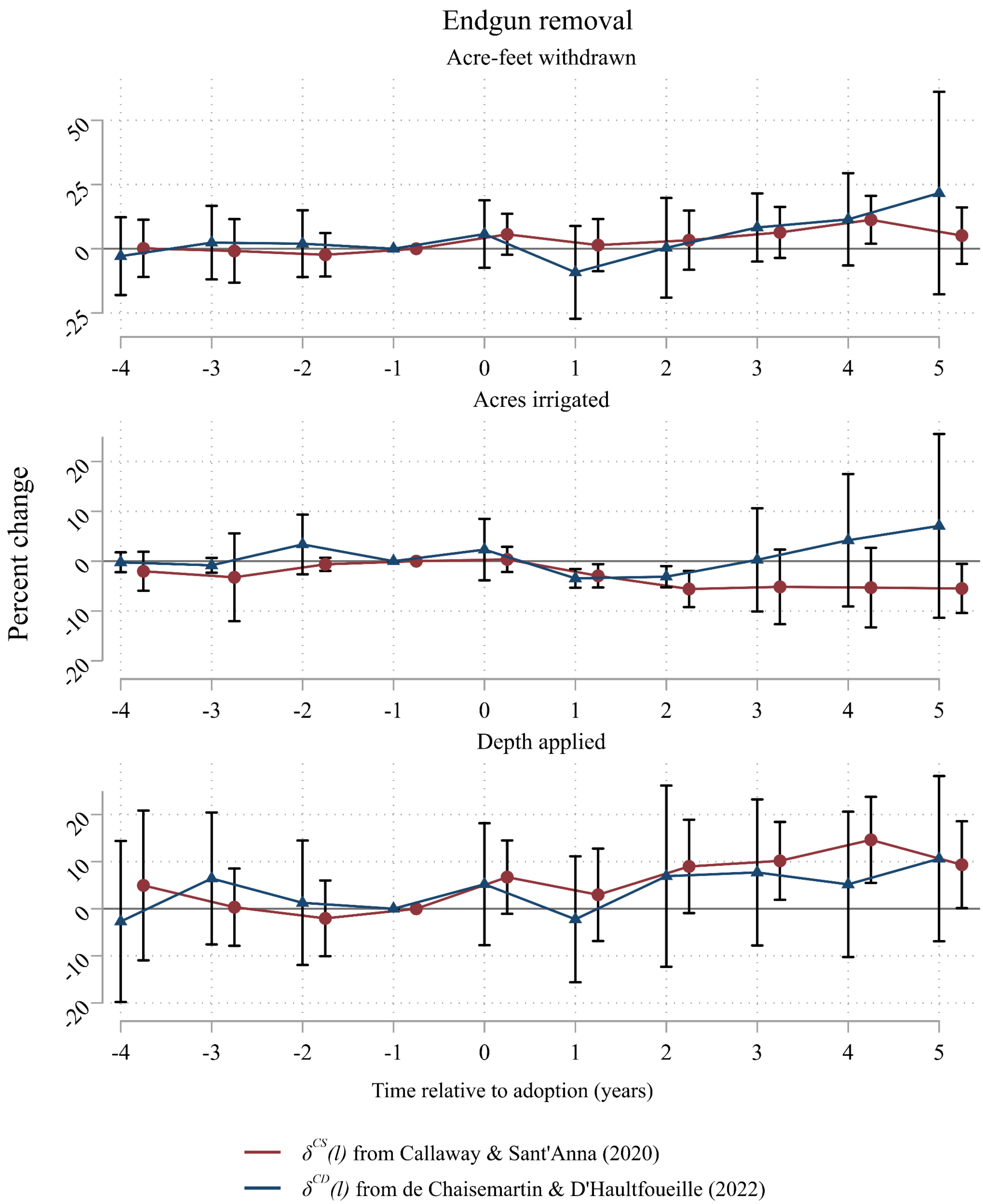

Figure 1.8. Dynamic treatment effects due to removing endguns from center pivot systems. Effects are expressed as a percent change relative to the sample mean of each dependent variable.

1.6 Conclusion

We do not find evidence of rebound effects in great enough magnitude to produce Jevon's paradox in the three irrigation technology changes we investigate in Kansas. The possibility of such a feedback loop by which increasing efficiency encourages additional resource use is certainly cause for concern, but the empirical results presented here do not substantiate claims of its occurrence in the groundwater context. Instead, we find the conversion from flood to center pivot irrigation initially reduces groundwater withdrawals and leads to a persistent reduction in the intensity of irrigation, or depth-applied. However, this reduction in depth applied is offset by an increase in irrigated acres. For the conversion from traditional center pivot to LEPA, we find the average treatment effect is indistinguishable from zero and limited evidence of a reduction in groundwater withdrawals ten years after adoption. As in our simulated example, these results demonstrate how TWFE estimates can generate false evidence of Jevon's paradox when the staggered nature of adoption is not addressed. Finally, when irrigators remove endguns from their center pivot systems, we find that they reduce their irrigated acreage as expected. However, the significant increase in depth applied suggests any water saved in contracting irrigated acreage is applied to the smaller remaining acreage.

In comparing the average and dynamic treatment effects across all three technology adoptions, we do not find a consistent impact of increasing irrigation efficiency on groundwater use. Instead, our results highlight the heterogeneous impacts of different technologies. For some changes in technology, such as from flood to center pivot irrigation, there are significant reductions in groundwater withdrawals and the intensity of irrigation despite increasing irrigated acreage. Of the three transitions we study, switching from flood to center pivot irrigation involves the greatest reduction in evaporative losses and the largest difference in operation.

However, while the effect of converting from flood to center pivot irrigation produces near-term reductions in withdrawals and irrigation intensity, the flood irrigation systems in use today are likely to differ substantially from those in our dataset used by Kansas producers two or three decades ago. For the other two conversions, adoption of LEPA and removal of endguns for center pivot systems, the primary irrigation technology remains consistent and the efficiency gains due to reduced evaporative losses are smaller in magnitude. While the adoption of LEPA shows some evidence of leading to reductions in groundwater withdrawals as irrigators adjust over time, the removal of endguns may have the opposite effect. These heterogenous impacts, in combination with the hydrological and institutional concerns omitted from this work, necessitate situation specific analyses when considering irrigation technology improvements as a means of addressing water scarcity concerns.

Finally, while irrigation technology changes with small efficiency gains may not reduce groundwater withdrawals, this does not mean they are incongruent with the goals of sustainable intensification. While we do not consider the impact on yields in this work, these technologies may increase the quantity of crops produced per unit of groundwater extracted. For example, as is suggested by the dynamic effects of LEPA adoption, they may facilitate long-term adaptation that allows producers to maintain yields while reducing groundwater use over time. Alternatively, such as for the removal of endguns, they could increase the yield produced per unit of water or mitigate risk by reallocating the same quantity of groundwater to better use in a smaller area. For non-renewable groundwater resources in aquifers without significant return flows, such improvements in the "crop-per-drop" may be preferable to conservation depending on the investment opportunities available and discounting involved.

Chapter 2 - Where will the United States farm carbon?

2.1 Introduction

In the United States, sellers of agriculturally produced carbon credits provide a diversified product as each program is defined by a unique combination of payment mechanism, verification process, quantification protocol, and contract design (Plastina and Wongpiyabovorn 2021). As such, buyers of agricultural carbon credits make purchase decisions based on the price and quality of available credits. One of the crucial determinants of carbon credit quality is additionality, or the degree to which the carbon sequestration would not have occurred in the absence of the carbon credit program. However, current carbon farming programs' monitoring, reporting, and verification processes focus primarily on confirming the adoption and continued use of conservation agriculture practices, rather than evaluating if producers' net return to practice adoption would have been positive even in the absence of the program (Plastina 2021).

The significance of the additionality challenge for carbon credit quality is highlighted by two prior works which conclude adoption of the two most frequently covered practices due to government incentive programs was only moderately additional. In a study of the impact of cost-share programs on Iowa producers' adoption of cover-cropping, Sawadgo and Plastina (2021) find 54% of cover-crop adoption was additional. Similarly, Claassen, Duquette, and Smith (2018) estimate that only 47% of conservation tillage adoption by participants in federally funded voluntary agricultural conservation programs was additional due to the practice's relatively high short run profitability. In discussing how their results differ from a smaller estimate of 25% found in an earlier work (Mezzatesta, Newburn, and Woodward 2013), Claassen et al. highlight how heterogeneity in local growing conditions and an area's past production history drive differences in producers' net returns to adopting a practice.

However, even if administrators of voluntary programs address concerns of non-additional practice adoption, ensuring additionality of practice adoption is inconsequential if adopting the relevant conservation agriculture practices does not increase sequestration, or prevent emission, of greenhouse gases. Ultimately, the quality of a carbon credit is determined by the degree to which it accurately represents additional greenhouse gases being sequestered or retained in the soil, and the same two factors impacting producers' net returns to practice adoption, local growing conditions and past production history, are also principal determinants of a soil's carbon sequestration potential (Bruce et al. 1999).

In this work, we demonstrate how the interrelationship between these two aspects of additionality, additional practice adoption versus greenhouse gas sequestration or retention, has significant implications for the geographic variability in quality of agricultural carbon credits produced across the United States agricultural landscape. Furthermore, this variation in quality will also impact how cost-effective investments in carbon farming or conservation agriculture practices are at meeting carbon sequestration targets across the country. In the subsequent section, we review the literature concerning the present state of the voluntary agricultural carbon credit market and the biophysical conditions that determine a soil's carbon sequestration potential. Then, following a description of the datasets we employ, we turn our attention to our methodological approach. In brief, we use a machine learning approach to predict the current rate of conservation agriculture practice adoption and use these predicted rates of adoption to proxy for the expected net return to practice adoption. After combining these predictions with county-level estimates of the carbon sequestration rates associated with each practice, we generate bivariate maps comparing the predicted rates of adoption with the carbon sequestration rates.

In the results and discussion section, we focus on three categories of counties defined by their predicted rates of adoption and carbon sequestration potential. Using these categories, we examine how future carbon credit programs could target regions based on their objective, be it generating credits at least cost or ensuring credit quality above all else. For example, producers in the Pacific Northwest would fall into the first category defined by high predicted rates of adoption but low carbon sequestration. While such producers may require a smaller incentive to adopt conservation tillage, these would be low quality, non-additional credits. Regions in the second category comprise the opposite case with a low predicted rate of adoption and high carbon sequestration potential. After discussing the unique challenges and opportunities posed by areas in the third category with high predicted rates of adoption and carbon sequestration, we conclude by discussing the implications of our results for the future of voluntary agricultural carbon credit markets and conservation programs in the United States

2.2 Literature Review

2.2.1 Current state of voluntary agricultural carbon credit programs

Current carbon farming programs compensate producers for carbon sequestration or offsets in one of two ways: paying for practice(s) adopted on a per-acre basis or paying for the carbon offset or sequestered on a per-unit basis (Plastina and Wongpiyabovorn 2021). While the agronomic practices eligible for participation in each type of program are similar, one of the principal transaction costs involved in carbon farming, the cost of verification, is much lower for payment-for-practice style programs. Verification involves confirming implementation of the practice on a per-acre basis for payment-for-practice programs, but for per-unit compensation programs it requires compiling the detailed records necessary for measuring and modeling the amount of carbon sequestered. For reference, a verifier can be the same entity selling the carbon credits, a single third-party entity, or various individuals who bid on the opportunity (Plastina 2021).

Despite the efficiency gains carbon credit purchasers can realize even when they bear the burden of verification in per-unit programs (Antle et al. 2003), producers bear some or all of the verification costs in several per-unit carbon farming programs at present (Plastina and Wongpiyabovorn 2021). As such, while carbon markets are in their nascent stage, the alternative payment-for-practice programs may play a significant role given that transaction costs are a major barrier to participation in carbon markets and depress supplies of carbon credits (Antle et al. 2007). For producers farming areas with lower carbon sequestration potential, the reduction in transaction costs associated with payment-for-practice programs may be the crucial determinant leading to their participation.

Finally, even if participating in a carbon farming program is insufficient to overcome the cost of adopting a practice by itself, many of the current programs allow producers to “stack” benefits by participating in other government of private conservation programs (Plastina and Wongpiyabovorn 2021). For example, a producer may receive assistance in making the conversion to no-till agriculture from a program, like the Environmental Quality Incentives Program (EQIP), and concurrently receive compensation from a carbon credit program for the carbon sequestration this conversion produces. While stacking programs could bolster participation in carbon credit programs, it also increases the inefficiency that occurs if producers’ adoption behavior is not additional. This possibility is especially salient given the $8.45 billion dollars appropriated for EQIP in the Inflation Reduction Act, a portion of which will be used to encourage adoption of conservation agriculture practices (117th Congress 2022).

2.2.2 Biophysical determinants of the carbon sequestration potential of cropland

Aside from the management practices carbon farming programs focus on, there are three main determinants of soil carbon sequestration potential: climate, soil characteristics, and the current stock of carbon (Bruce et al. 1999). Of the three, climate is the dominant driver as it determines growing season length and precipitation patterns, the principal drivers of primary productivity (Weil and Brady 2017). In general, soil carbon is more responsive to management in tropical and moist climates because the abundant precipitation and warm temperature increase net primary productivity in these regions (Ogle, Breidt, and Paustian 2005). In temperate regions, like the United States, precipitation is the more impactful climate driver because higher temperatures may reduce biomass production and increase the rate of carbon decomposition (Crowther et al. 2016; Delgado-Baquerizo et al. 2018).

After climate, the silt and clay content of a soil serves as another indicator of coil carbon sequestration potential (Sims and Nielsen 1986; Augustin and Cihacek 2016). These particles serve to protect soil carbon, thereby increasing the mineral-associated organic carbon (MOC) content (Cai et al. 2016). The final factor to consider is the current carbon content of the soil (Lal 2004). Soil carbon experiences logistic growth following adoption of a practice, with the marginal change in soil carbon being greatest five to twenty years after adoption and decreasing as it reaches its new equilibrium after fifty to seventy years (West and Post 2002; Paustian 2014). As such, producers farming presently degraded soils with high potential for carbon sequestration have more to gain in a payment-per-unit carbon program than producers farming land with high soil carbon levels and low potential for carbon sequestration.

2.3 Data

2.3.1 Data on adoption of conservation agriculture practices

The USDA census of agriculture in 2012 and 2017 recorded the number of acres where cover-cropping, no-till, and conservation tillage practices were used in most counties within the continental United States. To provide a sense of current use of these practices, we use these data and data on total cropland acreage from the USDA National Agricultural Statistics Service (NASS) to compute the share of cropland acres in each practice for 2012 and 2017. Figures A.1, A.2, and A.3 in the appendix display the change in the share of cropland employing cover-cropping, no-till, and conservation tillage practices over this five-year period. As we cannot determine whether the change in acreage using no-till is due to producers switching from conventional or conservation tillage, we combine the acreage under no-till and conservation tillage into one category representing the use of any practice with reduced tillage. Figure A.4 displays the chance in a county's cropland acreage using either no-till or conservation tillage between 2012 and 2017.

2.3.2 Data for carbon sequestration potential of practices

The COMET Planner tool contains county level estimates of the carbon sequestration potential for conservation agricultural practices (Swan et al. 2020). The values contained within represent county level aggregations of estimates from the COMET Farm tool, a field-scale platform for estimating carbon fluxes using the DayCent process-based model (Paustian et al. 2017). To create a single estimate for the adoption of no-till or conservation tillage practices, we take the mean of all county-level estimates for scenarios with USDA-NRCS Conservation Practice Standard (CPS) numbers 345 and 329. These scenarios are a change from intensive tillage to reduced tillage and a change from intensive tillage to no-till or strip-till respectively, and the mean expected sequestration in metric tons of carbon dioxide equivalents per acre per year for them is displayed in Figure A.5 in Appendix A.

To generate the expected carbon sequestration associated with cover-cropping, we take the mean of COMET Planner estimates for scenarios under USDA-NRCS CPS number 340. These scenarios include the addition of a legume seasonal cover crop with a 50% reduction in nitrogen fertilizer or the addition of a non-legume seasonal cover crop with a 25% reduction in nitrogen fertilizer, and the mean expected sequestration for them is depicted in Figure A.6 in Appendix A. In creating the average sequestration values for a reduction in tillage or adoption cover crops, we do not include scenarios where multiple practices are adopted jointly. For instance, we do not include scenarios involving the joint adoption of no-till and cover-cropping contemporaneously when generating the expected carbon sequestration associated with tillage reduction in a county. When combining the carbon sequestration data with the rates of adoption to create the final maps displayed in the results section, the carbon sequestration values are divided into three categories by tercile.

2.3.3 Soil, weather, and growing region characteristics

We draw daily datapoints for precipitation, maximum temperature, and minimum temperature from the Parameter-elevation Regressions on Independent Slopes Model repository maintained by Oregon State University (PRISM Climate Group 2014a). To allow for a non-linear relationship between the use of conservation agriculture practices and the weather in the past five years, we then construct temperature and soil moisture exposure variables for each month in the year as in Schlenker and Roberts (2009). Specially, for each month we construct six soil moisture exposure bins indicating the days spent in each 0.10 cm^3/cm^3 interval beginning at zero and ending at 0.60 cm^3/cm^3. For the temperature exposure variables, we also construct monthly exposure bins such that the first bin aggregates all days with temperatures below 0 degrees Celsius, and the remaining five bins represent the days spent in each 10-degree interval ranging from zero to 50 degrees Celsius for a given month. Time invariant soil characteristics, such as the water holding capacity, are drawn from the Soil Survey Geographic (SSURGO) Database (Soil Survey Staff 2022). To ensure the values for each soil characteristic are not affected by soils that are not used for agricultural purposes, we overlay each county with a grid of one square mile cells and remove any cells that are less than 50% cropland using the Cropland Data Layer (USDA National Agricultural Statistics Service 2022).

2.4 Conceptual model and methodology

2.4.1 Conceptual model

Our conceptual model is presented in two stages. The first is designed to illustrate how the decisions of agents within a given county concerning the adoption of conservation agricultural practices generate the county-level adoption outcomes we employ. In doing so, we can then demonstrate how the rate of adoption in a county serves as a proxy for the constituent producers' net returns to adoption, and consequently the likelihood of incentivized adoption being additional as well. This component, concerning the adoption of practices, is inspired by the threshold model of technology diffusion described in Sunding and Zilberman (2001) and the model of voluntary opt-in presented in van Benthem and Kerr (2013). The second stage of our conceptual model adds to the first by also incorporating heterogeneity in the carbon sequestration potential of the practices being adopted. Using the inferences from our conceptual model of practice adoption, we create a diagram illustrating how the interaction between the pace of adoption and carbon sequestration potential of a practice in a county produces carbon credits of varying quality. Furthermore, we can use these same inferences to predict how cost-effective operating a carbon farming program in a given region will be at meeting carbon sequestration goals.

Consider two counties, A and B, with identically sized populations of 1,000 profit-maximizing producers indexed by i. The producers in each county must decide whether to adopt a conservation practice in each year, $t \in [0,2]$, based on their heterogenous net returns to adoption, $r_{i,t}^{A}$ or $r_{i,t}^{B}$ for producers in county A or B respectively. A producer will decide to adopt the practice each year if the net return to adoption is positive, $r_{i,t}^{A}, r_{i,t}^{B} > 0$. The net return to adopting the practice in year t is the sum of a time invariant component reflecting the individual producer's biophysical conditions and operational characteristics, r_i^A and r_i^B, and a time varying component, L_t. As such, the distribution of net returns for counties A and B, $g\left(r_{i,t}^{A}\right)$ and $g\left(r_{i,t}^{B}\right)$ respectively, will shift over time but retain a constant shape determined by the distribution of their producer's time invariant characteristics: $f\left(r_i^A\right) \sim N(-2/3\,,1)$; $f(r_i^B) \sim N(-2,1)$. In both counties, the time-varying component to net returns evolves identically, $L_{t=0} = 0$; $L_{t=1} = \frac{2}{3}$; $L_{t=2} = \frac{4}{3}$, such that the only difference between the counties is in the mean values of their respective distributions. Panels (A) and (B) of Figure 2.1 depict the cumulative distribution functions for net returns at times $t = 0$ to 2 in counties A and B, $\mathrm{G}\left(r_{i,t}^{A}\right)$ and $\mathrm{G}\left(r_{i,t}^{B}\right)$.

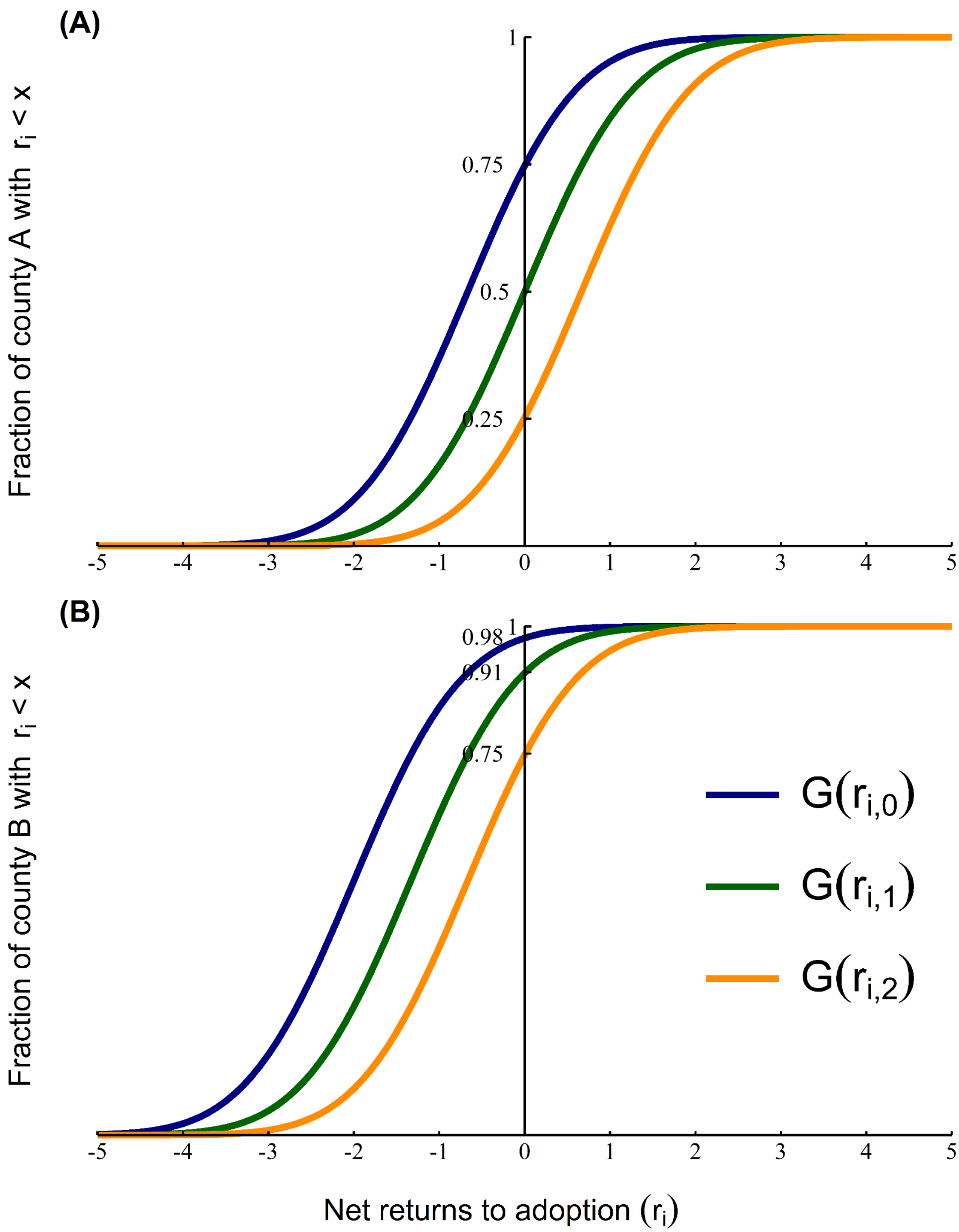

Figure 2.1. Cumulative distribution functions for the net return to adopting conservation agriculture practices in example counties A and B. The y-intercept of the function in each time period indicates the portion of the population not using the practice.

In both panels of Figure 2.1, the increase in net returns to adoption caused by the growth of L_t over time produces the rightward shift of the $G(r_{i,t})$ curves relative to the $G(r_{i,t-1})$ curves. In Panel (A), this shift causes the fraction of the population adopting the practice in county A, $1 - G(r_{i,t}^A = 0)$, to double from a quarter to half of the producers between times $t = 0$ and $t = 1$. In Panel (B), an equivalent increase in L_t between times $t = 0$ and 1 increases adoption from roughly two percent to just over nine percent of producers. Then, after an equivalent increase in L_t between times $t = 1$ and 2, adoption increases by 25 percent in county A and 16 percent in county B. So, the rate of change in adoption between one period and the next serves as an indicator of the size of the population with net returns near zero, and thus the expected magnitude of adoption in the near future as well. Furthermore, if we were to continue increasing L_t and plot the resulting adoption, we would recreate the characteristic "S" shaped adoption curve that typifies the adoption of new agricultural technologies (Griliches 1957; 1958; Feder, Just, and Zilberman 1985). Lastly, while we assume a normal distribution for demonstration purposes, these dynamics will hold so long as the distribution of net returns is unimodal (Sunding and Zilberman 2001).

To demonstrate how this relationship relates to additionality, consider an administrator operating a payment-for-practice carbon credit program which pays producers a fixed incentive, $P_2 = \$1/4$, if they adopt the practice between time $t = 1$ and time $t = 2$. In addition to the cost of compensating producers, the program administrator also incurs a fixed cost, $F_c = \$50$, when establishing the program in a county. As shown by the resulting cumulative distribution functions plotted in Panels (A) and (B) of Figure 2.2, indicated by $G^P(r_{i,2})$, an additional seven percent of county A's producers and nine percent of county B's producers would enroll and adopt the practice relative to the counterfactual scenario indicated by the orange dashed $G(r_{i,2})$

curves. The program administrator, however, cannot determine if producers would have adopted the practice in the absence of the incentive without accurately knowing each producer's individual net return to adoption. As a result, the 25% of producers who would have adopted the practice between times $t = 1$ and 2 in the absence of an incentive will also enroll in the program, meaning the administrator pays the $\$1/4$ incentive to 32% of producers in county A. On average, deploying the program in county A costs the administrator nearly \$0.41 per enrollee. But, with only 7% of this adoption being additional, the policy maker pays roughly \$1.86 per additional adoptee in county A.

In county B, the policy maker pays just under a quarter of producers to adopt the practice, and there are almost twice as many non-additional adopters (16%) as additional adopters (9%). While the payment-per-practice is identical between counties, the average cost per enrollee is greater at \$0.45 in county B due to the smaller number of total program enrollees and identical fixed cost. The average cost of an additional adoptee, however, is cheaper in county B at \$1.25. This is because in the county with a lower rate of adoption between the first and second periods, a greater portion of the program administrator's expenditure on incentivizing adoption in the second period goes toward efficient, additional adoption. In summary, a lower (greater) rate of adoption suggests the portion of non-adopting producers in an area with net returns near zero is small (large) and, consequently, programs incentivizing adoption in the region are more (less) likely to create additional adoption.

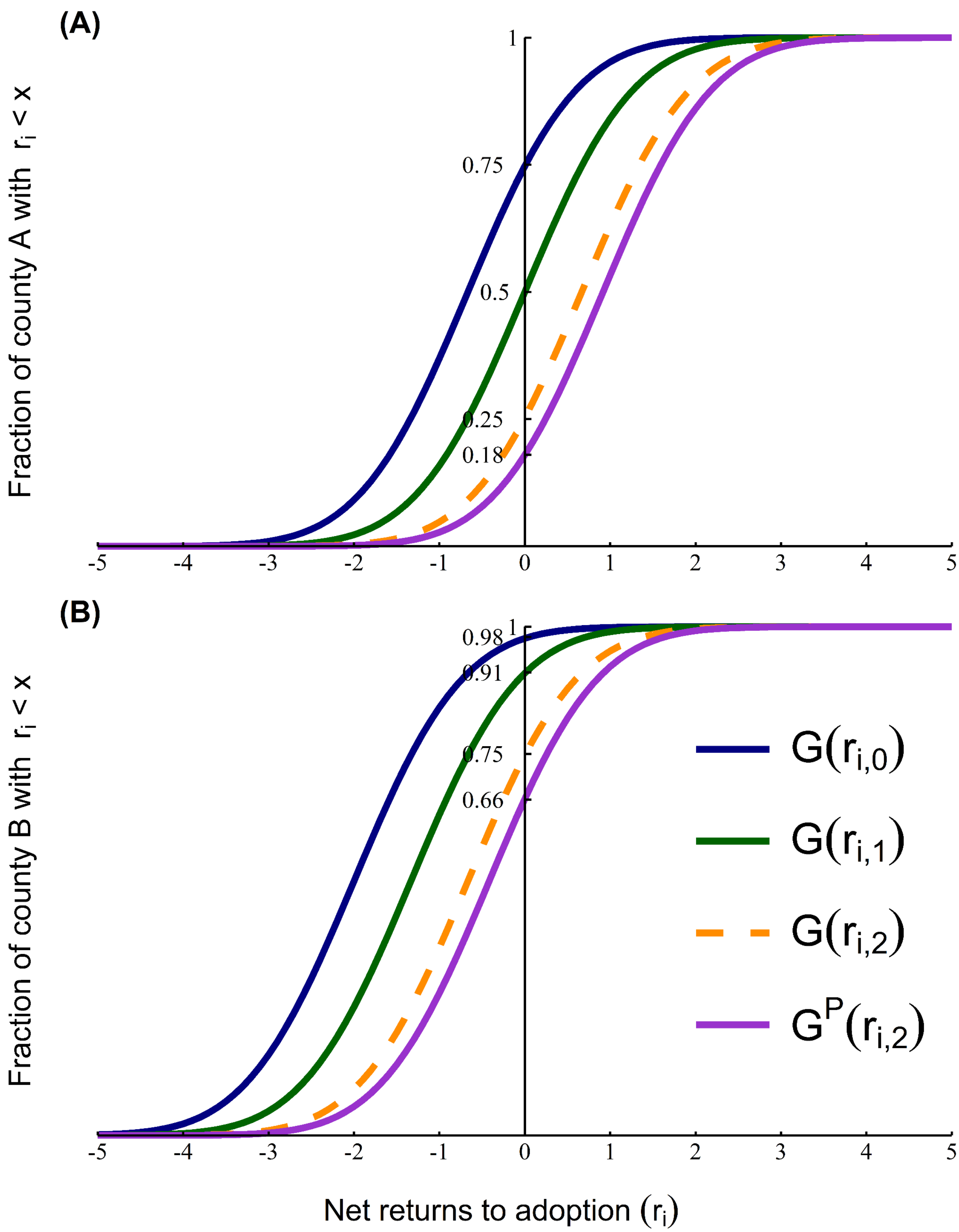

Figure 2.2. Cumulative distribution functions for the net return to adopting conservation agriculture practices in example counties A and B when a payment-for-practice program is available. The dotted graph indicates the counterfactual outcome without the program.

2.4.2 Heterogenous carbon sequestration potential

With this inverse relationship between the rate of adoption and the likelihood of additionality established, we now introduce heterogeneity in the carbon sequestration potential of conservation practice adoption. Consider the same scenario depicted in Figure 2.2 except there are now two versions of each county and the producers in one version of each county sequester more carbon on average if they adopt the conservation practice. We will assume there are 1,000 producers in each of these four counties: A^{low}, A^{high}, B^{low} and B^{high}. Counties A^{low} and A^{high} have the same distributions of net returns as county A in Figure 2.2, and counties B^{low} and B^{high} have distributions of net returns identical to county B. For simplicity, producers in counties with low sequestration potential, A^{low} and B^{low}, sequester 1 ton of carbon if they adopt the practice, and producers in counties with high sequestration potential, A^{high} and B^{high}, sequester 2 tons of carbon. The evolution of L_t, the program administrator's fixed cost of running the program in each county, and the fixed incentive $P_2 = \$1/4$ for adopting the practice between the second and third periods remain unchanged.

Unlike the earlier program where the goal was increasing adoption of the practice, the objective of the policy maker in this case is to produce carbon credits. The policy maker will generate one carbon credit for each ton of carbon sequestered by a program enrollee. Note that, because the policy maker is still unable to distinguish between additional and non-additional enrollees, it will issue carbon sequestration credits for every enrolled producer. For example, the policy maker would issue 320 credits for the one unit of carbon sequestered by each of the 320 producers in county A^{low} who adopt the practice between times $t = 1$ and 2, even though 250 of these producers would have adopted the practice and sequestered 250 tons of carbon even if the program was unavailable. In the same way, the policy maker produces 640, 250, and 500 credits in counties A^{high}, B^{low}, and B^{high} respectively. The average costs of producing a carbon credit (AC_{cred}) in the four counties, as calculated by the policy maker, are: $AC_{cred}(A^{low}) = \$0.41;$ $AC_{cred}(B^{low}) = \$0.45; AC_{cred}\left(A^{high}\right) = \$0.21;$ and $AC_{cred}\left(B^{high}\right) = \$0.23$. For the naïve policy maker, there are two clear implications. First, in choosing between two counties with the same carbon sequestration potential (A^{high} vs. B^{high}; A^{low} vs. B^{low}), targeting the county with a higher rate of practice adoption (A^{high}; A^{low}) will generate a greater number of credits at a lower per-unit cost. Second, deploying the policy in areas with high sequestration potential is an unambiguous improvement in cost effectiveness:

$$AC_{cred}(B^{low}) > AC_{cred}(A^{low}) > AC_{cred}(B^{high}) > AC_{cred}(A^{high}) \qquad \text{(Eq. 2.1)}$$

Alternatively, the efficiency of the program can be measured based on the average cost of sequestering an additional ton of carbon, $AC_{add\ seq}$. Using the \$1.86 average cost of an additional adopter for county A from earlier, the average costs of sequestering an additional ton of carbon in the two counties with high rates of adoption are: $AC_{add\ seq}(A^{low}) = \$1.86$; $AC_{add\ seq}\left(A^{high}\right) = \$0.93$. In contrast, the policy maker pays \$1.25 and \$0.63 per additional unit of carbon sequestered in counties B^{low} and B^{high}. Unlike the average cost of producing a credit, the average cost of an additional ton of carbon sequestered is lower for a county with a lower rate of adoption when comparing two counties with identical rates of sequestration:

$$AC_{add\ seq}(A^{low}) > AC_{add\ seq}(B^{low}) > AC_{add\ seq}(A^{high}) > AC_{cred}(B^{high}) \qquad \text{(Eq. 2.2)}$$

Additionally, consider a hypothetical county, B^{mid}, with the same low rate of adoption as counties B^{low} and B^{high} but having an intermediate sequestration rate of 1.5 tons of carbon. Despite having a lower carbon sequestration rate, the average cost per ton of additional sequestration would be less than that of a county with greater sequestration, A^{high}:

$$AC_{add\ seq}(A^{high}) = \$0.93 > AC_{cred}(B^{mid}) = \$0.83 \qquad \text{(Eq. 2.3)}$$

So, if using the average cost of additional sequestration as the efficiency metric, targeting areas with the highest rates of carbon sequestration is not always optimal.

The diagram in Figure 2.3 displays the adoption rate and carbon sequestration characteristics for counties A^{low}, A^{high}, B^{low} and B^{high}, along with the average cost per credit calculated by the policy maker and the average cost per additional ton of carbon sequestration. To facilitate interpretation of the maps presented in the results section, we label each quadrant of Figure 2.3 based on these characteristics. The upper left quadrant of Figure 2.3, for example, describes county A^{low} with its high rate of adoption and low carbon sequestration potential. In comparison to the lower left quadrant, the lower average cost of a credit indicates producing carbon credits in this county will be relatively cheaper for the policy maker.

If the policy maker sells the credits produced in each county at their respective average costs, then every dollar of carbon credits produced in county A^{low} purchased by an entity represents 0.54 tons of carbon sequestered. In comparison to the credits produced in B^{low}, where every dollar of credits represents 0.8 tons of carbon sequestered, the quality of the credits produced in county A^{low} is relatively poor. By making similar comparisons between the remaining counties, we can see that the program administrator's cost of producing carbon credits is decreasing in the northeasterly direction of Figure 2.3, but the quality of the credits produced is increasing in the southeasterly direction. These relationships are the basis for our categorization of counties in the maps we present in the results section.

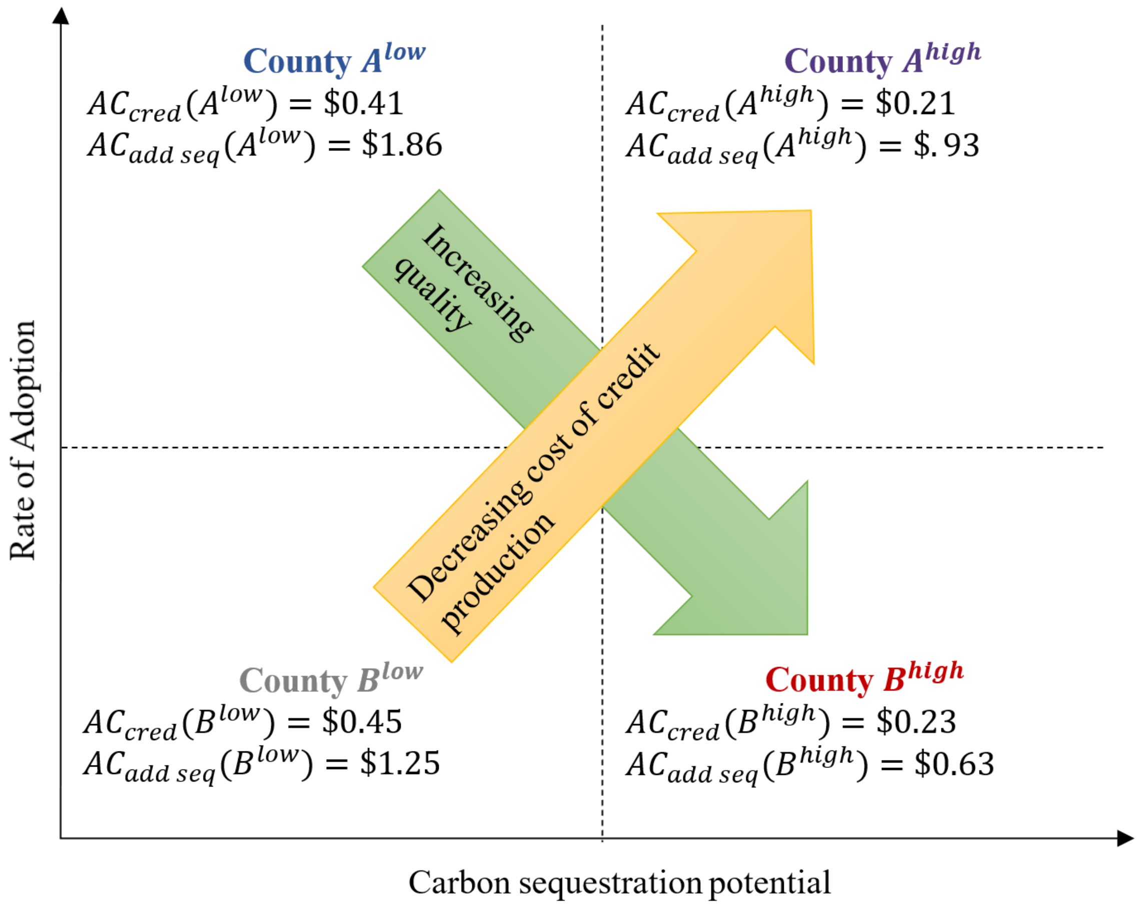

Figure 2.3. Carbon credit quality and cost as a function of the rate of practice adoption and carbon sequestration in the area where the program occurs.

In Figure 2.3, the clear implication of our conceptual model is that targeting areas lying in the eastern quadrants, regions with greater carbon sequestration potential, will generally improve the cost-efficiency of a carbon offset program. This conceptual model and the associated results are meant to demonstrate this general principle for voluntary carbon sequestration programs, rather than making predictions or prescriptions for specific circumstances. For example, if the policy maker is deciding between two counties with identical carbon sequestration potential but differing rates of adoption, we cannot predict that targeting the region with a lower rate of adoption will reach a sequestration target in a more cost-efficient manner. We can make comparisons between the four example counties in Figure 2.3 solely because of the distributional assumptions we make concerning producers' returns and the chosen fixed cost of operating the program. As such, the choice between targeting regions with low or high rates of adoption will depend on the specifics of a potential program's design and the policy maker's objectives. For instance, a policy maker prioritizing additionality would likely target regions with lower rates of adoption, while one seeking to minimize costs would likely target regions with greater rates of adoption.

2.5 Methods

2.5.1 Predicting the rate of practice adoption between 2017 and 2022

As we have the expected carbon sequestration of practices from the COMET Planner tool, the remaining data we require to populate our conceptual framework is the rate of practice adoption. More specifically, our objective is to parameterize the future rate of adoption because it will determine the additionality of current programs encouraging conservation agriculture practices. If we assume the adoption of conservation agriculture practices follows the typical "S" shaped adoption curve, then the rate of adoption will have a non-linear relationship with the current level of adoption. Due to this relationship, we anticipate using the raw data on practice use for 2012 and 2017 would not accurately reflect the future evolution of conservation agriculture practice use. Instead, we use a random forest to predict the rate of adoption for cover-cropping and tillage reduction practices between 2017 and 2022 due to the exceptional predictive accuracy random forests provide (Storm, Baylis, and Heckelei 2020).

As we use it for regression analysis, the algorithm for the random forest predictor involves constructing a specified number of regression trees for a series of bootstrapped samples from the true data (Breiman 2001). The structure of each individual regression tree is determined by recursively partitioning the bootstrapped sample data such that the partitions optimize the Classification and Regression Tree (CART) criterion from Breiman et al. (1984) for a random sample of the possible predictors. The number of predictors tested at each node, or partition, is also specified by the user. The recursive splitting of variables makes random forests especially effective when there are non-linear relationships between variables, such as the one previously described between past and future rates of adoption. Another advantage of the random forest approach is that the algorithm involved includes testing the out-of-sample prediction capacity of the model, because each tree in the forest is trained using a modified training dataset constructed by bootstrapping (Efron and Hastie 2016).

For this work, we combine the soil, weather, and growing region variables with the observed adoption behavior in 2012 and 2017 from the USDA Census of Agriculture to construct our training dataset. The objective of the training exercise is to minimize the squared prediction error when modeling the following relationship:

$$ln\left(\frac{y_{i,t}}{y_{i,t-\tau}}\right) = f\left(ln\left(y_{i,t-\tau}\right), X_{i,t}\right);\ \tau = 5 \qquad \text{(Eq. 2.4)}$$

In Equation 2.4, $y_{i,t}$ is acres using a conservation practice in county i for year t, and $X_{i,t}$ contains the growing condition variables. Within $X_{i,t}$ are daily soil moisture and temperature exposure variables averaged over the five years preceding t, the log transformed maximum county cropland acres, the USDA-ERS farm resource region indicator variables, and the time-invariant soil characteristics. The interval between time points, $\tau = 5$, in Equation 2.4 reflects the gap between USDA Census datapoints. By taking the arithmetic mean of the soil moisture and temperature variables across these five years, we are assuming each of the five years preceding an observation carries equal weight in the decisions of a given county's producers.

To generate the random forest, we use the *randomForest* R package by Liaw and Wiener (2002) which implements the algorithm defined by Breiman (2001). The random forest is grown with 500 trees and at each node we sample 50 of the 153 variables in the dataset. Figure 2.4 displays the relative importance for the top 15 most important variables in the random forest predicting the county-level rate of change in acres using a reduced tillage practice between 2012 and 2017.

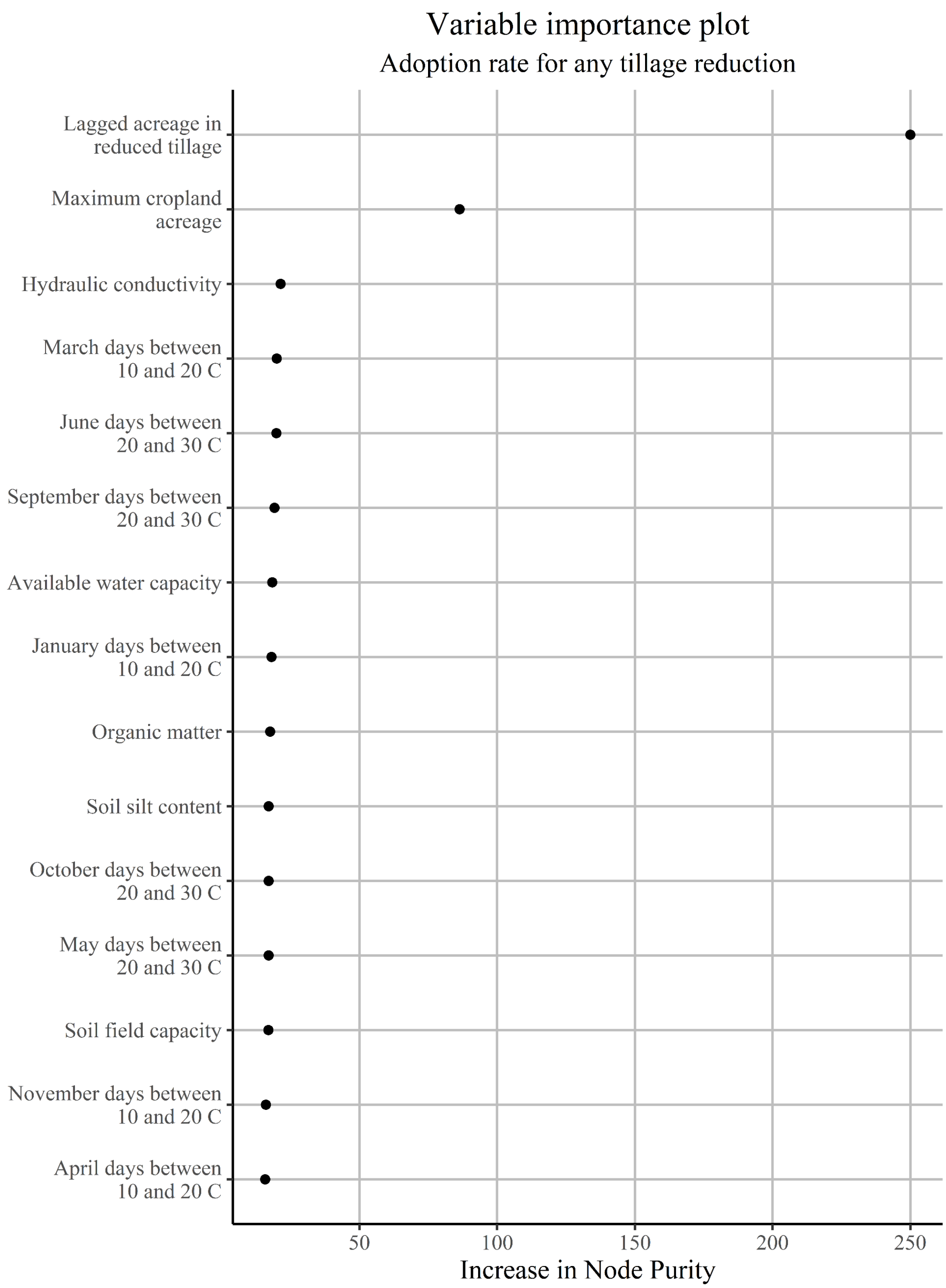

Figure 2.4. Variable importance plot for the random forest predicting the county-level rate of change in acreage using a reduced tillage practice between 2012 and 2017.

In Figure 2.4, the fact that lagged acreage using reduced tillage practices is the most important variable in the random forest supports our decision to predict the rate of change between 2017 and 2022. It suggests that if we were to rely solely on the data between 2012 and 2017, we would not account for the effect that the level of adoption in 2017 has on future rates of change. After generating the predicted rates of adoption for tillage reduction and cover-cropping practices using the random forest, the predicted rates are divided into terciles just as the carbon sequestration estimates were. Figures A.7 and A.8 in Appendix A display the two maps detailing the predicted rates of adoption and sequestration for tillage reduction practices as an example of how the two-way choropleths presented in the results section are created.

2.6 Results and Discussion

2.6.1 No-till, conservation tillage, or reduced tillage practices

In Figure 2.5, we display the predicted rate of change in use of reduced tillage practices between 2017 and 2022 along with the rate of greenhouse gas sequestration at the county-level. First consider the counties in light purple, those in the lowest terciles for both the rate of adoption and sequestration. Using the relationships from our conceptual model, we classify carbon credit production in these areas as a very expensive endeavor resulting in low-quality credits. The low rate of adoption suggests it will take a large incentive for producers to change their practices, and the marginal increase in carbon sequestration for each producer who does make the change will be quite small. Aside from the southern portion of the Appalachian Mountain range, most of the counties with these characteristics are in the western states. While most counties in western states have low sequestration rates for tillage reductions, as shown in Figure A.7, those with low or negative rates of adoption are concentrated in arid or semi-arid regions. Many of the same semi-arid and arid states in the west also contain counties with low sequestration rates but high rates of predicted adoption, indicated by the bright blue color in the top left of Figure 2.5’s legend. In addition, there are also pockets of such counties in the Great

Lakes region, coastal Texas, and Florida. While encouraging adoption in these areas would be relatively cheap due to the high predicted rate of adoption, the credits produced would be very low in quality due to the combination of a low expected sequestration rate and high likelihood of adoption being non-additional.

Counties in bright red in Figure 2.5, with a predicted rate of adoption in the lowest tercile but expected sequestration in the highest tercile, are concentrated in two areas: the lower Mississippi River Basin and the Chesapeake Bay watershed. In these areas, the sequestration associated with reducing tillage is very high in part because COMET Planner incorporates the reduction in nutrient pollution and erosion these practices entail. Given the low rates of adoption, producers will likely require a large incentive to reduce their tillage intensity, but the producers who are incentivized to adopt the practice are likely to be additional. So, while encouraging adoption may be somewhat expensive, the credits produced would be very high in quality. The final category of counties we focus on, those in dark purple in Figure 2.5, have a high predicted rate of adoption and high expected carbon sequestration for tillage reduction practices. These counties are predominantly located in the upper Mississippi river basin, although there are several in northern New York, Maine, and the upper Ohio River basin. Due to their high adoption rate, incentivizing producers to adopt tillage reduction practices would be inexpensive. While this high rate of adoption means a reduced likelihood of additionality, the high expected carbon sequestration associated with adoption in these regions serves to counterbalance the inefficiency introduced by non-additional adopters. Taken together, operating an agricultural carbon credit program in counties with high adoption rates and high expected sequestration for the tillage reduction practice is likely to be a relatively inexpensive endeavor resulting in credits of moderate quality.

Predicted adoption rate versus sequestration potential

Tillage reduction practices

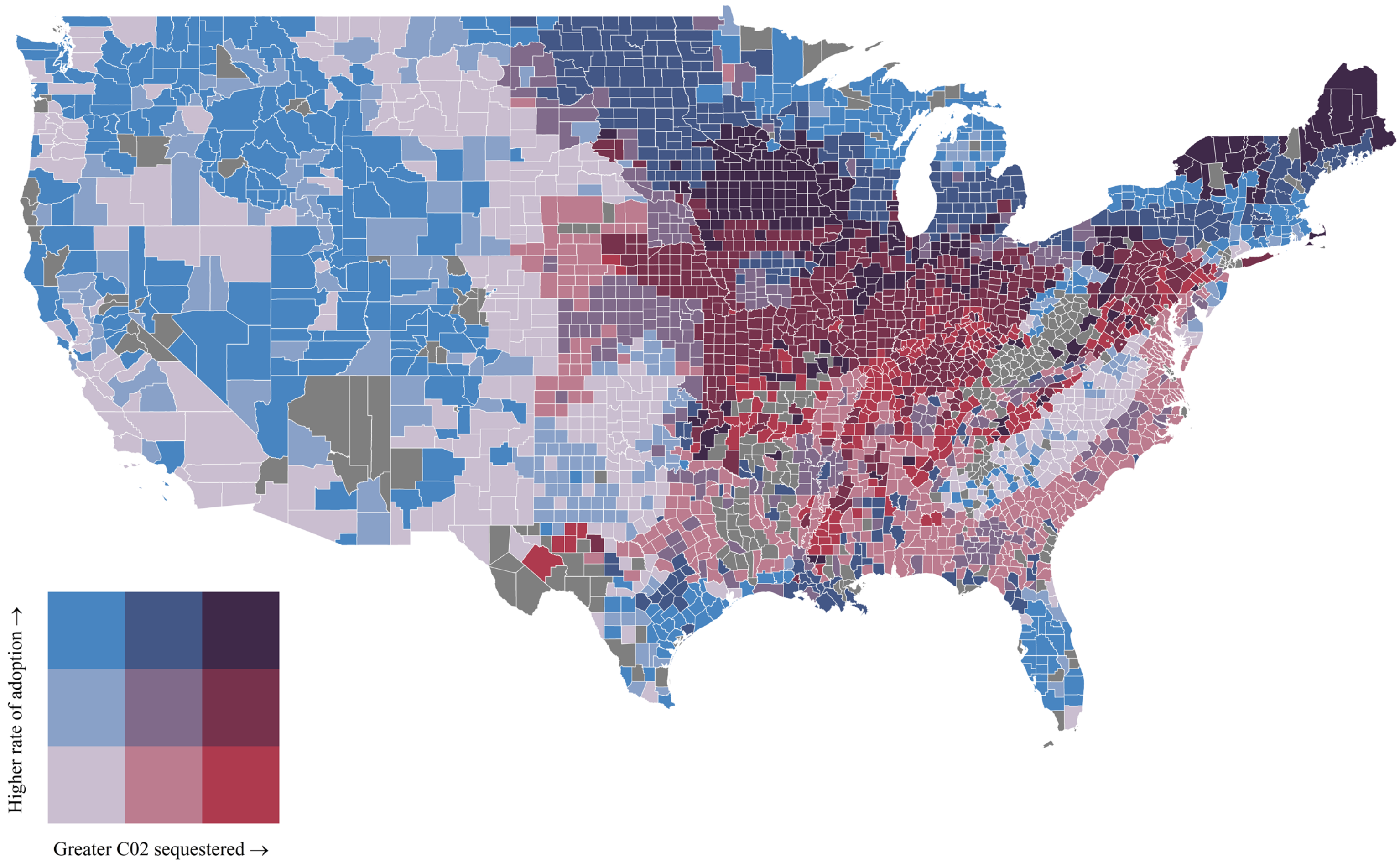

Figure 2.5. Two-way choropleth depicting predicted rates of adoption for tillage reduction practices against the expected sequestration due their use by terciles.

2.6.2 Cover-cropping

One of the main similarities between the results for tillage reduction practices in Figure 2.5 and the results for cover-cropping depicted in Figure 2.6 is the prevalence of counties with low predicted rates of adoption and low expected carbon sequestration in the western United States. Given the additional water demand cover-crops represent, the greater number of counties with low predicted rates of adoption for cover-cropping is not surprising in these arid and semi-arid regions. If rainfall and access to irrigation are insufficient to support cover-cropping, producers will require exorbitant incentives to adopt cover-cropping as it will actively detract from the growth of the crops they plan to harvest. This is likely the reason why some of the counties with high predicted rates of adoption for reduced tillage practices in the Rocky Mountains have low predicted rates of adoption for cover-crops. In general, as is the case for tillage reduction practices, carbon credit programs targeting cover crop adoption in many western states would be expensive efforts producing little in the way of sequestered carbon or avoided emissions.

One of the most interesting differences between Figures 2.5 and 2.6 is the greater prevalence of counties with low carbon sequestration but high rates of predicted adoption for cover-crops in the Upper Mississippi River Basin. Parts of North Dakota, South Dakota, Minnesota, and Iowa have expected sequestration values in the lowest tercile for cover-cropping but in the middle or highest terciles for tillage reduction practices. While encouraging adoption of both types of practices may be equally inexpensive in such areas, the expected contribution of cover-cropping to greenhouse gas sequestration or emission reductions goals would be lower for a program targeting cover-crop adoption. For programs which compensate producers based on the expected sequestration for a suite of practices, the inclusion of cover-crops may dilute the contributions of enrolled producers who adopt both practices. This is especially concerning if incorporating cover-crops into a reduced tillage system drastically increases the uncertainty of the resulting sequestration.

As is the case for tillage reduction practices in Figure 2.5, many of the counties with high expected rates of carbon sequestration, or avoided emissions, and low rates of predicted adoption are located within the southern portion of the Mississippi River basin in Figure 2.6. In contrast to Figure 2.5, there are a greater number of counties with high predicted rates of adoption and high rates of carbon sequestration further south in the Mississippi River basin in Figure 2.6. Given the high expected sequestration for both types of practices in this region, the discrepancy in predicted rates of adoption between tillage reduction and cover-cropping practices represents a significant opportunity for program administrators. If a program were to couple the incentives for adopting cover-crops and reduced tillage practices, it is possible the net return of jointly adopting both practices could become positive despite the individual return to tillage reduction practices being negative. Even though the incentive required might be larger than what is necessary to incentivize cover-cropping in isolation, the cumulative return in carbon sequestration could offset this additional cost given the high sequestration values for both practices in the region.

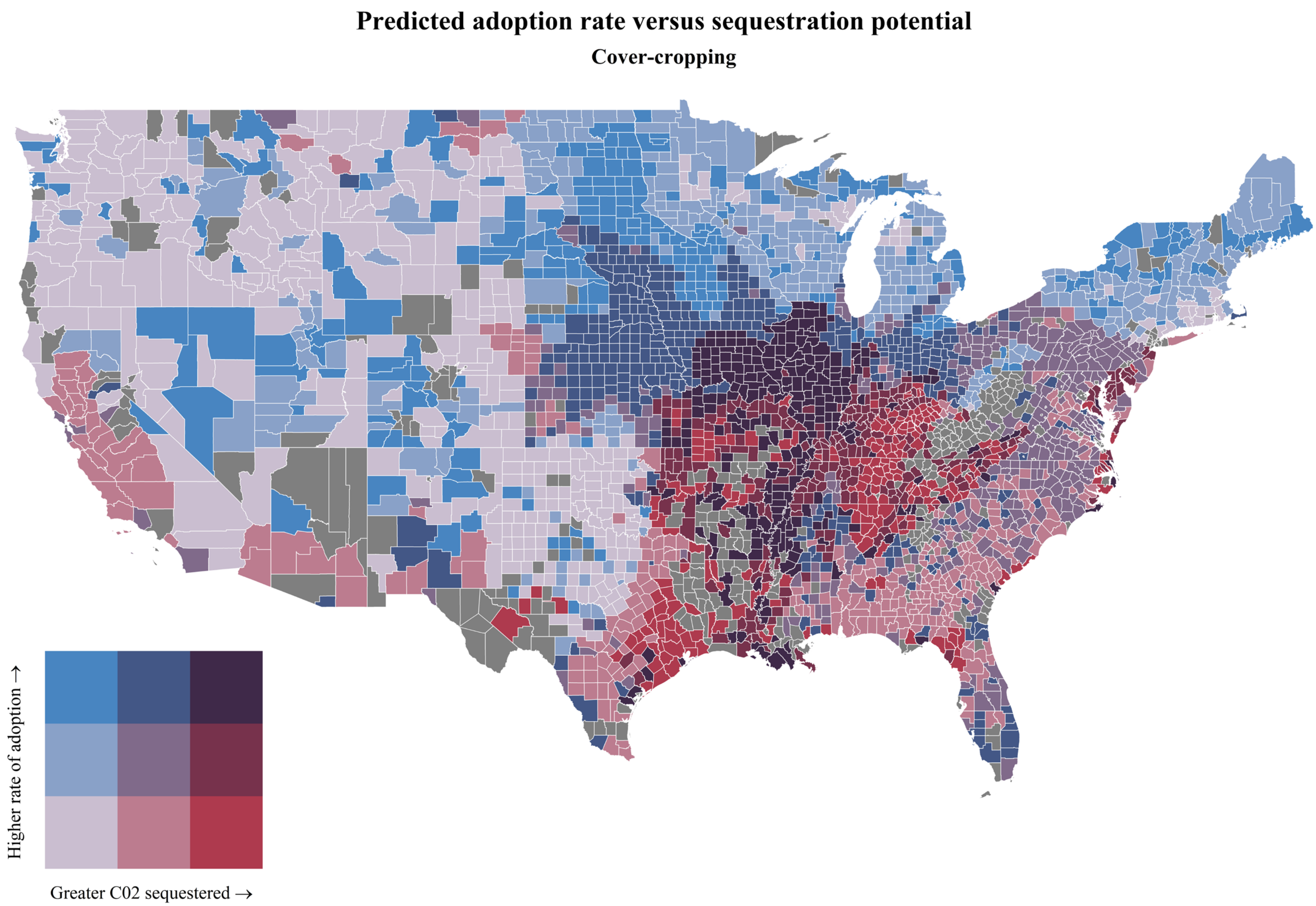

Figure 2.6. Two-way choropleth depicting predicted rates of adoption for cover-cropping against the expected sequestration due to its use by terciles.

2.7 Conclusion

One of the unifying features amongst current voluntary agricultural carbon credit programs is their use of past adoption behavior to determine whether producers are eligible for participation. This focus on the prior use of practices frames the additionality of program outcomes in terms of the number of enrolled producers adopting a conservation agriculture practice, rather than the number of enrolled producers who adopted a practice and would not have done so if the program did not exist. We demonstrate how these non-additional producers, while inexpensive for the program administrator to enroll, detract from the quality of a credit and the efficiency with which purchasers of carbon credits can offset their emissions or reach sequestration objectives. We then illustrate how the interaction between additionality and heterogeneity in the carbon sequestration expected of a given practice can cause carbon credit production in some regions to be attractive to program administrators despite resulting in poor quality carbon credits.

To depict how this interaction may play out in the United States context, we use county-level data on the number of cropland acres where cover-cropping and reduced tillage practices were used to predict the rate of change in use of these practices between 2017 and 2022. After combining these predictions with county-specific estimates of the carbon sequestration expected of cover-cropping and tillage reduction practices, we map the interaction of these two qualities for the continental United States. In many arid and semi-arid areas of the western United States, the predicted rate of adoption and expected sequestration for cover-cropping and tillage reduction practices is very low. Given the high cost of encouraging adoption and little return in terms of carbon sequestration, we do not expect significant enrollment in agricultural carbon credit programs in these areas. For some states in the western United States, coastal Texas, and

Florida, where tillage reduction practices have a high predicted rate of adoption and low expected carbon sequestration, we anticipate producers will be most interested in programs that reward practice adoption rather than carbon sequestration outcomes. While operating a program that solely focuses on practice adoption in these areas may be relatively inexpensive, the high likelihood of adoption being non-additional and the low carbon sequestration expected per adopter would drastically reduce the quality of any credits produced there.

For both practices, much of the corn belt has a high predicted rate of adoption and high expected carbon sequestration, indicating carbon credits of intermediate or good quality could be produced there relatively cheaply. Counties further south in the Mississippi River basin have similarly high rates of expected sequestration when adopting tillage reduction practices, but the lower rates of adoption predicted there suggest incentivizing adoption is likely to be more expensive. However, this low rate of adoption indicates any adoption that occurs is more likely to be additional, so the increase in costs could be offset by the improvement in quality increased additionality affords.

Ultimately, the quality of agricultural carbon credits produced in any region of the United States will depend on the issuer's ability to credibly substantiate the additional change in soil carbon their credits represent. While this could entail program administrators targeting regions with low rates of adoption and high expected carbon sequestration, common means of mitigating risk could prevent regions with low carbon sequestration potential or high rates of adoption from being excluded. For example, credits from different regions could be pooled to reduce the risk of non-additionality for consumers while decreasing the cost of incentivizing adoption born by program administrators. Alternatively, the trading ratios proposed by several authors to combat non-additionality could be adjusted based on carbon credits' region of origin (Chung 2007;

Environmental Defense Fund 2007; Schneider 2009). Finally, the maps displayed in this work rely on county-level data which obscures local heterogeneity. While the average expected sequestration or rate of adoption for a practice in a county may be high, local site conditions could cause individual producers to have a significant advantage in carbon credit production relative to their neighbors.

Chapter 3 - Compensation mechanism, program scale, and the efficiency of voluntary carbon offset programs

3.1 Introduction

Adoption of conservation agriculture activities on U.S. agricultural land could sequester 234 Tg CO_2 Eq. per year, roughly 4% of the United States' annual emissions in 2020 (Sperow 2016; EPA 2020). Reductions in tillage intensity, which contribute the majority of this sequestration potential, are one of the more economical sources of sequestration costing roughly $14 to $55 / tonne CO_2 Eq. for the central United States (Antle et al. 2007). While the estimated cost is slightly lower for forest-based sequestration at between $8 and $25 / tonne CO_2 Eq. (Richards and Stavins 2005), public and private interest has increasingly focused on the carbon sequestration potential of agricultural land in recent year. For example, all four of the voluntary agriculture conservation programs receiving funds under the 2022 Inflation Reduction Act specify that projects reducing emissions or increasing sequestration of greenhouse gases must be prioritized when considering program applications (117th Congress 2022). In the private sector, there are now at least 11 voluntary agricultural carbon credit programs purchasing offsets generated by agricultural producers using conservation agriculture practices in the United States (Plastina and Wongpiyabovorn 2021).

Regardless of whether a program is public or private, voluntary programs suffer from an inherent information asymmetry between the policy maker and potential participants. Ideally, the policy maker would only enroll participants whose behavior is additional, meaning they would not have altered their emissions or sequestration in the baseline, counterfactual scenario where the policy is unavailable. In an analysis of U.S. voluntary agriculture conservation programs, Claassen, Duquette, and Smith (2018) find that just 47% of payments for conservation tillage practices produced additional adoption. Similarly, Sawadgo and Plastina (2021) conclude that 54% of the expenditures on cost-share agreements in their study of cover-crop adoption in Iowa were additional.

Programs producing offsets, like the 11 agricultural carbon credit programs analyzed in Plastina and Wongpiyabovorn (2021), combat this problem of non-additionality by comparing producers' behavior to an assigned baseline, or benchmark, based on their production history for example. However, even with extensive information on their past behavior, the baseline assigned by the policy maker is still an estimate of producers' counterfactual. The variety of approaches for setting baselines among current voluntary agricultural carbon credit programs, and the magnitude of error each entails, are the result of each entity weighing the cost of uncovering additional information on potential participants against the benefits of reducing spurious expenditures (Montero 2000).

In addition to determining how baselines will be set, administrators of agricultural carbon offset programs must also decide how they will compensate producers for the carbon sequestration or avoided emissions resulting from their production practices. One option is to compensate each producer on a per-acre basis for its use of practices above an assigned baseline based on the policy maker's prediction of counterfactual practice use in acres. Alternatively, producers could be compensated for results they produce on a per-unit of sequestration basis. In theory, a payment-for-sequestration policy is more efficient than the alternative payment-for-practice option because producers' per-unit compensation can be equated with the marginal social benefit of carbon sequestration (Sandmo 1975). So, even for difficult-to-measure ecosystem services, like biodiversity, results-based payment for ecosystem services programs are generally considered to be more cost effective than action-based programs (Hanley et al. 2012). But, while Wuepper and Huber (2021) find this to be true in an empirical study comparing action-based and results-based payments for biodiversity outcomes in Switzerland, they mention that their results are limited by the fact that they do not observe the outcome variable, changes in biodiversity, directly. This limitation highlights the principal difficulty of results-based programs, accurately measuring outcomes.

This issue of measuring outcomes is especially pertinent for agricultural carbon offset programs due to the large uncertainties in carbon sequestration outcomes. Regional uncertainties in such estimates can range between $\pm 114\%$ to 118%, but site scale uncertainties can be between $\pm 674\%$ and $\pm 739\%$ (Ogle et al., 2010). One solution to this issue, originally proposed to mitigate error in estimates of producers' baselines for offset programs, is to increase the scale of the program and take advantage of the inverse relationship between the scale of analysis and uncertainty. Using a model in which agents are offered a fixed conservation incentive akin to a payment-for-practice program, van Benthem and Kerr (2013) demonstrate that increasing the scale at which baselines are set and objectives measured improves the efficiency, offset quality, and average cost for a carbon offset program. In their model, increasing the program's scale involves aggregating the predicted counterfactual behavior across groups of agents, and then compensating the group for its performance relative to this aggregate baseline.

In this paper, we build on the approach of van Benthem and Kerr (2013) and investigate how the choice of compensation mechanism, payment-for-practice or payment-for-sequestration, affects the benefits of increasing the program's baseline scale when producers' have heterogenous carbon sequestration outcomes. We begin by determining whether an individual-level program (i) paying for practice(s) adopted on a per-acre basis or (ii) paying for the carbon offset, or sequestered, on a per-unit basis generates greater efficiency gains when the policy maker estimates producers' baseline behavior with error. Then, we introduce error in the policy maker's estimates of producers' carbon sequestration and compare the programs' performance. Due to the significance both types of error have for determining the relative performance of the compensation mechanisms, we test the impact of program scale for various combinations of carbon sequestration and baseline estimation error.

The two most similar works concerning agricultural carbon sequestration are the empirical work of Antle et al. (2003) and theoretical investigation by van Benthem and Kerr (2013). Antle et al. (2013) include heterogeneous outcomes and producers' costs of adoption in their model using site-specific estimates of carbon sequestration and producer characteristics. They find a per-acre payment-for-practice program could be as much as five times more costly than a per-ton payment-for-sequestration program depending on the acceptable level of uncertainty in carbon sequestration estimates. Unlike in our model, the two programs in Antle et al. (2013) do not measure performance relative to a counterfactual scenario. So, there is no role for error in the policy maker's estimates of producers' behavior in the absence of the program. In addition, the carbon sequestration estimates in Antle et al. (2013) are formulated at a regional scale based on the USDA's Major Land Resource Areas and the effect of program scale on these estimates is not investigated. As mentioned earlier, the work of van Benthem and Kerr (2013) focuses primarily on the effect of program scale on the efficiency of a carbon offset program that compares producers' behavior to a predicted counterfactual scenario. As the carbon sequestration potential of all agents in van Benthem and Kerr's (2013) model is assumed to be identical, they do not consider how the choice of compensation mechanism or uncertainty in carbon sequestration outcomes modifies the effect of increasing program scale.

Our work addresses the limitations of these two previous works by uniting the heterogenous sequestration outcomes and competing programs designs of Antle et al. (2003) with the baseline prediction and program scale features of van Benthem and Kerr (2013). In doing so, we make three contributions to our understanding of carbon offset policies. First, if the policy maker knows the sequestration potential of producers with certainty, we find the payment-for-sequestration program is more efficient regardless of the policy maker's error in estimating baselines because the per-unit compensation mechanism selects for individuals with greater carbon sequestration potential. Second, when we introduce uncertainty in the policy maker's estimates of sequestration outcomes, the payment-for-practice program is more efficient under certain conditions. For example, when the policy maker's error in estimating sequestration is high relative to its error estimating producer's baseline adoption behavior, the average cost of additional sequestration is lower for the payment-for-practice program. Finally, increasing the program scale reduces the average cost of additional sequestration for both compensation mechanisms, but the current level of adoption in the population determines which program benefits to a greater degree. In areas with low levels of adoption, scaling up the program benefits the payment-for-practice program to a greater extent, and the opposite holds true for the payment-for-sequestration program.

3.2 Model of an individual-scale voluntary carbon credit program

The initial formulation of our model modifies the individual-level model from van Benthem and Kerr (2013) for a hypothetical agricultural context where adoption of a practice produces a positive marginal carbon externality by increasing carbon sequestration. To demonstrate how the model functions, envision an example case where the agents in our model are producers considering switching from conventional tillage to no-till for the next growing season. Next growing season, even in the absence of any program, some of these hypothetical producers may be planning to adopt no-till next growing season purely because of its agronomic benefits. We refer to this counterfactual scenario as the baseline scenario in our model, and the number of producers using the practice in the baseline scenario is determined by each individual producer's net return to adoption r_i. We assume producers are profit maximizing, so the baseline behavior for producers with net returns greater than zero is to use the practice, indicated by the binary variable BL_i equaling one:

$$BL_i = \begin{cases} 1 \;\; if \; r_i > 0 \\ 0 \;\; if \; r_i \leq 0 \end{cases}. \qquad \text{(Eq. 3.1)}$$

We construct this baseline scenario for a group of N producers with identically sized parcels, such that parcel and producer are synonymous and indexed by i. Defining the distribution of net returns across i as f_r, then the portion of producers adopting the practice in this baseline scenario can be expressed as: $\frac{1}{N}\sum_i^N BL_i \sim \int_0^\infty f_r(r)dr$.

Next, we introduce a payment-for-practice program, in which the policy maker offers a payment equal to p_p to producers who it predicts would not adopt the practice in the absence of the program. In the no-till example, this program would pay producers p_p for adopting no-till in the coming growing season only if the policy maker predicts they were going to use conventional tillage in the absence of a program. The policy maker in our model predicts producers' baseline use of the conservation agriculture practice based on its estimates of their net returns, defined as $\hat{r}_i$ for producer i. We use the symbol $\widehat{\;}$ to indicate when variables are estimates made by the policy maker instead of their true values.

In the case of the policy maker's estimates of net returns ($\hat{r}_i$), the policy maker's estimates include some error ($\varepsilon_{i,r}$) such that: $\hat{r}_i = r_i + \varepsilon_{i,r}$. This could be because the policy maker, for instance, uses the state level change in no-till adoption to predict the likelihood of individual producers in that state using no-till next year. While this prediction may be correct on average, the policy maker will likely make mistakes for individual producers in the state because of their unique production characteristics. In a similar fashion, we assume the error in the policy maker's estimate of net returns, $\varepsilon_{i,r}$, is symmetric around zero, so the policy maker's estimates are unbiased. We also assume the error in predicting a producer's net returns, $\varepsilon_{i,r}$, is independent of the producer's true return to adoption drawn from f_r, and that it has standard deviation σ_r. The policy maker's predicted baseline behavior for producer i is therefore:

$$\widehat{BL}_i = \begin{cases} 1 & if\ \hat{r}_i > 0 \\ 0 & if\ \hat{r}_i \leq 0 \end{cases}. \qquad \text{(Eq. 3.2)}$$

We can now illustrate two inefficiency categories created by the policy maker's net return observation error, inefficient participants and non-participants, by examining a representative producer's adoption decision when this payment-for-practice program is available. Given its assigned baseline, a producer will enroll in the program if the net return to participating is positive, and its assigned baseline indicates they would not have used the practice in the absence of the program: $r_i + p_p > 0$ and $\widehat{BL}_i = 0$. However, the producer will only be efficiently enrolled in the program if its true baseline behavior was to not adopt the practice, meaning $r_i + p_p > 0$, $\widehat{BL}_i = 0$, and $BL_i = 0$.

Inefficient participants are the non-additional producers who are enrolled even though they would have adopted the practice if the program was unavailable, $r_i + p_p > 0$, $\widehat{BL}_i = 0$, and $BL_i = 1$. For the example concerning the adoption of no-till, these would be producers who were planning to adopt no-till next growing season, but the policy maker mistakenly predicts that they were going to use conventional tillage. As a result, the policy maker lets them participate and they are paid for adopting a practice they had already planned to use. In our model, these non-additional, inefficient participants can enroll because the policy maker underestimates their net return to using the practice.

Conversely, the inefficient nonparticipants occur because the policy maker overestimates some producers' net return to adoption, so the resulting assigned baselines prevent producers from efficiently participating: $r_i + p_p > 0$, $\widehat{BL}_i = 1$, and $BL_i = 0$. These would be producers who were going to use conventional tillage next season, but the policy maker predicts that they're going to use no-till in the example case. As a result, these producers are not offered the payment p_p, even though it would have caused them to switch from conventional tillage to no-till next season.

3.3 Heterogeneity in the marginal carbon externality from adoption

We now allow the societal benefit of an individual opting-in, the carbon sequestered due to adopting a practice in this context, to vary across agents. For the no-till adoption example, this change mimics the variation in carbon sequestration that occurs when producers with two different soil types and production histories adopt no-till. In our model, we define the carbon sequestered by producer i if it adopts the practice as c_i, and we initially assume that the policy maker knows c_i with certainty. If this were the case for the no-till adoption example, the policy maker would know exactly how much carbon each individual could sequester if they switched to no-till.

Due to the heterogeneity in carbon sequestration among producers, the outcomes of the program and efficiency metrics are defined in terms of the carbon sequestered from this point onward. For example, the policy maker no longer uses the number of producers adopting the practice above their assigned baseline behavior to calculate how much carbon the program sequesters. Instead, the policy maker computes the outcome of its program based on the carbon which is sequestered in excess of what producers were expected to sequester in the baseline scenario. To demonstrate this distinction, we construct the true baseline carbon sequestration for producer i (BLC_i), along with the carbon sequestration expected by the policy maker ($\widehat{BLC}_i$) given its prediction of producer i's net returns:

$$BLC_i = \begin{cases} c_i \ \ if\ r_i > 0 \\ 0 \ \ if\ r_i \leq 0 \end{cases} \qquad \text{(Eq. 3.3)}$$

and

$$\widehat{BLC}_i = \begin{cases} c_i \ if\ \hat{r}_i > 0 \\ 0 \ \ if\ \hat{r}_i \leq 0 \end{cases}. \qquad \text{(Eq. 3.4)}$$

3.3.1 Payment-for-practice program

As described earlier, the payment-for-practice program offers the fixed level of compensation p_p to producers who adopt the practice when their assigned baseline indicates they would not have otherwise. Let the binary variable E_i^p indicate the enrollment status of producer i, having a value of one if producer i enrolls, according to:

$$E_i^p = \begin{cases} 1\ if\ r_i + p_p > 0\ and\ \widehat{BL}_i = 0 \\ 0\ if\ r_i + p_p \leq 0\ or\ \widehat{BL}_i = 1 \end{cases}. \qquad \text{(Eq. 3.5)}$$

In Equation 3.5, the superscript p is used to indicate variables associated with the payment-for-practice program. Notice that E_i^p can only take a value of 1, indicating the producer is participating, if the policy maker predicts that the producer would not use the practice in the baseline scenario ($\widehat{BL}_i = 0$). A producer may not be participating, $E_i^p = 0$ in Equation 3.5, for one of two reasons. First, it could be because the incentive p_p is too small to make adoption profitable, meaning $p_p < |r_i|$. Or a producer may not be participating because the policy maker estimates that it would adopt the practice in the baseline scenario $\widehat{BL}_i = 1$, so it is not eligible for the payment. For simplicity, we assume perfect compliance by producers meaning all producers who accept the payment and participate will use the practice without fail.

Given the variation in producers carbon sequestration outcomes, we also define the sequestration that occurs based on producers' behavior when the payment-for-practice program is available as:

$$c_i^p = \begin{cases} c_i \ \ if\ r_i + E_i^p * p_p > 0 \\ \ \ 0 \ \ if\ r_i + p_p \leq 0 \end{cases} . \qquad \text{(Eq. 3.6)}$$

In Equation 3.6, there are three cases in which a producer sequesters carbon when the payment-for-practice program is available, such that $c_i^p = c_i$. First, there are the efficient enrollees, as defined above, who would not have adopted the practice in the baseline so $BLC_i = 0$. Second, there are inefficient participants who would have adopted the practice in the baseline scenario, meaning $BLC_i = c_i$. Last, notice in Equation 3.6 that a producer may sequester c_i even if they are not enrolled, meaning $E_i^p = 0$. This type of producer has positive net returns to adopting the practice, so they adopt the practice even though their correctly assigned baseline of $\widehat{BL}_i = 1$ prevents them from receiving the incentive p_p. When calculating the additional sequestration that the payment-for-practice program produces, the second and third cases where $c_i^p = c_i$ necessitate subtracting producers' true baseline carbon sequestration values, BLC_i, from the sequestration that occurs when the program is available, c_i^p. We define the additional sequestration that occurs because of the payment-for-practice program as AS^P, and it is calculated using the following equation:

$$AS^P = \sum_i^N c_i^p - BLC_i. \qquad \text{(Eq. 3.7)}$$

For our no-till example, Equation 3.7 is the mathematical equivalent of only counting the sequestration that occurs because of individuals who made the switch from conventional tillage to no-till because of the incentive provided to them.

The policy maker, however, cannot calculate the true amount of carbon sequestration occurring due to the policy using Equation 3.7, because it does not know producers' true returns to adopting the practice and, consequently, their true baseline sequestration values, BLC_i. Instead, the policy maker estimates the amount of additional sequestration generated by its payment-for-practice program using its estimates of producers' baseline sequestration according to:

$$\widehat{AS}^P = \sum_i^N c_i^p - \widehat{BLC}_i. \qquad \text{(Eq. 3.8)}$$

Notice that even though the policy maker knows the carbon sequestration potential of producers with certainty, the policy maker's estimate of carbon sequestration resulting from the payment-for-practice program in Equation 3.8 may be significantly different than the true additional sequestration in Equation 3.7 depending on how poorly the policy maker estimates producers net returns.

If the policy maker's estimate differs from the true additional sequestration caused by the program, meaning $AS^P \neq \widehat{AS}^P$, this difference is due to inefficient participants who receive an incorrect baseline of zero carbon sequestration from the policy maker. The fraction of carbon sequestration estimated by the policy maker that is based on the behavior of these inefficient participants (IP^p) is expressed as:

$$IP^p = 1 - \frac{\sum_i^N (c_i^p - BLC_i)}{\widehat{AS}^P}. \qquad \text{(Eq. 3.9)}$$

We will also refer to the quantity of estimated sequestration contributed by inefficient participants (IP^p) as the portion of offsets from inefficient enrollees. This definition reflects the policy maker's objective of incentivizing carbon sequestration, or purchasing carbon offsets, such that Equation 3.9 represents the portion of offsets purchased by the policy maker which are non-additional. For the example involving the adoption of no-till, Equation 3.9 would indicate the fraction of the carbon sequestration that the policy maker thinks they have incentivized but is truly from individuals who were planning to adopt no-till regardless. As such, it indicates the portion of the outcome estimated by the policy maker which would have occurred in the absence of any payments.

The other source of inefficiency described earlier, inefficient nonparticipants, occurs because the policy maker overestimates some producers' net returns and assigns them a positive baseline carbon sequestration value when the true value is zero. These would be individuals who were not planning on adopting no-till next season in our example, but the policy maker does not offer them an incentive to adopt because it incorrectly predicts that they were going to use no-till in the baseline scenario. We define the efficiency loss (EL^p) as the ratio of the carbon that is not sequestered because inefficient nonparticipants receive incorrect assigned baselines to the total carbon that would be efficiently sequestered if the policy maker observed producers' net returns without error:

$$EL^p = \frac{\sum_i^N (c_i - BLC_i \mid r_i + p_p > 0 \ \& \ \widehat{BL}_i = 1)}{\sum_i^N (c_i - BLC_i \mid r_i + p_p > 0 \ \& \ BL_i = 0)}. \qquad \text{(Eq. 3.10)}$$

The final metric we construct to measure the program's success is the average cost per unit of additional sequestration (AC^p), which is calculated by dividing the total transfers from the policy maker to producers by the additional sequestration the program results in:

$$AC^p = \frac{p_p * \sum_i^N E_i^p}{AS^P}. \quad \text{(Eq. 3.11)}$$

3.3.2 Payment-for-sequestration program

The second distinguishing feature of our model is an alternative payment-for-sequestration program which offers to compensate each producer p_c per unit of carbon it sequesters above its assigned baseline carbon sequestration, $\widehat{BLC}_i$. The programs are offered in two separate states of the word, so that either the payment-for-practice or payment-for-sequestration program is available in any given scenario. As with the payment-for-practice program, we assume perfect compliance, so any producer enrolling in the payment-for-sequestration program uses the conservation agriculture practice. Additionally, we continue to assume the policy maker knows the carbon sequestration of individual producers with certainty. Given these assumptions and the per unit of sequestration incentive p_c, the binary variable E_i^c is defined as follows to indicate whether producer i enrolls in the payment-for-sequestration program:

$$E_i^c = \begin{cases} 1 \ if \ r_i + p_c * \left(c_i - \widehat{BLC}_i\right) > 0 \ and \ \widehat{BL}_i = 0 \\ 0 \ if \ r_i + p_c * \left(c_i - \widehat{BLC}_i\right) \leq 0 \ or \ \widehat{BL}_i = 1 \end{cases}. \quad \text{(Eq. 3.12)}$$

Note, in equation 3.12, that the payment-for-sequestration program shares the enrollment criteria of the payment-for-practice program requiring producers to have an assigned baseline of not using the practice ($\widehat{BL}_i = 0$). Unlike the payment-for-practice program, the compensation an individual producer receives varies as a function of its carbon sequestration potential, c_i. We define the carbon sequestration for producers when the payment-for-sequestration program is available as:

$$c_i^c = \begin{cases} c_i \ \ if \ r_i + p_c * (c_i - \widehat{BLC}_i) > 0 \\ 0 \ \ if \ r_i + p_c * (c_i - \widehat{BLC}_i) \leq 0 \end{cases}. \quad \text{(Eq. 3.13)}$$

Given these outcomes, we can define the same four metrics used to evaluate the payment-for-practice program. The additional carbon sequestration which occurs (AS^C) due to the payment-for-sequestration program is:

$$AS^C = \sum_i^N c_i^c - BLC_i. \qquad \text{(Eq. 3.14)}$$

But the sequestration estimated by the policy maker for the payment-for-sequestration program is:

$$\widehat{AS}^c = \sum_i^N c_i^c - \widehat{BLC}_i. \qquad \text{(Eq. 3.15)}$$

The fraction of sequestration estimated by the policy maker which is due to adoption by inefficient participants (IP^c) and the efficiency loss which occurs due to the incorrectly assigned baselines for inefficient nonparticipants (EL^c) are:

$$IP^c = 1 - \frac{\sum_i^N (c_i^c - BLC_i)}{\widehat{AS}^c} \qquad \text{(Eq. 3.16)}$$

and

$$EL^c = \frac{\sum_i^N (c_i - BLC_i \mid r_i + p_c * (c_i - BLC_i) > 0 \,\&\, E_i^c = 0)}{\sum_i^N (c_i - BLC_i \mid r_i + p_c * (c_i - BLC_i) > 0 \,\&\, BL_i = 0)}. \qquad \text{(Eq. 3.17)}$$

Finally, the last metric used to evaluate the success of the program is the average cost of an additional unit of sequestration:

$$AC^c = \frac{p_c * \sum_i^N E_i^c * (c_i - \widehat{BLC}_i)}{AS^C}. \qquad \text{(Eq. 3.18)}$$

3.3.3 Simulation of program performance as a function of net return observation error

To test the performance of the two program types when the policy maker knows producers' carbon sequestration with certainty, we create a simulation of our conceptual model and examine how the four metrics defined for each compensation mechanism vary as a function of the policy maker's error in observing producers' net returns, $\varepsilon_{i,r}$. The simulation is parameterized by assigning values for the two incentives, $p_p = 0.5$ and $p_c = 1$, and making three distributional assumptions for r_i, c_i, and $\varepsilon_{i,r}$.

First, we assume producers' net returns to adopting the practice in the absence of any program are distributed normally such that: $f_r(r_i) \sim N(-0.5, 1^2)$. On average, this means that roughly 31% of producers would adopt the practice without any incentive. Second, we assume producer's carbon sequestration due to practice adoption, c_i, is drawn from a standard uniform distribution, $f_f(c_i) \sim U[0,1]$. Given the incentives, $p_p = 0.5$ and $p_c = 1$, the expected compensation is equal between the two program types as: $E[f_f(c_i)] = 0.5$. Last, the policy maker's error in estimating producers' net returns, $\varepsilon_{i,r}$, is assumed to be a normally distributed random variable with mean zero and standard deviation, σ_r.

For both payment mechanisms, the simulation is run for 31 values of σ_r running from zero to 1.5 in increments of 0.05, with 800 replicates for each value of σ_r. In every model run, there are a total of 10,000 producers, or plots, simulated. The average efficiency loss (EL), average cost of a unit of additional sequestration (AC), percent of offsets from inefficient enrollees (IP), and additional sequestration (AS) per plot across all model runs for the two program types are displayed in Figure 3.1.

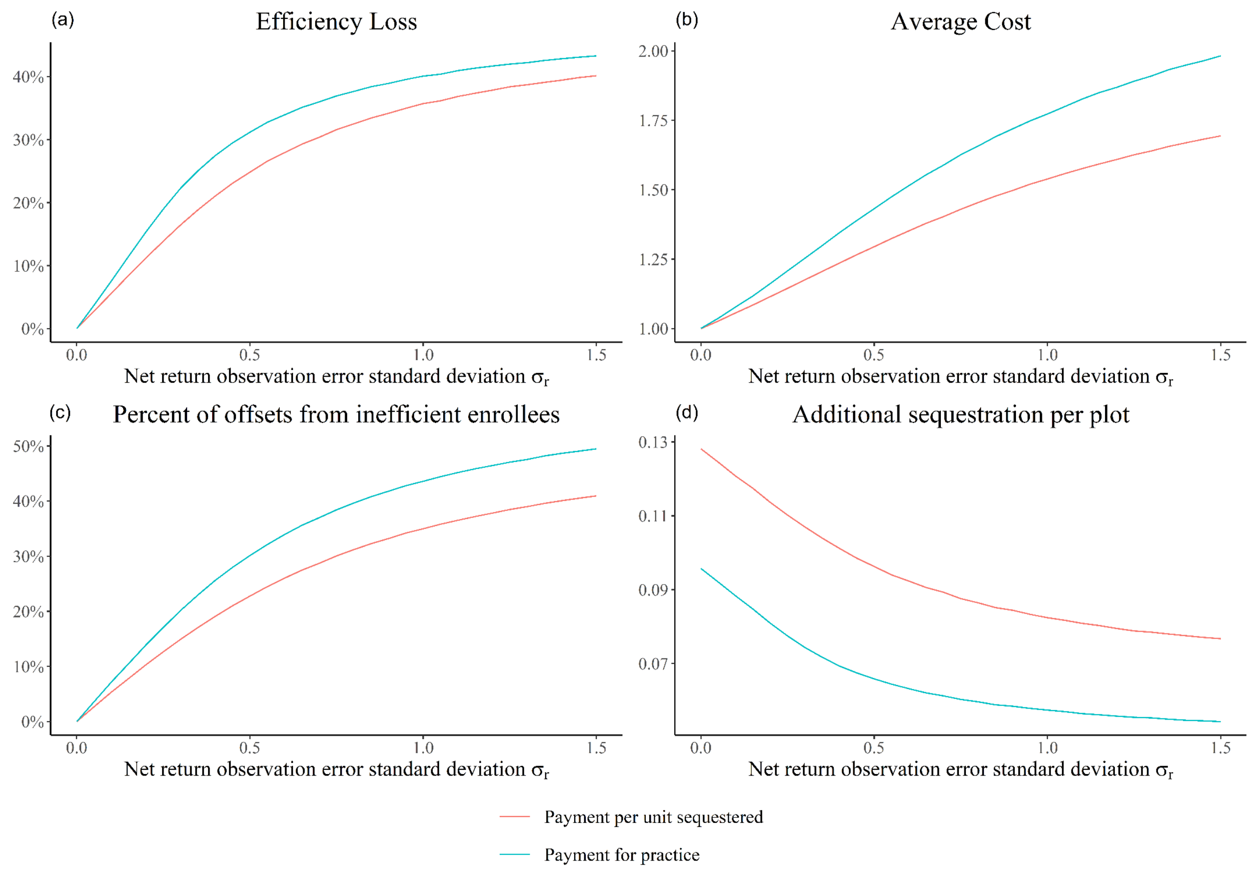

Figure 3.1. Effect of error in policy maker's estimates of producers' net returns to practice adoption. The x-axis indicates the standard deviation of the policy maker's error.

In Figure 3.1, the payment-for-sequestration program outperforms the payment-for-practice program in every metric for all values of $\sigma_r > 0$ when the policy maker knows producers' carbon sequestration with certainty. The difference between the two compensation mechanisms' performance is strictly increasing in the standard deviation of the net return observation error for two of the metrics examined, the average cost of an additional unit of sequestration in panel (b) and the percent of estimated sequestration from inefficient enrollees in panel (c) of Figure 3.1. The efficiency loss in panel (a) of Figure 3.1 initially increases more slowly for the payment-for-sequestration program as σ_r grows, but it then accelerates at intermediate levels of σ_r such that it approaches the efficiency loss of the payment-for-practice program at extreme values of σ_r. In contrast to the behavior of the other metrics, the difference in additional sequestration per plot between the two program types, displayed in panel (d) of Figure 3.1, remains constant as σ_r increases.

The reason for the improved performance of the payment-for-sequestration program and the behavior of the four metrics just described, is the built-in selection criteria of results-based compensation mechanisms. To demonstrate how this selection occurs, consider the potential carbon sequestration values for an example producer x. Letting c_x indicate the carbon producer x would sequester if it adopts the practice, then c_x can take any value on the interval $c_x \in [0,1]$ given the assumed standard uniform distribution for c_i. Now, assume producer x would efficiently enroll in the payment-for-practice program if given the opportunity. Producer x's carbon sequestration, c_x could still lie anywhere on the closed interval between zero and one, meaning $c_x \in [0,1]$, because eligibility and profitability are solely determined by producer x's returns to adoption, r_x, for the payment-for-practice program. However, if we assume producer x would efficiently enroll in the payment-for-sequestration program, the potential values of c_x are constrained to $c_x \in (|r_x|/p_c, 1]$ due to the producer's profit maximizing behavior. For producer x will only efficiently enroll if $r_x + p_c * c_x > 0$. So, when the policy maker knows producers' carbon sequestration with certainty, the payment-for-sequestration program effectively filters participants such that the average carbon sequestered by a participant will always be greater than that of the payment-for-practice program.

3.4 Including observation error for carbon sequestration

As mentioned in the introduction, programs using results-based payments include an additional informational demand because the program administrator rarely observes the outcome with certainty. The final modification we make to our model addresses this aspect of results-based payment mechanisms by introducing error in the program administrator's estimates of producers' carbon sequestration. So, like the estimates of net returns, the policy maker estimates producers carbon sequestration with some error such that: $\hat{c}_i = c_i + \varepsilon_{i,c}$. We assume the error in the policy maker's estimate of net returns, $\varepsilon_{i,c}$, is symmetric around zero, independent of f_r and $\varepsilon_{i,r}$, and has standard deviation σ_c. This does not change the values of the true baseline sequestration BLC_i from Equation 3.3, but it does alter the assigned baselines such that:

$$\widehat{BLC}_i^{\sigma_c} = \begin{cases} \hat{c}_i \; if \; \hat{r}_i > 0 \\ 0 \;\; if \; \hat{r}_i \leq 0 \end{cases}. \qquad \text{(Eq. 3.19)}$$

Now, as shown in equation 3.19, the policy maker assigns baseline carbon outcomes based on the predicted rather than true carbon sequestration for producers. We use σ_c in the superscript of terms to indicate that there is error in the policy maker's estimates of producers' carbon sequestration potential.

3.4.1 Payment-for-practice program with carbon observation error

Producers' enrollment decision is identical for the payment-for-practice program when the policy maker observes their carbon sequestration with error because compensation is based on producers' use of the practice relative to their assigned practice use baselines. This means equations 3.5, 3.6, and 3.7 produce the enrollment status (E_i^{p,σ_c}), true carbon sequestration outcomes for producers (c_i^{p,σ_c}), and total additional sequestration (AS^{p,σ_c}) when the payment-for-practice program is available despite the policy maker's error in estimating carbon sequestration. The efficiency loss (EL^{p,σ_c}) and average cost of an additional unit sequestered (AC^{p,σ_c}) are calculated using equations 3.10 and 3.11 for the same reason.

In contrast, the outcomes reported by the policy maker for the payment-for-practice program change significantly. First, we define $\hat{c}_i^{p,\sigma_c}$ as the carbon sequestration that is predicted due to producers' use of the conservation agriculture practice when the payment-for-practice program is available:

$$\hat{c}_i^{p,\sigma_c} = \begin{cases} \hat{c}_i \;\; if\; r_i + E_i^p * p_p > 0 \\ 0 \;\; if\; r_i + E_i^p * p_p \leq 0 \end{cases}. \qquad \text{(Eq. 3.20)}$$

As the policy maker no longer knows producers' sequestration potential with certainty, the policy maker's estimate for the carbon sequestered due to the program is now a function of the predicted carbon sequestration outcomes $\hat{c}_i^{p,\sigma_c}$ for enrolled producers:

$$\widehat{AS}^{p,\sigma_c} = \sum_i^N \hat{c}_i^{p,\sigma_c} * E_i^{p,\sigma_c}. \qquad \text{(Eq. 3.21)}$$

Consequently, the fraction of sequestration expected by the policy maker based on the sequestration of inefficient participants (IP^p, σ_c) becomes:

$$IP^{p,\sigma_c} = \frac{\sum_i^N (\hat{c}_i^{p,\sigma_c} * E_i^{p,\sigma_c} \mid BL_i = 1)}{\widehat{AS}^{p,\sigma_c}}. \qquad \text{(Eq. 3.22)}$$

3.4.2 Payment-for-sequestration with carbon observation error

For the payment-for-sequestration program, there are changes to both the producers' enrollment decisions and the outcomes reported by the program administrator. Each producer's enrollment decision is altered because its potential compensation is calculated using the policy maker's estimates of carbon sequestration:

$$E_i^{c,\sigma_c} = \begin{cases} \left(1 - \widehat{BL}_i\right) if\; r_i + p_c * \left(\hat{c}_i - \widehat{BLC}_i^{\sigma_c}\right) > 0 \;and\; \left(\hat{c}_i - \widehat{BLC}_i^{\sigma_c}\right) > 0 \\ 0\; if\; r_i + p_c * \left(\hat{c}_i - \widehat{BLC}_i^{\sigma_c}\right) \leq 0 \;or\; \left(\hat{c}_i - \widehat{BLC}_i^{\sigma_c}\right) \leq 0 \end{cases}. \qquad \text{(Eq. 3.23)}$$

In addition to using the estimated carbon sequestration potential for producers, the enrollment decision in equation 3.23 differs from the case without carbon sequestration error due to the restrictions on the predicted carbon sequestration above the assigned baseline, $\left(\hat{c}_i - \widehat{BLC}_i^{\sigma_c}\right)$.

These restrictions account for the possibility that the estimated carbon sequestration associated with using a practice is negative due to the policy maker's error in estimating a producer's carbon sequestration potential. If an eligible producer's return to adopting a practice is positive but the predicted carbon sequestration is negative, we assume the producer would use the conservation agriculture practice but would not enroll in the program as: $r_i > r_i + p_c * (\hat{c}_i - \widehat{BLC}_i^{\sigma_c})$. Using equation 3.23, we construct the true and estimated carbon sequestration by producers when the payment-for-sequestration program is available:

$$c_i^{c,\sigma_c} = \begin{cases} c_i \ \ if\ r_i + p_c * \hat{c}_i * E_i^{c,\sigma_c} > 0 \\ 0 \ \ if\ r_i + p_c * \hat{c}_i * E_i^{c,\sigma_c} \leq 0 \end{cases}; \qquad \text{(Eq. 3.24)}$$

$$\hat{c}_i^{c,\sigma_c} = \begin{cases} \hat{c}_i \ \ if\ r_i + p_c * \hat{c}_i * E_i^{c,\sigma_c} > 0 \\ 0 \ \ if\ r_i + p_c * \hat{c}_i * E_i^{c,\sigma_c} \leq 0 \end{cases}. \qquad \text{(Eq. 3.25)}$$

The true additional sequestration which occurs (AS^{c,σ_c}) and the sequestration estimated by the policy maker ($\widehat{AS}^{c,\sigma_c}$) for the payment-for-sequestration program are therefore:

$$AS^{c,\sigma_c} = \sum_i^N c_i^{c,\sigma_c} - BLC_i\ ; \qquad \text{(Eq. 3.26)}$$

$$\widehat{AS}^{c,\sigma_c} = \sum_i^N \hat{c}_i^{c,\sigma_c} * E_i^c\ . \qquad \text{(Eq. 3.27)}$$

Constructing the average cost per unit of additional sequestration for the payment-for-sequestration program with error in estimating carbon sequestration involves minor changes to equation 3.18:

$$AC^{c,\sigma_c} = \frac{p_c * \sum_i^N E_i^{c,\sigma_c} * (\hat{c}_i^{c,\sigma_c} - \widehat{BLC}_i)}{AS^{c,\sigma_c}}\ . \qquad \text{(Eq. 3.28)}$$

To define the fraction of offsets due to inefficient participation (IP^{c,σ_c}) and the efficiency loss (EL^{c,σ_c}) in this context, we must first consider how including carbon sequestration error affects each of the two categories of inefficient producers. It is still possible to have non-additional enrollment by producers if the policy maker incorrectly assigns a baseline of not using the practice when the opposite is true.

But unlike when the policy maker knows the carbon sequestration potential of producers with certainty, there is also another group of inefficient participants who enroll and alter their behavior relative to their baselines because the policy maker overestimates the carbon they will sequester. These participants are inefficient because the profit-maximizing choice for them, in the absence of error estimating carbon sequestration, would be to refrain from using the practice. Despite this, the sequestration occurring due to this second type of inefficient participants is included in AS^{c,σ_c} because it truly is additional relative to these producers' baselines. With error in carbon sequestration estimates, the fraction of the sequestration estimated by the policy maker due to inefficient participants combines the non-additional enrollees as well as the producers inefficiently participating due to overestimated values for their carbon sequestration potential:

$$IP^{c,\sigma_c} = \frac{\sum_i^N\left(\hat{c}_i^{c,\sigma_c} * E_i^{c,\sigma_c} \mid BL_i=1\right)}{\widehat{AS}^{c,\sigma_c}} + \frac{\sum_i^N\left(\hat{c}_i^{c,\sigma_c} * E_i^{c,\sigma_c} \mid BL_i=0 \ \& \ r_i + p_c * c_i \le 0\right)}{\widehat{AS}^{c,\sigma_c}}. \qquad \text{(Eq 3. 29)}$$

The first term in equation 3.29 is the fraction offsets issued based on the predicted sequestration of non-additional enrollees whose baselines were incorrectly estimated, and the second term is the fraction of offsets issued based on the estimated carbon sequestration of producers who were assigned and overly generous carbon sequestration value.

The inclusion of error in carbon sequestration estimates necessitates a similar adjustment to calculate efficiency loss (EL^{c,σ_c}) as well. There is still a possibility of producers inefficiently not participating in the program because of an incorrectly assigned baseline that indicates they would have used the practice in the absence of the program when the opposite is true. But, even if the policy maker correctly estimates returns to adoption, it is possible that a producer inefficiently refrains from enrolling enroll because the policy maker underestimates its carbon sequestration potential. The efficiency loss for the payment-for-sequestration program when there is error in the policy maker's estimates of producers' carbon sequestration potential (EL^{c,σ_c}) includes the carbon sequestration missed out on from both types of inefficient non-participants:

$$EL^{c,\sigma_c} = \frac{\sum_i^N (c_i - BLC_i \mid r_i + p_c * (c_i - BLC_i) > 0 \;\&\; E_i^{c,\sigma_c} = 0)}{\sum_i^N (c_i - BLC_i \mid r_i + p_c * (c_i - BLC_i) > 0 \;\&\; BL_i = 0)} \qquad \text{(Eq. 3.30)}$$

3.4.3 Simulation of program performance as a function of net return and carbon sequestration observation error

As earlier, we now test the performance of the two program types when the policy maker estimates producers' sequestration potential with error. We use the four metrics defined above for the case involving carbon observation error and examine how they vary as a function of the policy maker's errors in observing producers' net returns, $\varepsilon_{i,r}$, and carbon sequestration potential, $\varepsilon_{i,c}$. We use the same values as in the earlier simulation for the two incentives, $p_p = 0.5$ and $p_c = 1$, and the same distributional assumptions for r_i, c_i, and $\varepsilon_{i,r}$: $f_r(r_i) \sim N(-0.5, 1^2)$, $f_f(c_i) \sim U[0,1]$, and $\varepsilon_{i,r} \sim N(0, \sigma_r)$. For the policy maker's estimates of producers' carbon sequestration potential $\varepsilon_{i,c}$, we assume their error in observing c_i is drawn from a normal distribution with mean zero and standard deviation σ_c.

In figure 3.2, we depict the average efficiency loss (EL^{σ_c}), average cost of a unit of additional sequestration (AC^{σ_c}), percent of offsets from inefficient enrollees (IP^{σ_c}), and additional sequestration (AS^{σ_c}) per plot as a function of the standard deviation in the net return observation error (σ_r) for three values of the standard deviation in carbon sequestration error: $\sigma_c = 0, 0.2, \& \ 0.4$. We do not display the values of σ_c for the payment-for-practice program results because they are identical across all values of σ_c. This is because the policy maker's estimates of carbon sequestration are unbiased given $\varepsilon_{i,c} \sim N(0, \sigma_c)$.

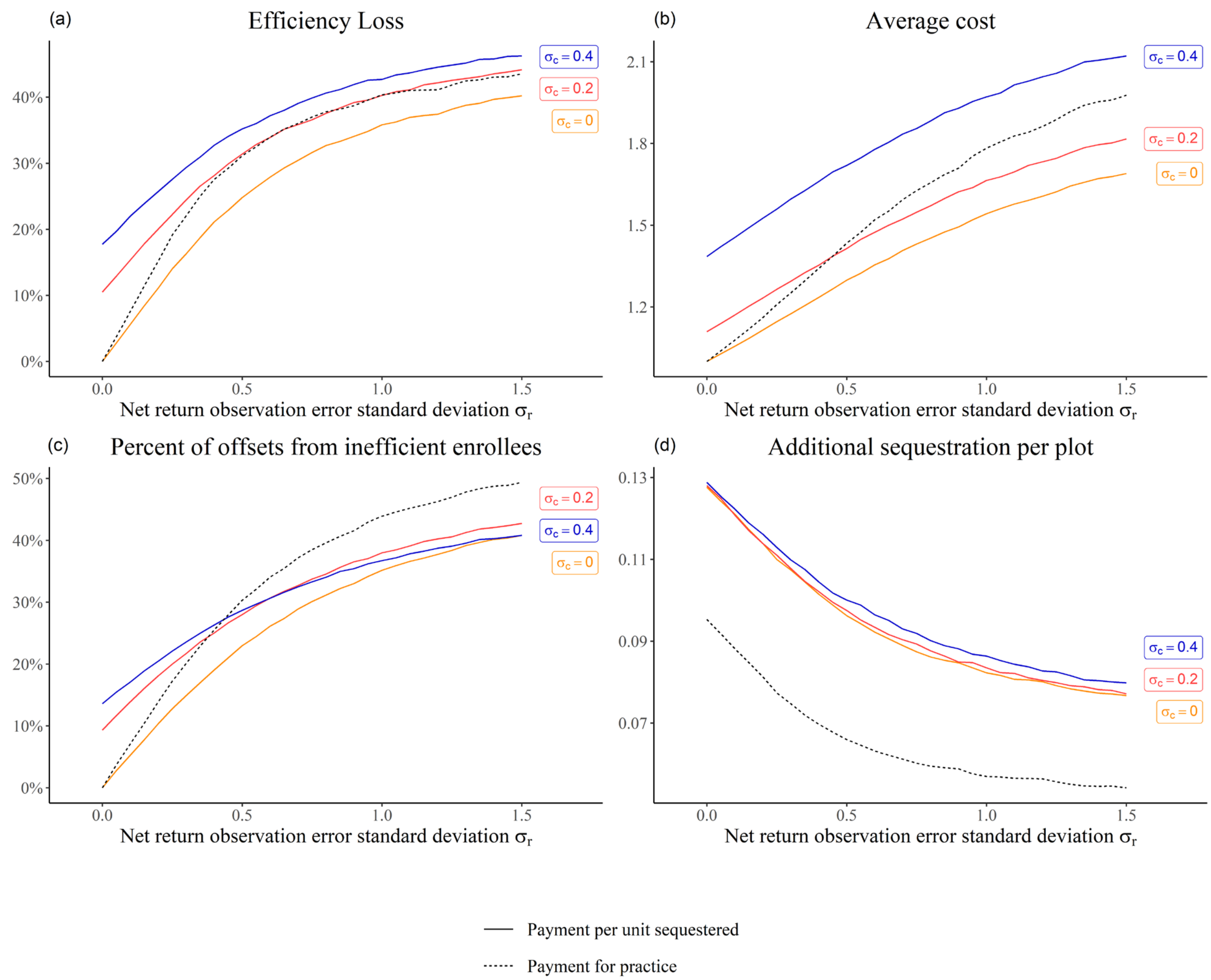

Figure 3.2. Effect of error in policy maker's estimates of producers' net returns to, and carbon sequestration produced by, practice adoption. The x-axis indicates the standard deviation of the policy maker's error, $\boldsymbol{\sigma_r}$. Individual lines are labelled with their respective values for the standard deviation of the policy maker's error in estimating carbon sequestration, $\boldsymbol{\sigma_c}$.

As in Figure 3.1, the payment-for-sequestration program results in panel (d) of Figure 3.2 depict greater additional sequestration for all values of net return observation error (σ_r), and this holds true for all values of carbon sequestration observation error (σ_r). For the rest of the metrics, however, the payment-for-practice program sometimes outperforms the payment-for-sequestration program. For example, the average cost of an additional unit of sequestration in panel (b) and the percent of offsets from inefficient enrollees in panel (c) of Figure 3.2 are both lower for the payment-for-practice program when the net return observation error (σ_r) is relatively low in comparison to the carbon sequestration observation error (σ_c). Conversely, when the net return observation error is relatively high, the payment-for-sequestration program outperforms the payment-for-practice program in terms of its average cost and percent of offsets from inefficient enrollees. This does not hold true, however, at very high levels of error in the policy maker's estimates of carbon sequestration. For the highest value of carbon sequestration observation error, $\sigma_c = 0.4$, the payment-for-practice program's efficiency loss and average cost are lower for all values of net return observation error, σ_r.

From the results in Figure 3.2, the payment-for-sequestration program's efficiency is clearly reduced by the presence of error in the policy maker's estimates of carbon sequestration. When there is any error in carbon sequestration estimates ($\sigma_c > 0$), the payment-for-sequestration program suffers large efficiency losses even in the absence of net return observation error ($\sigma_r = 0$). These efficiency losses represent the carbon which is not sequestered because producers received carbon sequestration estimates below their real sequestration potential. However, as shown in panel (d) of Figure 3.2, the greater additional sequestration for the payment-for-sequestration program indicates the carbon represented by these efficiency losses was somehow recouped. The curves for the payment-for-sequestration program in panel (c) of Figure 3.2 depict the new source of additional carbon: producers who would not have participated in the absence of carbon sequestration error but who enroll in the program because the policy maker overestimates their carbon sequestration. The carbon sequestered by these individuals is truly additional, but the policy maker is overpaying these individuals due to its overly generous sequestration estimates. The cumulative result of these effects is the increasing average cost of the payment-for-sequestration program.

3.5 Effect of increasing baseline scale on voluntary carbon credit program performance

To examine the impact of scale on the efficiency of the payment-for-practice and payment-for-sequestration programs, we first redefine the program enrollment decision and outcomes at a jurisdictional level. We consider a setting identical to that of the previous section involving error in policy maker's estimates of producers' net returns to practice adoption and the sequestration due to each producer's use of the practice. However, instead of evaluating outcomes across all individuals indexed by i, we now consider the aggregate outcomes for all producers in identically sized jurisdictions indexed by j. We hold the total number of producers (N) constant and define the number of jurisdictions as N_j, so the number of producers in each jurisdiction is $N_{i,j} = N/\ N_j$.

Within each jurisdiction, the individual producers are indexed by i running from $i = 1$ to $i = N_{i,j}$. The characteristics of individual producer's, $r_{i,j}$ and $c_{i,j}$, share the same features as their counterparts in the individual-scale model, r_i and c_i. The baseline adoption behavior for the conservation agriculture practice in jurisdiction j is the sum of the individual baselines for all producers in jurisdiction j such that the true baseline use of the practice in jurisdiction j is:

$$BL_j = \sum_{i=1}^{N_{i,j}} BL_{i,j}\ ;\ where\ BL_{i,j} = \begin{cases} 1 & if\ r_{i,j} > 0 \\ 0 & if\ r_{i,j} \leq 0 \end{cases}. \quad \text{(Eq. 3.31)}$$

The carbon that would be sequestered in jurisdiction j based on its baseline practice use is:

$$BLC_j = \sum_{i=1}^{N_{i,j}} BLC_{i,j}\ ;\ where\ BLC_{i,j} = \begin{cases} c_{i,j} & if\ r_{i,j} > 0 \\ 0 & if\ r_{i,j} \leq 0 \end{cases}. \quad \text{(Eq. 3.32)}$$

Like in the individual-scale case, the policy maker assigns a baseline level of practice use for each jurisdiction based on its estimates of producers' returns to adopting the practice, $\hat{r}_{i,j} = r_{i,j} + \varepsilon_{i,j,r}$. We continue to assume the error in the policy maker's estimate of net returns, $\varepsilon_{i,j,r}$, is symmetric around zero, independent of f_r, and has standard deviation σ_r. The policy maker constructs its prediction of jurisdiction j's baseline by adding up the assigned baselines of all the constituent producers:

$$\widehat{BL}_j = \sum_{i=1}^{N_{i,j}} \widehat{BL}_{i,j}\,;\ where\ \widehat{BL}_{i,j} = \begin{cases} 1 \ \ if\ \hat{r}_{i,j} > 0 \\ 0 \ \ if\ \hat{r}_{i,j} \leq 0 \end{cases}. \qquad \text{(Eq. 3.33)}$$

In similar fashion, the policy maker continues to estimate producers' carbon sequestration with some error such that: $\hat{c}_{i,j} = c_{i,j} + \varepsilon_{i,j,c}$. We likewise assume the error in the policy maker's estimate of net returns, $\varepsilon_{i,j,c}$, is symmetric around zero, independent of f_r and $\varepsilon_{i,j,r}$, and has standard deviation σ_c. As such, the baseline carbon sequestration that the policy maker assigns for jurisdiction j is:

$$\widehat{BLC}_j = \sum_{i=1}^{N_{i,j}} \widehat{BLC}_{i,j}\,;\ where\ \widehat{BLC}_{i,j} = \begin{cases} \hat{c}_{i,j} \ \ if\ \hat{r}_{i,j} > 0 \\ 0 \ \ if\ \hat{r}_{i,j} \leq 0 \end{cases}. \qquad \text{(Eq. 3.34)}$$

3.5.1 Jurisdiction scale payment-for-practice program

For the jurisdictional payment-for-practice program, each jurisdiction is offered a fixed compensation p_p times the number of its constituent producers who adopt the conservation agriculture practice above the assigned jurisdictional baseline. As an example, consider a program in which countries are compensated on a per-acre basis for reforestation efforts. The jurisdiction, a country in this example, would be assigned a baseline that reflects the expected forest acreage across its entire geographic extent in the absence of any reforestation initiatives. To determine its compensation, the policy maker would subtract this country-level baseline from the total forested acreage observed after the country institutes its payment-for-practice reforestation program and multiply the result by p_p. Instead of measuring whether individual producers in each country reforest their particular plots of land, this jurisdictional approach focuses on the aggregate outcomes across all members of jurisdictions.

Note that if each jurisdiction contains a single producer ($N_{i,j} = 1$), this is identical to the individual-scale payment-for-practice program examined above. But, when the jurisdiction size is greater than one, it is necessary to define how each jurisdiction decides whether to participate or not. To do so, similar to van Benthem and Kerr (2013), we assume that jurisdictions know the returns to adoption with certainty and that they are informed of their assigned jurisdictional baseline for practice use ($\widehat{BL}_j$) by the policy maker. In essence, the assumption that jurisdictions know their constituents returns to adoption with certainty is akin to stating each jurisdiction can perfectly predict the total cost of adoption among its constituents for any given level of adoption.

To demonstrate the jurisdictions' enrollment decision-making process, consider a jurisdiction j which is offered the payment-for-practice incentive by the policy maker p_p. Jurisdiction j computes the potential benefit to enrollment by totaling how many of its constituent producers would adopt the practice given incentive p_p. We define $N_j^{p_p}$ as this total number of producers in jurisdiction j who would adopt given the incentive p_p, and note that this sum includes producers who would have adopted without an incentive. Then, jurisdiction j's potential return to enrollment is: $p_p * \left(N_j^{p_p} - \widehat{BL}_j\right)$. To determine the cost of enrolling, each jurisdiction then totals the individual cost of adoption across all its constituents who would alter their behavior when offered p_p, yielding: $\sum_{i|r_{i,j} \in (-p_p, 0]}^{N_{i,j}} r_{i,j}$. A jurisdiction will enroll in the payment-for-practice program, reflected by the indicator variable E_j^p equaling one, if the potential benefits from enrollment are greater than the total cost of adoption born by its constituents:

$$E_j^p = \begin{cases} 1 \ if \ p_p * \left(N_j^{p_p} - \widehat{BL}_j\right) > \sum_{i|r_{i,j} \in (-p_p, 0]}^{N_{i,j}} r_{i,j} \\ 0 \ if \ p_p * \left(N_j^{p_p} - \widehat{BL}_j\right) \leq \sum_{i|r_{i,j} \in (-p_p, 0]}^{N_{i,j}} r_{i,j} \end{cases} . \quad \text{(Eq. 3.35)}$$

In equation 3.35, notice that the jurisdiction does not make its enrollment decision based on the expenditures it would make internally for each of its constituents who use the practice, so we are not modeling how the jurisdiction disperses payments amongst producers. Instead, we assume that all $N_j^{p_p}$ producers in jurisdiction j will adopt the conservation agriculture practice if the jurisdiction enrolls in the payment-for-practice program, meaning $E_j^p = 1$.

To determine the efficiency of the jurisdiction-scale payment-for-practice program in the presence of error in both the estimates of net returns and carbon sequestration potential, we focus our analysis on the average cost of an additional unit of sequestration ($AC^{p,j}$). The j component in the superscript of $AC^{p,j}$ indicates this term refers to the average cost of the payment-for-practice program at a jurisdictional scale. To calculate the average cost, we add up the cost of the program to the policy maker across all jurisdictions to produce:

$$TC^{p,j} = \sum_{j=1}^{N_j} E_j^p * p_p * \left(N_j^{p_p} - \widehat{BLC}_j\right) . \quad \text{(Eq. 3.36)}$$

Then, we find the total additional sequestration that occurs because of jurisdictions' enrollment decisions using:

$$AS^{p,j} = \sum_{j=1}^{N_j} E_j^p * \sum_{i|r_{i,j} \in (-p_p, 0]}^{N_{i,j}} c_{i,j} . \quad \text{(Eq. 3.37)}$$

The average cost per unit of additional sequestration is then:

$$AC^{p,j} = \frac{\sum_{j=1}^{N_j} E_j^p * p_p * \left(N_j^{p_p} - \widehat{BLC}_j\right)}{\sum_{j=1}^{N_j} E_j^p * \sum_{i|r_{i,j} \in (-p_p, 0]}^{N_{i,j}} c_{i,j}} . \quad \text{(Eq. 3.38)}$$

3.5.2 Jurisdiction scale payment-for-sequestration program

When the jurisdictional payment-for-sequestration program is available, each jurisdiction is offered compensation p_c times the amount of carbon the policy maker estimates is sequestered above the assigned jurisdictional baseline ($\widehat{BLC_j}$). Like for the jurisdictional payment-for-practice program, we assume the jurisdiction knows the net returns of its producers with certainty and determines its returns to participation based on its constituent producers' potential returns given the payment-for-sequestration incentive it receives from the policy maker, p_c. As was the case for the jurisdictional payment-for-practice program, if jurisdictions are comprised of one producer ($N_{i,j} = 1$), the jurisdiction-scale and individual-scale payment-for-sequestration programs are identical. However, unlike producers' returns to practice adoption, we do not assume that the jurisdiction knows its constituents' true carbon sequestration potential ($c_{i,j}$) with certainty. Instead, we assume the jurisdiction relies on the policy maker's estimates of its constituent producers' carbon sequestration potential ($\hat{c}_{i,j}$) when calculating individual producers' returns in the jurisdictional payment-for-sequestration program.

When $N_{i,j} > 1$ for the jurisdictional payment-for-sequestration program, jurisdictions compute their return to enrollment by totaling the estimated carbon sequestration of all constituent producers who would adopt the practice when offered $p_c * \hat{c}_{i,j}$. Let $\hat{C}_j^{p_c}$ be this total estimated carbon sequestration of producers in jurisdiction j who would participate given the incentive $p_c * \hat{c}_{i,j}$, defined as:

$$\hat{C}_j^{p_c} = \sum_{i|r_{i,j}+p_c*\hat{c}_{i,j} > 0 \,\&\, \hat{c}_{i,j}> 0}^{N_{i,j}} \hat{c}_{i,j} \ . \qquad \text{(Eq. 3.39)}$$

Jurisdiction j's potential return to enrollment is then $p_c * \left(\hat{C}_j^{p_c} - \widehat{BLC_j}\right)$, and the total cost of participating summed across all individuals in a jurisdiction who would alter their behavior is

$\sum_{i|r_{i,j}\in(-p_c*\hat{c}_{i,j},0]\ \&\ \hat{c}_{i,j}>0}^{N_{i,j}} r_{i,j}$. Taken together, these two quantities determine if jurisdiction j will enroll in the jurisdictional payment-for-sequestration program, reflected by the indicator variable E_j^c equaling one, according to:

$$E_j^c = \begin{cases} 1\ if\ p_c * \left(\hat{C}_j^{p_c} - \widehat{BLC}_j\right) > \sum_{i|r_{i,j}\in(-p_c*\hat{c}_{i,j},0]\ \&\ \hat{c}_{i,j}>0}^{N_{i,j}} r_{i,j} \\ 0\ if\ \ p_c * \left(\hat{C}_j^{p_c} - \widehat{BLC}_j\right) \leq \sum_{i|r_{i,j}\in(-p_c*\hat{c}_{i,j},0]\ \&\ \hat{c}_{i,j}>0}^{N_{i,j}} r_{i,j} \end{cases}. \qquad \text{(Eq. 3.40)}$$

So, when a jurisdiction chooses to participate, meaning $E_j^c = 1$, the jurisdiction's total estimated carbon sequestration across its constituent producers equals $\hat{C}_j^{p_c}$. Like for the payment-for-practice jurisdictional program, we focus our analysis on the average cost of an additional unit of sequestration ($AC^{c,j}$). The total cost of the program born by the policy maker across all jurisdictions is defined as:

$$TC^{c,j} = \sum_{j=1}^{N_j} E_j^c * p_c * \left(\hat{C}_j^{p_c} - \widehat{BLC}_j\right). \qquad \text{(Eq. 3.41)}$$

While the total additional sequestration due to the jurisdictional payment-for-sequestration program is:

$$AS^{c,j} = \sum_{j=1}^{N_j} E_j^p * \sum_{i|r_{i,j}\in(-p_c*\hat{c}_{i,j},0]\ \&\ \hat{c}_{i,j}>0}^{N_{i,j}} c_{i,j}. \qquad \text{(Eq. 3.42)}$$

The average cost per unit of additional sequestration is then:

$$AC^{c,j} = \frac{\sum_{j=1}^{N_j} E_j^c * p_c * \left(\hat{C}_j^{p_c} - \widehat{BLC}_j\right)}{\sum_{j=1}^{N_j} E_j^p * \sum_{i|r_{i,j}\in(-p_c*\hat{c}_{i,j},0]\ \&\ \hat{c}_{i,j}>0}^{N_{i,j}} c_{i,j}}. \qquad \text{(Eq. 3.43)}$$

3.5.3 Numerical simulation of jurisdictional voluntary carbon credit programs

To test the performance of the jurisdictional versions of the payment-for-practice and payment-for-sequestration programs, we narrow our focus to the impact of jurisdiction size ($N_{i,j}$) on the average cost of an additional unit of sequestration ($AC^{p,j}, AC^{c,j}$). To account for the impact of error in the policy maker's estimates of net returns and carbon sequestration, we test this relationship at several values for the standard deviations in returns to adoption (σ_r) and carbon sequestration (σ_c). We make the same distributional assumptions as in the earlier simulations such that $f_r(r_{i,j}) \sim N(-0.5, 1^2)$, $f_f(c_{i,j}) \sim U[0,1]$, $\varepsilon_{i,j,r} \sim N(0, \sigma_r)$, and $\varepsilon_{i,j,c} \sim N(0, \sigma_c)$. We also use the same values for the two incentives, $p_p = 0.5$ and $p_c = 1$, and hold the total number of producers in the simulation constant across runs at $N = 10{,}000$.

In figure 3.3, we depict the average cost of an additional unit of sequestration ($AC^{p,j}, AC^{c,j}$) as function of the jurisdiction size ($N_{i,j}$) for nine different scenarios. The nine scenarios depicted are defined by the combination of three values for the standard deviation in policy maker's estimates of returns to adoption ($\sigma_r = 0.1, 0.3, 0.5$) and carbon sequestration ($\sigma_c = 0.1, 0.2, 0.3$). The plotted lines depict the mean average cost of an additional unit of sequestration ($AC^{p,j}, AC^{c,j}$) across the 50 simulation runs performed for each combination of jurisdiction size ($N_{i,j}$), standard deviation in return observation error (σ_r), and standard deviation in carbon sequestration error (σ_c).

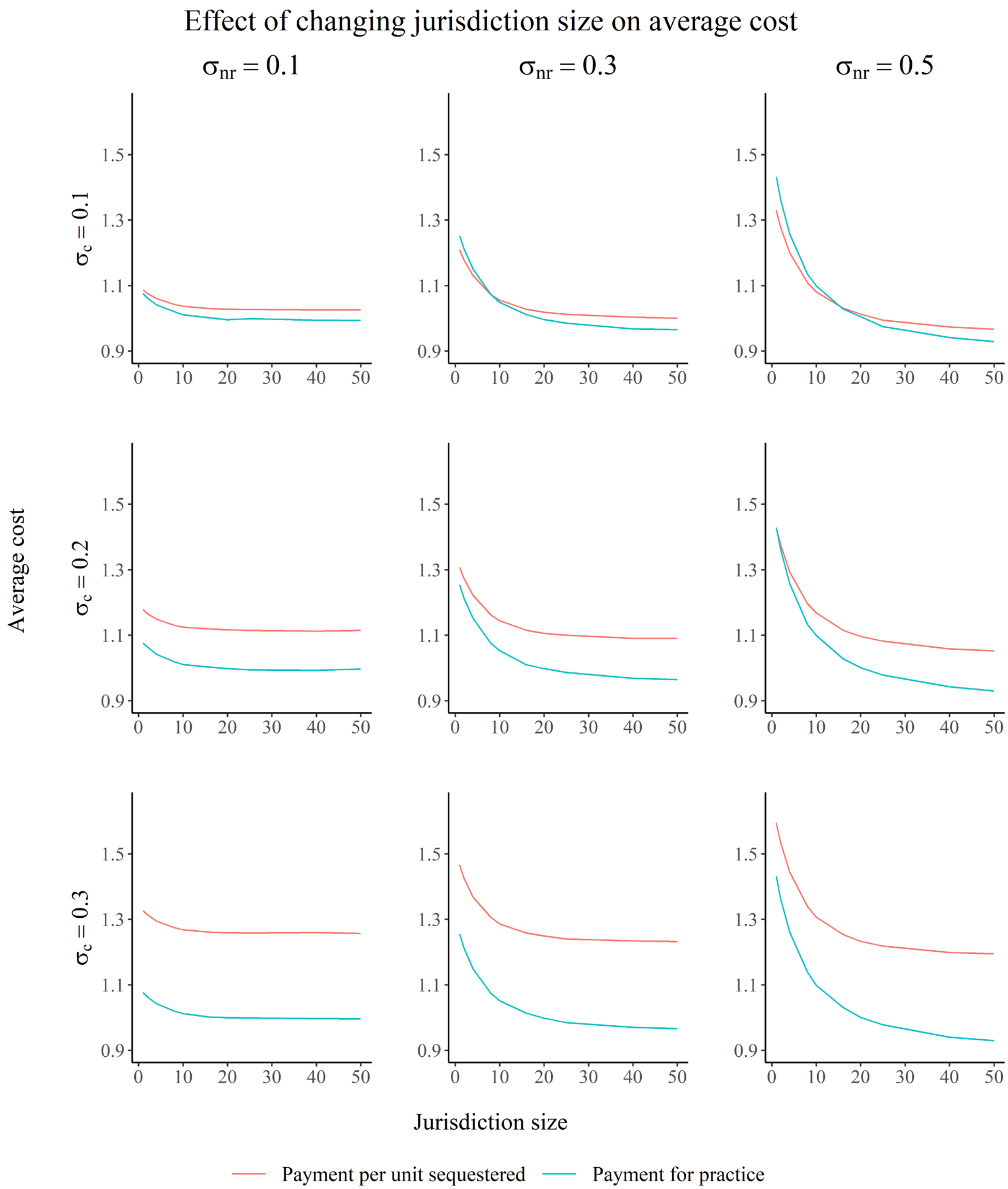

Figure 3.3. Effect of jurisdiction size on the average cost of an additional unit of sequestration by compensation mechanism.

The top row of graphs in figure 3.3 depict the effect of increasing the size of jurisdictions when the error in policy maker's estimates is lowest ($\sigma_c = 0.1$). As expected from the earlier results displayed in figure 3.2, the average cost of an additional unit of sequestration is lower for the payment-for-sequestration program when jurisdictions are small in size and there is little error in carbon sequestration estimates relative to the error in estimates of net returns. But, across all three values for σ_r, the payment-for-practice program eventually becomes the more cost-effective program as the size of jurisdictions increases when $\sigma_c = 0.1$. One potential reason for this is the selection for greater carbon contributions involved in the payment-for-sequestration payment mechanism. In comparing the graphs in the top rows of figure 3.3, it becomes clear that error in the policy maker's estimates of net returns (σ_r) mediates the relationship between scale and average cost of the payment-for-practice program. For example, holding the error for carbon sequestration constant at $\sigma_c = 0.1$, an increase in σ_r from 0.2 to 0.3 causes the jurisdiction size at which the payment-for-practice program outperforms the payment-for-sequestration program to increase.

For the intermediate or high levels of error in estimates of carbon sequestration ($\sigma_c = 0.2, 0.3$), shown in the second and third rows of graphs in figure 3.3, the greater impact of increasing jurisdiction size on the payment-for-practice program causes it to outperform the payment-for-sequestration program in almost every circumstance. The exception to this statement is the case with the intermediate level of error in carbon sequestration estimates ($\sigma_c = 0.2$), very small jurisdiction sizes, and the highest level of error in estimates of producers' net returns, ($\sigma_5 = 0.5$). One reason for the markedly superior performance of the payment-for-practice program is our assumption that the jurisdiction uses the policy maker's estimates of carbon sequestration ($\hat{c}_{i,j}$) to calculate the returns to participation for its constituents. This assumption implies that the jurisdiction does not have additional information on the rates of sequestration amongst its producers in comparison to the policy maker.

Overall, the effect of increasing program scale mitigates the inefficiencies caused by the policy maker's error in estimating producers' net returns to a greater degree. As the payment-for-practice program is only affected by the inefficiencies related to net return observation error, its performance improves to a greater degree as the scale of the program increases. The insensitivity of the payment-for-practice program's efficiency to the policy maker's error in estimating carbon sequestration is readily apparent when viewing a single column of figure 3.3. In contrast, the payment-for-sequestration program, being affected by the inefficiencies born of carbon sequestration error, benefits from the effect of increasing the program's scale to a lesser degree.

3.6 Conclusion

In this work we first demonstrate that the superior performance of a payment-for-sequestration, or results-based, carbon offset program relies on the policy maker's ability to accurately predict the sequestration outcomes of individual producers. The payment-for-sequestration program excels because it selects for producers whose actions produce larger results, a greater increase in carbon sequestration due to adoption of a conservation agriculture practice in this context. Furthermore, the selection caused by the payment-for-practice compensation mechanism causes the results-based policy to outperform the actions-based offset policy even when the policy maker predicts producers' counterfactual behavior with significant error. However, the unequivocal, superior performance of the payment-for-sequestration program only occurs when the policy maker knows the results of producers' actions, or the sequestration each producer would contribute, with certainty.

When we introduce error in the policy makers estimates of the carbon sequestration resulting from producers using the practice, our results indicate that the payment-for-practice program outperforms the payment-for-sequestration program when the uncertainty in carbon sequestration values is high relative to the uncertainty in estimates of producers' returns to adoption. This result builds on the conclusions of White and Hanley (2016) who find that an input-based contract, analogous to the payment-for-practice program, is more cost effective than an output-based contract, similar to our payment-for-sequestration program, because it has a lower cost approach to adverse selection. With our model, we demonstrate that the cost of either approach to adverse selection, meaning the compensation mechanism chosen, will depend on the policy maker's ability to accurately estimate crucial producer characteristics. For example, the payment-for-sequestration mechanism selecting for producers with greater sequestration potential becomes increasingly costly as the error in carbon sequestration estimates rises. This increasing cost of mitigating adverse selection is due to the growth in costly contributions of inefficient participants who are assigned inflated carbon sequestration baselines.

Finally, in examining the impact of a jurisdictional approach to setting baselines and evaluating outcomes, we find that the positive relationship between program scale and efficiency established in van Benthem and Kerr (2013) applies to both the payment-for-practice and payment-for-sequestration programs. In addition, we find that increasing the program scale improves the cost-efficiency of the payment-for-practice program to a greater degree, such that it becomes the more cost-effective jurisdictional program in the majority of scenarios. Future work will clarify whether this result is due to the assumptions made in our work concerning the independence of model parameters or the distributions used in the numerical simulations.

Bibliography

117th Congress. 2022. *Inflation Reduction Act of 2022*.

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *Review of Economic Studies* 72: 1–19.

Alcott, Blake. 2005. "Jevons' Paradox." *Ecological Economics* 54 (1): 9–21.

Allen, Richard G, Luis Pereira, and Martin Smith. 1998. *Crop Evapotranspiration: Guidelines for Computing Crop Requirements*. Rome: Food and Agriculture Organization of the United Nations.

Antle, John, Susan Capalbo, Siân Mooney, Edward Elliott, and Keith Paustian. 2003. "Spatial Heterogeneity, Contract Design, and the Efficiency of Carbon Sequestration Policies for Agriculture." *Journal of Environmental Economics and Management* 46 (2): 231–50.

Antle, John M., Susan M. Capalbo, Keith Paustian, and Md Kamar Ali. 2007. "Estimating the Economic Potential for Agricultural Soil Carbon Sequestration in the Central United States Using an Aggregate Econometric-Process Simulation Model." *Climatic Change* 80 (1–2): 145–71.

Augustin, Christopher, and Larry J. Cihacek. 2016. "Relationships between Soil Carbon and Soil Texture in the Northern Great Plains." *Soil Science* 181 (8): 386–92.

Benthem, Arthur van, and Suzi Kerr. 2013. "Scale and Transfers in International Emissions Offset Programs." *Journal of Public Economics* 107: 31–46.

Berbel, Julio, Carlos Gutiérrez-Martín, Juan A. Rodríguez-Díaz, Emilio Camacho, and Pilar Montesinos. 2015. "Literature Review on Rebound Effect of Water Saving Measures and Analysis of a Spanish Case Study." *Water Resources Management* 29 (3): 663–78.

Borusyak, Kirill. 2018. "Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume."

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32.

Breiman, Leo, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone. 1984. *Classification and Regression Trees*. 1st Editio. New York: Routledge.

Brookes, Leonard. 2000. "Energy Efficiency Fallacies Revisited." *Energy Policy* 28 (6): 355–66.

Bruce, James P, Michele Frome, Eric Haites, Henry Janzen, Rattan Lal, and Keith Paustian. 1999. "Carbon Sequestration in Soils." *Journal of Soil and Water Conservation* 54 (1): 382–89.

Cai, Andong, Wenting Feng, Wenju Zhang, and Minggang Xu. 2016. "Climate, Soil Texture, and Soil Types Affect the Contributions of Fine-Fraction-Stabilized Carbon to Total Soil Organic Carbon in Different Land Uses across China." *Journal of Environmental Management* 172: 2–9.

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2020. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230.

Chaisemartin, Clement de, and Xavier D'Haultfoeuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–96.

Chaisemartin, Clément de, and Xavier D'Haultfoeuille. 2022. "Difference-in-Differences Estimators of Intertemporal Treatment Effects."

Chaisemartin, Clément de, Xavier D'Haultfoeuille, and Yannick Guyonvarch. 2019. "DID_MULTIPLEGT: Stata Module to Estimate Sharp Difference-in-Difference Designs with Multiple Groups and Periods." Statistical Software Components S458643, Boston College Department of Economics.

Chan, Nathan W., and Kenneth Gillingham. 2015. "The Microeconomic Theory of the Rebound Effect and Its Welfare Implications." *Journal of the Association of Environmental and Resource Economists* 2 (1): 133–59.

Chung, Rae Kwon. 2007. "A CER Discounting Scheme Could Save Climate Change Regime after 2012." *Climate Policy* 7 (2): 171–76.

Claassen, Roger, Eric N. Duquette, and David J. Smith. 2018. "Additionality in U.S. Agricultural Conservation Programs." *Land Economics* 94 (1): 19–35.

Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100 (1): 35–69.

Crowther, T. W., K. E.O. Todd-Brown, C. W. Rowe, W. R. Wieder, J. C. Carey, M. B. MacHmuller, B. L. Snoek, et al. 2016. "Quantifying Global Soil Carbon Losses in Response to Warming." *Nature* 540 (7631): 104–8.

Cunningham, Scott. 2021. *Causal Inference*. *The Mixtape*. Yale University Press.

Delgado-Baquerizo, Manuel, Senani B. Karunaratne, Pankaj Trivedi, and Brajesh K. Singh. 2018. "Climate, Geography, and Soil Abiotic Properties as Modulators of Soil Carbon Storage." In *Soil Carbon Storage: Modulators, Mechanisms and Modeling*, 137–65. Elsevier Inc.

Droppers, Bram, Iwan Supit, Michelle T.H. Van Vliet, and Fulco Ludwig. 2021. "Worldwide Water Constraints on Attainable Irrigated Production for Major Crops." *Environmental Research Letters* 16 (5).

Efron, Bradley, and Trevor Hastie. 2016. *Computer Age Statistical Inference*. Cambridge University Press.

Environmental Defense Fund. 2007. "CDM and the Post-2012 Framework." Vienna.

EPA. 2020. "Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2020."

Feder, G., R. E. Just, and D. Zilberman. 1985. "Adoption of Agricultural Innovations in Developing Countries: A Survey." *Economic Development & Cultural Change* 33 (2): 255–98.

Gillingham, Kenneth, Matthew J. Kotchen, David S. Rapson, and Gernot Wagner. 2013. "The Rebound Effect Effect Is Ovrplayed." *Nature, London* 493 (7433): 475–76.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77.

Graham, Bryan S., Cristine Campos De Xavier Pinto, and Daniel Egel. 2012. "Inverse Probability Tilting for Moment Condition Models with Missing Data." *Review of Economic Studies* 79 (3): 1053–79.

Greening, Lorna A., David L. Greene, and Carmen Difigli. 2000. "Energy Efficiency and Consumption — the Rebound Effect — a Survey." *Energy Policy* 28 (6–7): 389–401.

Griliches, Zvi. 1957. "Hybrid Corn : An Exploration in the Economics of Technological Change." *Econometrica* 25 (4): 501–22.

———. 1958. "Research Costs and Social Returns: Hybrid Corn and Related Innovations." *Journal of Political Economy* 66 (5): 419–31.

Hanley, Nick, Simanti Banerjee, Gareth D. Lennox, and Paul R. Armsworth. 2012. "How Should We Incentivize Private Landowners to 'produce' More Biodiversity?" *Oxford Review of Economic Policy* 28 (1): 93–113.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64 (4): 605–54.

Hendricks, Nathan P., and Jeffrey M. Peterson. 2012. "Fixed Effects Estimation of the Intensive and Extensive Margins of Irrigation Water Demand." *Journal of Agricultural and Resource Economics* 37 (1): 1–19.

Hrozencik, R. Aaron. 2022. "Trends in U.S. Irrigated Agriculture: Increasing Resilience Under Water Supply Scarcity." *SSRN Electronic Journal*, no. 229.

Huffaker, Ray. 2008. "Conservation Potential of Agricultural Water Conservation Subsidies." *Water Resources Research* 44 (7): 1–8.

Jevons, William Stanley. 1865. *The Coal Question ; an Enquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-Mines* . London: Macmillan.

Khazzoom, J Daniel. 1980. "Economic Implications of Mandated Efficiency in Standards for Household Appliances Published by : International Association for Energy Economics Economic Implications of Mandated Efficiency in Standards for Household Appliances" 1 (4): 21–40.

King, Gary, and Richard Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27 (4): 435–54.

Lal, R. 2004. "Soil Carbon Sequestration Impacts on Global Climate Change and Food Security." *Science* 304 (5677): 1623–27.

Leuven, Edwin, and Barbara Sianesi. 2003. "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing." Statistical Software Components, Boston College Department of Economics.

Li, Haoyang, and Jinhua Zhao. 2018. "Rebound Effects of New Irrigation Technologies: The Role of Water Rights." *American Journal of Agricultural Economics* 100 (3): 786–808.

Liaw, A., and M. Wiener. 2002. "Classification and Regression by RandomForest." *R News* 2 (3): 18–22.

Martin, Derrel, William L Kranz, T. Smith, Suat Irmak, C. Burr, and R. Yoder. 2017. *Center Pivot Irrigation Handbook*. Lincoln, NE: University of Nebraska-Lincoln, Extension.

Mezzatesta, Mariano, David A. Newburn, and Richard T. Woodward. 2013. "Additionality and the Adoption of Farm Conservation Practices." *Land Economics* 89 (4): 722–42.

Montero, Juan Pablo. 2000. "Optimal Design of a Phase-in Emissions Trading Program." *Journal of Public Economics* 75 (2000): 273–91.

NRCS. 1997. "National Engineering Handbook Irrigation Guide. Technical Report." *Natural Resources Conservation Service, United States Department of Agriculture.*

Ogle, Stephen M., F. Jay Breidt, and Keith Paustian. 2005. "Agricultural Management Impacts on Soil Organic Carbon Storage under Moist and Dry Climatic Conditions of Temperate and Tropical Regions." *Biogeochemistry* 72 (1): 87–121.

Ogle, Stephen M., F. Jay Breidt, Mark Easter, Steve Williams, Kendrick Killian, and Keith Paustian. 2010. "Scale and Uncertainty in Modeled Soil Organic Carbon Stock Changes for US Croplands Using a Process-Based Model." *Global Change Biology* 16 (2): 810–22.

Paul, Carsten, Anja Kristina Techen, James Scott Robinson, and Katharina Helming. 2019. "Rebound Effects in Agricultural Land and Soil Management: Review and Analytical Framework." *Journal of Cleaner Production* 227: 1054–67.

Paustian, K. 2014. "Soil: Carbon Sequestration in Agricultural Systems." In , 5:140–52. Elsevier Inc.

Paustian, Keith, Mark Easter, Kevin Brown, Adam Chambers, Marlen Eve, Adriane Huber, Ernie Marx, et al. 2017. "Field- and Farm-scale Assessment of Soil Greenhouse Gas Mitigation Using COMET-Farm." In , 341–59. Madison, WI, USA: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.

Pérez-Blanco, C. Dionisio, Arthur Hrast-Essenfelder, and Chris Perry. 2020. "Irrigation Technology and Water Conservation: A Review of the Theory and Evidence." *Review of Environmental Economics and Policy* 14 (2): 216–39.

Perry, Chris, and Pasquale Steduto. 2017. *Does Improved Irrigation Technology Save Water? A Review of the Evidence*. *Food and Agriculture Organization of the United Nations, Cairo*. Vol. 42.

Peterson, Jeffrey M., and Ya Ding. 2005. "Economic Adjustments to Groundwater Depletion in the High Plains: Do Water-Saving Irrigation Systems Save Water?" *American Journal of Agricultural Economics* 87 (February): 147–59.

Pfeiffer, Lisa, and C. Y.Cynthia Lin. 2014. "Does Efficient Irrigation Technology Lead to Reduced Groundwater Extraction? Empirical Evidence." *Journal of Environmental Economics and Management* 67 (2): 189–208.

Plastina, Alejandro. 2021. "How Do Data and Payments Flow through Ag Carbon Programs ?" *Ag Decision Maker. Iowa State University Extension and Outreach*.

Plastina, Alejandro, and Oranuch Wongpiyabovorn. 2021. "How to Grow and Sell Carbon Credits in US Agriculture." *Ag Decision Maker. Iowa State University Extension and Outreach*.

PRISM Climate Group. 2014a. "PRISM Daily Data." Oregon State University. 2014. https://prism.oregonstate.edu.

———. 2014b. "PRISM Monthly Data." Oregon State University. 2014. https://prism.oregonstate.edu.

Richards, Kenneth R, and Robert N Stavins. 2005. “The Cost of U.S. Forest-Based Carbon Sequestration.” Arlington, Virginia.

Rogers, Danny H, Jonathan Aguilar, Isaya Kisekka, and Freddie R Lamm. 2017. “Center Pivot Irrigation System Losses and Efficiency.” *Proceedings of the 29th Annual Central Plains Irrigation Conference* Kansas Sta (Feb. 21-22).

Rosa, Lorenzo, Maria Cristina Rulli, Kyle Frankel Davis, Davide Danilo Chiarelli, Corrado Passera, and Paolo D’Odorico. 2018. “Closing the Yield Gap While Ensuring Water Sustainability.” *Environmental Research Letters* 13 (10).

Rosenberg, Andrew B. 2020. “Targeting of Water Rights Retirement Programs: Evidence from Kansas.” *American Journal of Agricultural Economics* 102 (5): 1425–47.

Sandmo, Agnar. 1975. “Optimal Taxation in the Presence of Externalities.” *The Swedish Journal of Economics* 77 (1): 86–98.

Sant’Anna, Pedro H.C., and Jun Zhao. 2020. “Doubly Robust Difference-in-Differences Estimators.” *Journal of Econometrics* 219 (1): 101–22.

Saunders, Harry D. 1992. “The Khazzoom-Brookes Postulate and Neoclassical Growth Author (s): Harry D . Saunders Published by : International Association for Energy Economics Stable URL : Https://Www.Jstor.Org/Stable/41322471” 13 (4): 131–48.

Sawadgo, Wendiam, and Alejandro Plastina. 2021. “Do Cost-Share Programs Increase Cover Crop Use? Empirical Evidence from Iowa.” *Renewable Agriculture and Food Systems* 1 (2016): 1–9.

Schlenker, Wolfram, and Michael J. Roberts. 2009. “Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (37): 15594–98.

Schneider, Lambert. 2009. “Assessing the Additionality of CDM Projects: Practical Experiences and Lessons Learned.” *Climate Policy* 9 (3): 242–54.

Sims, Z. R., and G. A. Nielsen. 1986. “Organic Carbon in Montana Soils as Related to Clay Content and Climate.” *Soil Science Society of America Journal* 50: 1269–71.

Słoczynski, Tymon. 2020. “Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights.” *Review of Economics and Statistics*.

Soil Survey Staff. 2022. “Soil Survey Geographic (SSURGO) Database.” Natural Resources Conservation Service, United States Department of Agriculture. 2022. https://sdmdataaccess.sc.egov.usda.gov.

Sorrell, Steve, John Dimitropoulos, and Matt Sommerville. 2009. “Empirical Estimates of the Direct Rebound Effect : A Review” 37: 1356–71.

Sperow, Mark. 2016. “Estimating Carbon Sequestration Potential on U.S. Agricultural Topsoils.” *Soil and Tillage Research* 155: 390–400.

Storm, Hugo, Kathy Baylis, and Thomas Heckelei. 2020. “Machine Learning in Agricultural and Applied Economics.” *European Review of Agricultural Economics* 47 (3): 849–92.

Sun, Liyang, and Sarah Abraham. 2021. “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects.” *Journal of Econometrics* 225 (2): 175–99.

Sunding, David, and David B T - Handbook of Agricultural Economics Zilberman. 2001. “Chapter 4 The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector.” In *Agricultural Production*, 1:207–61. Elsevier.

Swan, Amy, Mark Easter, Adam Chambers, Kevin Brown, Stephen Williams, Jeff Creque, John Wick, and Keith Paustian. 2020. “COMET-Planner Carbon and Greenhouse Gas Evaluation for NRCS Conservation Practice Planning.”

United States Department of Agriculture. 2020. “2018 Irrigation and Water Management Survey.” *2017 Census of Agriculture*. Vol. 3.

USDA National Agricultural Statistics Service. 2022. “Published Crop-Specific Data Layer [Online].” Cropland Data Layer. 2022. https://nassgeodata.gmu.edu/CropScape/.

Ward, Frank A., and Manuel Pulido-Velazquez. 2008. “Water Conservation in Irrigation Can Increase Water Use.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (47): 18215–20.

Weil, Raymond, and Nyle Brady. 2017. *The Nature and Properties of Soils. 15th Edition*.

West, Tristram O., and Wilfred M. Post. 2002. “Soil Organic Carbon Sequestration Rates by Tillage and Crop Rotation: A Global Data Analysis.” *Soil Science Society of America Journal* 66 (6): 1930–46.

White, Ben, and Nick Hanley. 2016. “Should We Pay for Ecosystem Service Outputs, Inputs or Both?” *Environmental and Resource Economics* 63 (4): 765–87.

Wilson, B, J Bartley, K Emmons, J Bagley, J Wason, and S Stankiewicz. 2005. “Water Information Management and Analysis System, Version 5, for the Web. User Manual.” *Kansas Geological Survey Open File Report 2005-30.*, 37.

Wilson, Janet. 2022. “Senators Add $4 Billion for Colorado River Drought Relief into Inflation Reduction Act.” *The Desert Sun*, August 7, 2022.

Wuepper, David, and Robert Huber. 2021. “Comparing Effectiveness and Return on Investment of Action- and Results-Based Agri-Environmental Payments in Switzerland.” *American Journal of Agricultural Economics*, no. January: 1585–1604.

Appendix A - Supplemental materials for Chapter 2

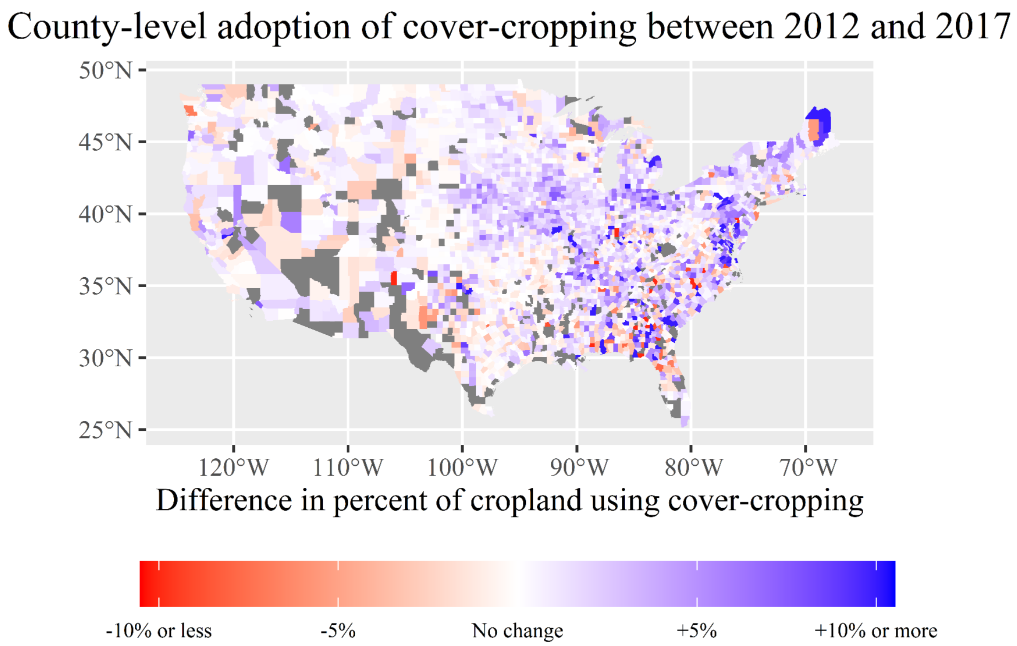

Figure A.1: Change in share of cropland acreage employing cover cropping between 2012 and 2017.

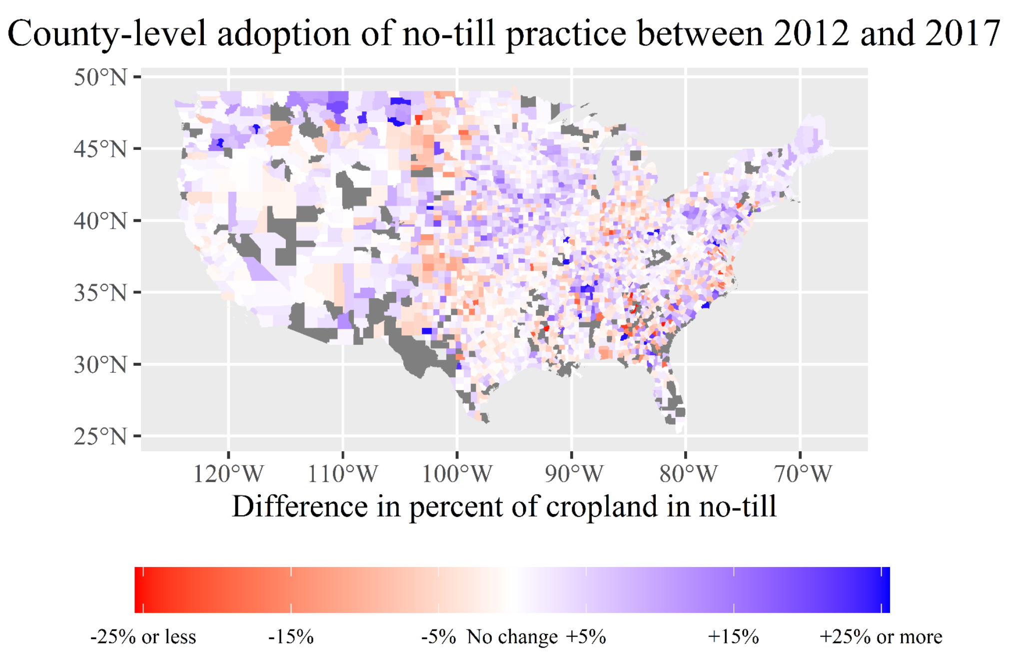

Figure A.2: Change in share of cropland acreage using no-till between 2012 and 2017.

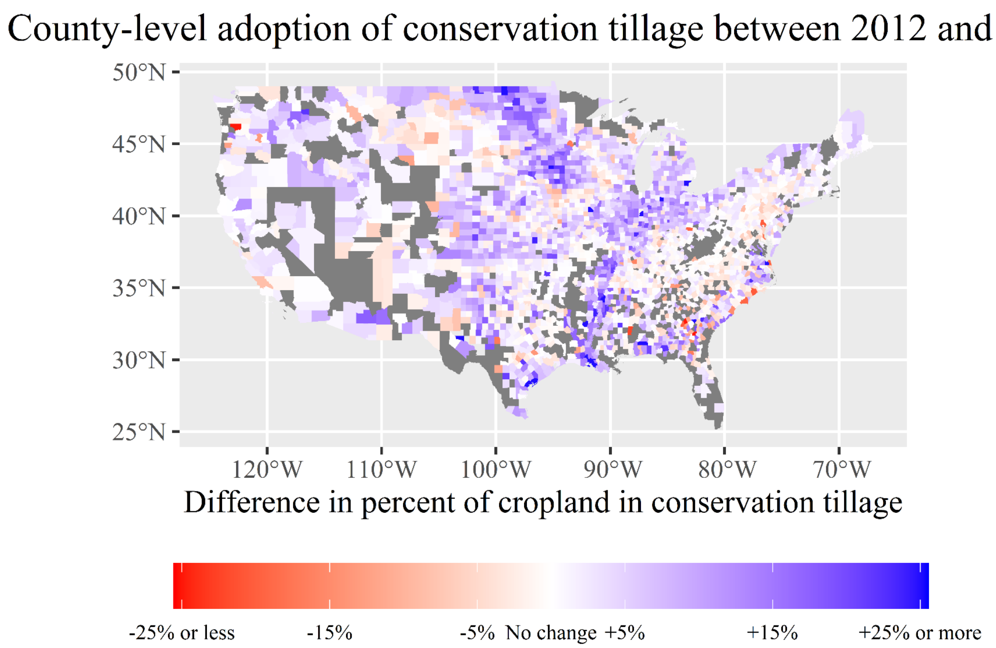

Figure A.3: Change in share of cropland acreage using conservation tillage between 2012 and 2017.

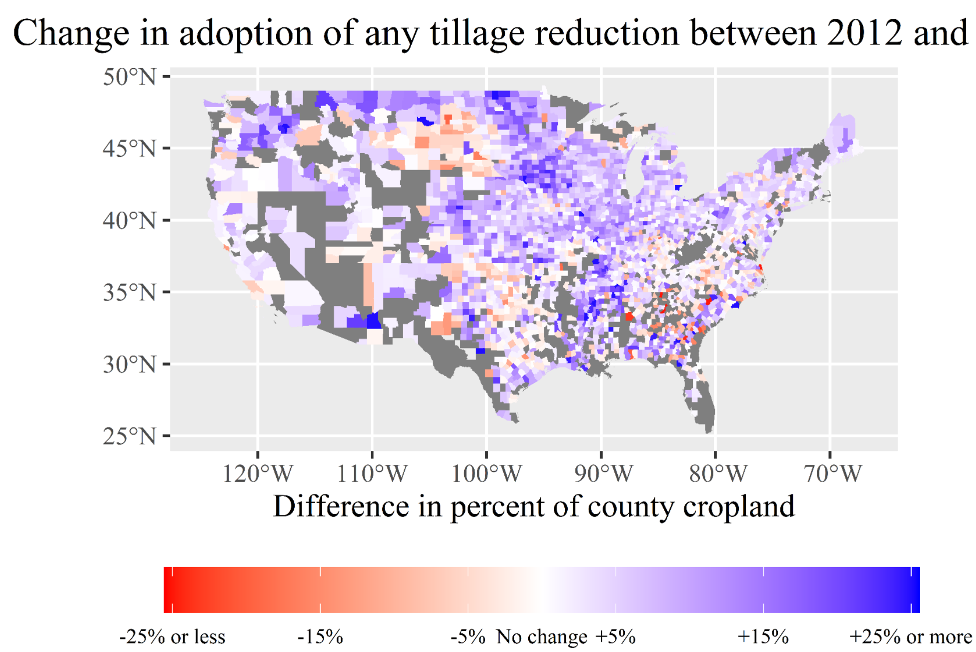

Figure A.4: Change in share of cropland acreage using either no-till or conservation tillage between 2012 and 2017.

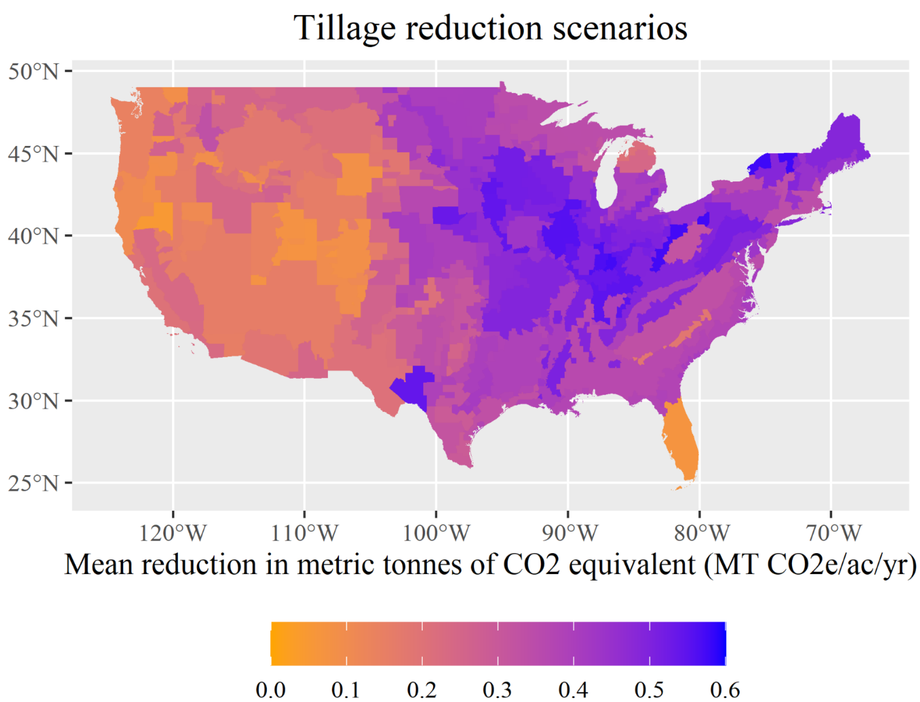

Figure A.5: Average carbon sequestration for COMET Planner scenarios involving a reduction in tillage intensity (CPS numbers 329 and 345).

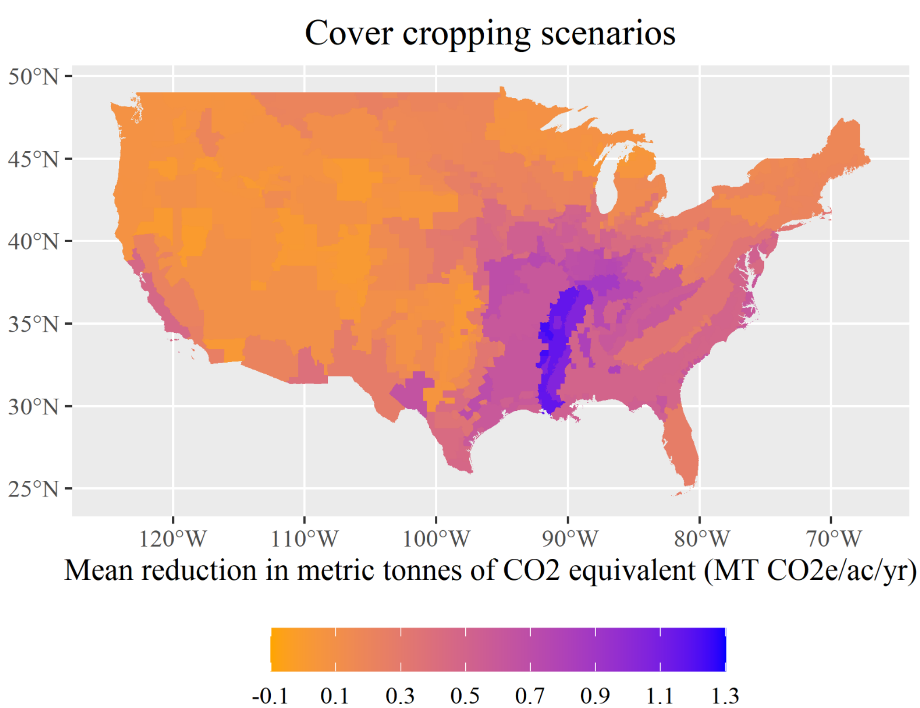

Figure A.6: Average carbon sequestration for COMET Planner scenarios involving cover-cropping (CPS number 340).

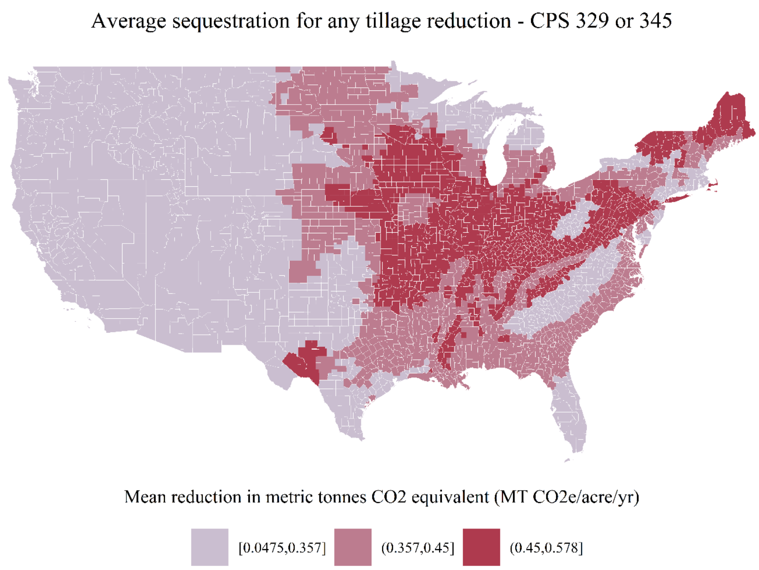

Figure A.7: Average carbon sequestration due to tillage reduction practices by tercile.

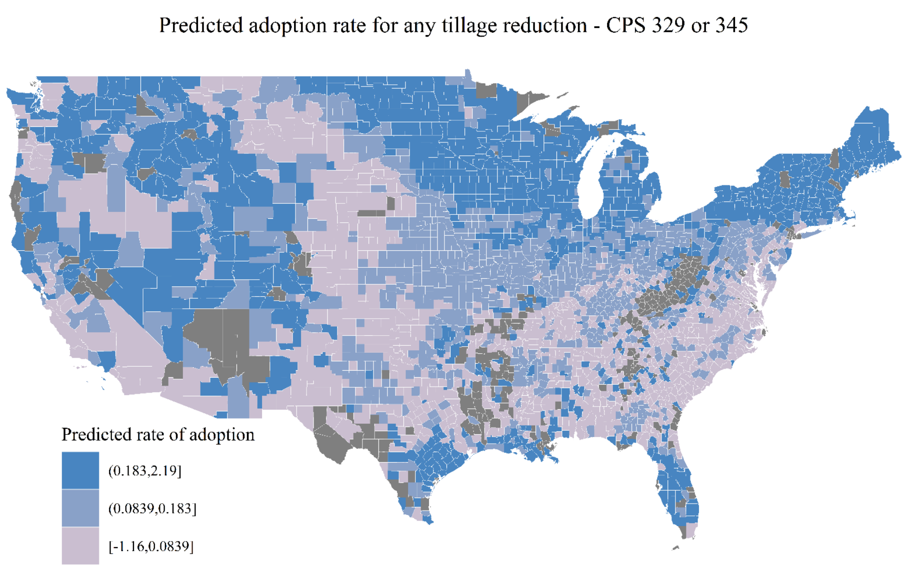

Figure A.8: Predicted rate of adoption for tillage reduction practices by tercile.