A Clean SWEEP? Estimating the Impacts of Irrigation Efficiency Subsidy Programs on Water and Energy Use

Isabelle Picciotto

February 9, 2025

Link to Current Version

Abstract

State and federal governments spend billions of dollars on agricultural conservation programs that provide subsidies to agricultural producers for on-farm actions to improve environmental outcome. Evaluating these programs is econometrically complex due to the voluntary nature of participation and the "black box" evaluation of applicants. These challenges make it difficult to evaluate the realized effects on natural resources. Programs may have unintended consequences. Behavioral responses to efficiency improvements, for example, may cause a rebound effect in resource use, eroding potential benefits. I use granular applicant-level data for California's State Water Efficiency and Enhancement Program (SWEEP) to account for program selection combined with large volume field-level geospatial data to estimate whether receiving irrigation efficiency subsidies result in adjustments to production decisions that may cause a rebound effect. Across the entire set of applicants, on average, I find that program enrollment has no significant effect on crop choice or water use. However, I find important heterogeneities in program impacts. Having access to surface water leads to funded applicants switching to more water-intensive crops compared to their unfunded counterparts. Funded projects that install micro irrigation technologies are effective in reducing consumptive water use however these effects are entirely offset by those projects that do not install micro irrigation technologies and, in fact, increase their consumptive water use. The purpose of SWEEP is to reduce water use and energy use. These results highlight key issues related to conservation program targeting. SWEEP funds a broad range of potential technologies, some of which reduce water use, but ultimately funding decisions are not made along the margins that most significantly increase water conservation.

1 Introduction

Water conservation in agriculture is becoming increasingly important. Global climate change is increasing the occurrence and severity of extreme weather events, such as heat waves and droughts. These events have a significant effect on surface and groundwater stores in the western United States. Recent droughts have increased surface water variability and scarcity, causing agricultural producers to rely more heavily on groundwater for irrigation. In California, agricultural irrigation can account for more than two-thirds of water use in a dry year (Mount and Hanak, 2016; Newman, Howitt, and MacEwan, 2018), and unsustainable groundwater extraction and droughts are two of the foremost issues affecting California's agricultural sustainability. Increased reliance on groundwater has led to an accelerating rate of groundwater depletion in California's Central Valley since 2003 (Liu et al., 2022).

In response to critical groundwater overdraft and increasingly scarce surface water supplies, California has introduced a number of water management policies. In 2014, California state legislature passed the Sustainable Groundwater Management Act (SGMA) to bring groundwater basins into balanced levels of pumping and recharge. In the same year, the California State Water Efficiency and Enhancement Program (SWEEP) was authorized under an emergency drought declaration. SWEEP subsizides adoption of more efficient irrigation technology on agricultural operations in California with the joint goals of saving water and reducing energy use. Since its inception in 2014, SWEEP has awarded over 1,000 projects in California, shelling out more than $123 million in funding. To put in perspective the magnitude of these types of conservation programs on a national scale, the USDA Environmental Quality Incentives Program (EQIP), which subsidizes, among other things, similar irrigation projects, received over $2 billion in funding in fiscal year 2023 just for irrigation related projects.

Due to the complex management of water and the lack of established water markets in many areas, groundwater extraction suffers from the tragedy of the commons; groundwater extraction by one agricultural producer necessarily restricts the availability of that resource for other uses. When property rights are not well-defined over a common resource, economic theory suggests that without policy intervention, producers will over-extract groundwater and under-invest in irrigation efficiency. Effective groundwater management is lacking in many agricultural regions, which exacerbates existing issues of groundwater overdraft. Government conservation programs, such as SWEEP, provide subsidies to promote adoption of more efficient technologies to close the gap between the social optimum and market equilibrium. While these programs have specific objectives, such as reducing water use, increasing carbon sequestration, or improving soil health, an open question in the literature is whether these programs are meeting their objectives and achieving conservation goals.

In this paper I use SWEEP to evaluate the effect of irrigation efficiency subsidies on crop production decisions with a focus on better understanding the resource rebound effect in agriculture and evaluating the effectiveness of these types of conservation programs. A rebound effect refers to when behavioral responses to efficiency improvements erode the potential benefits of increased efficiency. It was first documented by Jevons (1865) regarding the demand for energy after increased efficiency in coal consumption and has subsequently been heavily studied in the energy literature with respect to, for example, fuel efficient cars or energy saving light bulbs (For a review, see: Gillingham, Rapson, and Wagner, 2016; Greening, Greene, and Difiglio, 2000).

Since these type of voluntary conservation programs are one of the main ways that state and federal promote adoption of conservation practices this question has important policy implications. Evaluating these programs is econometrically complex due to the voluntary nature of participation and the uncertainty around applicant evaluation criteria by the researcher. I address these challenges by using a combination of granular applicant-level program data and field-level remotely sensed data on land use and water use to evaluate the effect of SWEEP on program outcomes. My research questions are: (1) Does participation in SWEEP change producer crop choice, and subsequently, crop water demand? (2) Does participation in SWEEP change producer water use and energy use? and (3) Do these effects vary depending on groundwater regulatory strength, the source of irrigation water, or technology?

Studies using simulations or mathematical programming methods find mixed evidence across counterfactual scenarios for irrigation efficiency improvements on water savings, estimates range from modest water savings to more than a 100% rebound in water use (Christian-Smith, Cooley, and Gleick, 2012; Medellín-Azuara, Howitt, and Harou, 2012; Berbel and Mateos, 2014; Freire-González, 2019). Two related empirical studies that use panel data on groundwater extraction and adoption of low energy precise application irrigation technology in Kansas find evidence that adoption leads to an increase in groundwater extraction along both the extensive and intensive margins (Pfeiffer and Lin, 2014; Li and Zhao, 2018). Pfeiffer and Lin (2014) indirectly discuss the role that subsidy programs play in adoption, since during their study period there was extensive take-up of EQIP for irrigation-related investments.

However, a recent study by Cameron-Harp and Hendricks (2024) find that these effects may be biased due to the use of two-way fixed effects estimation with heterogeneous treatment effects. Wallander and Hand (2011) specifically evaluate EQIP and find that adoption of irrigation practices through EQIP may incentivize an expansion in irrigated acreage and increase in water use, but these results became insignificant when controlling for self-selection into the program using matching methods.

These studies do not paint a consistent picture of how irrigation efficiency improvements interact with production decisions, and they illustrate the lack of policy evaluation of irrigation subsidy programs. These evaluations are important for future policy-making as funding through agricultural conservation programs is abundant; the Inflation Reduction Act directed an additional \$11.7 billion in 2023 to support USDA's working lands conservation programs over the next four years.

These types of agricultural subsidies are politically salient and it is likely the trend of increasing these subsidies to promote adoption of climate smart agriculture and conservation practices will continue. However, whether or not these programs deliver in terms of capturing all of the potential environmental benefits is still an open question. Adopting new management practices or production technologies change the profit maximization problem, which may in turn affect production decisions that increase or decrease the conservation savings from a particular practice. This could be due to changes in input use or land use decisions. Irrigation efficiency improvements can reduce labor costs and water costs, make water-intensive crops more feasible, or make less productive land more feasible for agricultural production. Any adjustments along these margins have the potential to erode the benefits of irrigation efficiency improvements.

Because conservation programs have voluntary enrollment, it is difficult to identify a causal effect of program participation on relevant outcomes. Individual producers apply based on unobservable factors that may not only influence their decision to participate, but may also influence their production choices before and after enrollment. I leverage two important characteristics of SWEEP to address these challenges. First, I obtained granular data at the SWEEP applicant level for both funded and unfunded projects. This data contains the application score, relevant score criteria, and additional demographic information, providing insight into program selection criteria and similarity of applicants. Second, because of state government funding limitations, SWEEP is consistently oversubscribed, with more eligible applicants than the program can fund in each year. I incorporate these characteristics into my identification strategy and use a nonparametric generalization of the difference-in-differences

(DiD) estimator proposed by Imai, Kim, and Wang (2021). This estimator relies on constructing a set of control units for each treated unit using existing matching methods and then applies the DiD estimator to each treated observation using the weighted average of matched control units. The treatment effect is the weighted average across all treated units. Treated units are defined as applicants to SWEEP that received funding and control units are defined as applicants to SWEEP that did not receive funding.

One of the characteristics of SWEEP is that applicants are able to install a variety of technologies, such as soil moisture sensors, variable frequency drive systems for irrigation, or solar-powered irrigation systems. I create a standardized categorization of SWEEP projects based on the types of technologies that are installed on each project based on project descriptions. Using natural language processing, I systematically categorize the technologies that applicants install through SWEEP into similar groupings to shed light on the most common practices funded through the program, which practices are generally installed together, and whether program effects vary by project type.

Using field-level crop choice data, county-level estimates of crop water demand (intensity), and field-level consumptive water use, I evaluate the effect of SWEEP enrollment on three outcomes: (1) The proportion of total acres planted to a water-intensive crop, (2) the average crop water intensity across all acreage, and (3) average consumptive water use. The first two outcomes inform whether program enrollment causes a shift in cropping decisions. I index field-level crop choice using estimates of seasonal irrigation water requirements by crop group and County. Indexing crop choice allows me to compare relative crop water-intensities and measure how crop switching may impact water use. I define a binary variable indicating whether a crop is water-intensive if it has an estimated seasonal irrigation water requirement greater than the County median. If SWEEP participants plant more (or less) acreage in water intensive crops following enrollment, the treatment effect will be positive (negative).

The third variable informs changes in actual water use. I use remotely sensed data on evapotranspiration (ET)and precipitation to estimate field-level consumptive water use.[1] Consumptive water use is the portion of water that is completely removed from the system, i.e., is not recoverable or reusable downstream. ET is the second largest component of the water cycle after precipitation and makes up the majority of consumptive water use in irrigated agriculture (Wilkening et al., 2021). I use ET and

---

[1] Evapotranspiration is water that evaporates from land and water sources or transpires from plants and re-enters the atmosphere.

precipitation data at the field level to estimate changes in actual consumptive water use. If irrigation efficiency improvements reduce non-beneficial transpiration from weeds or soil surfaces or improve irrigation scheduling to better meet crop water requirements, the treatment effect will be negative.

Across the entire set of applicants applicants, on average, I find no significant effect of program enrollment on crop choice or water use. In other words, I do not find evidence of a rebound effect through the channels of producers switching to more water-intensive crops or increasing their consumptive water use. However, I find distinct heterogeneities in how subsidies for efficient irrigation technologies affect crop choice and water use. For applicants that have access to surface water I find evidence of a rebound effect, demonstrated by funded applicants switching to more water-intensive crops. Conversely, I find precisely estimated null effects for applicants that report using only groundwater for irrigation. Intuitively, this reflects the incentive for producers to use their full endowment of surface water due to California's "beneficial use" water rights system.

When considering different irrigation technologies, I find that funded projects that install micro irrigation technologies are effective in reducing their consumptive water use relative to unfunded applicants. However, these water savings are almost entirely offset by funded projects that do not install micro irrigation technologies and increase their consumptive water use. The main goal of SWEEP is to reduce water use and energy use on agricultural operations, however program impacts vary significantly across different types of applicants and different project components such that the average effect of the program is zero. These results highlight some of the key issues that arise from a lack of targeting in conservation subside programs. These programs subsidize a broad range of technologies, some of which are effective at reducing water use, however if funding decisions are not made along margins that reduce water use then these programs are unlikely to meet their conservation goals.

This paper contributes to the literature on agriculture conservation program evaluation by causally estimating the effect of an agricultural subsidy program on production choices, specifically in response to improved irrigation efficiency. I use heterogeneity robust econometric methods to estimate the effect of irrigation efficiency subsidy programs on crop choice and water use, controlling for variation in treatment timing and individual-level characteristics. Subsidizing conservation practices is one of the main policy tools that is used by state and federal governments to promote adoption. These programs are expensive to implement and econometrically complex to evaluate. I use remotely sensed data to estimate consumptive water use, which advances our ability to evaluate these types of programs on resource use. Previous studies have evaluated how farm characteristics, such as farm size, education, and land quality influence the adoption of conservation practices (Schaible et al., 2015; Fuglie and Kascak, 2001). Others have evaluated how exposure to risk such as weather shocks, increased drought, and water scarcity function as incentive mechanisms to adopt more efficient irrigation (Shi, Wu, and Olen, 2022; Blumberg, Goemans, and Manning, 2022; Hagerty, 2021; Olen, Wu, and Langpap, 2016; Wallander et al., 2013).

This paper contributes to the literature on the rebound effect in natural resource efficiency-enhancing technology. I evaluate a larger selection of behavioral mechanisms that may cause a rebound effect in water use in agriculture to better understand the micro-level impacts of irrigation incentive programs. A large body of literature on the rebound effect focuses on energy efficiency (e.g., Gillingham, Rapson, and Wagner, 2016; Sorrell and Dimitropoulos, 2008; Brännlund, Ghalwash, and Nordström, 2007; Binswanger, 2001; Greening, Greene, and Difiglio, 2000) and other resource efficiency issues, including agricultural land use and groundwater extraction (Cameron-Harp and Hendricks, 2024; Li and Zhao, 2018; Pfeiffer and Lin, 2014) or modeling water demand and economy-wide impacts of irrigation efficiency improvements (Freire-González, 2019; Gómez and Pérez-Blanco, 2014). This paper further contributes to this literature by empirically evaluating an irrigation efficiency subsidy program in California ex post to estimate the relationship between subsidy programs on crop choice, water use, and energy use.

Finally, the effectiveness of large-scale agro-environmental subsidy programs has been widely debated in the literature in both the domestic and international contexts (Börner et al., 2017). This paper contributes to that debate by evaluating the behavioral response to SWEEP participation and its effectiveness in reducing water use and facilitating climate change adaptation in agriculture. Previous studies have quantified the environmental benefits of certain conservation programs, specifically related to water quality (Karwowski and Skidmore, 2023; Liu, Wang, and Zhang, 2022; Skidmore, Andarge, and Foltz, 2023). Whether these programs capture all potential environmental benefits and their effectiveness as climate change adaptation strategies more broadly, is an open question.

2 Background of California's State Water Efficiency and Enhancement Program

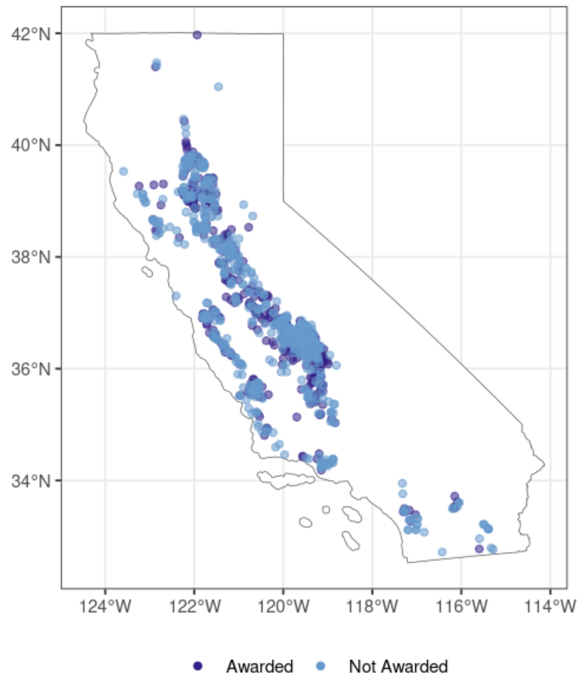
Since 2003, total irrigated acres in California have remained stable around 8.4 million acres. However, the type of irrigation technology used has shifted. Irrigated acreage transitioned from gravity irrigation, which accounted for about 63% of irrigated acres in 2003 and 46% of irrigated acres in 2018, to sprinkler irrigation systems and drip, trickle, or low-flow micro sprinklers, which accounted for about 41% of irrigated acres in 2003 and 60% of irrigated acres in 2018 (USDA, 2013, 2018).[2] More efficient irrigation systems improve the precision with which water is applied to target crops, reducing non-beneficial evaporation and runoff, and improving beneficial transpiration, which has been shown to increase yields (Evans and Sadler, 2008). In California, the majority of crop revenue originates from irrigated agriculture, and the composition of irrigated cropland has been shifting from field crops towards perennials over the same time period. Perennials limit producers' ability to adjust to water availability during times of drought, since these fields cannot be fallowed or switched to less water-intensive crops easily.

SWEEP is an agricultural subsidy program administered by the California Department of Food and Agriculture (CDFA) and was authorized in 2014 under an emergency drought declaration. Funding for the program comes from the state's Greenhouse Gas Reduction Fund in 2014, 2015 and 2016, the California Drought, Water, Parks, Climate, Coastal Protection, and Outdoor Access for All Act of 2018 in 2018 and 2019, and the Budget Act of 2021 in 2021. Participating agricultural producers are required to improve efficiency of their irrigation management systems to address two goals: save water and reduce greenhouse gas emissions. To be eligible, projects must demonstrate, through ex ante calculations, that the proposed changes have the potential to reduce both water and energy use. Eligible project components include technologies such as soil moisture monitoring, the use of weather or ET data for irrigation scheduling, drip or micro-irrigation systems, renewable energy, retrofitting or replacing pumps, low pressure irrigation systems, and variable frequency drives. Previous literature has estimated how these different components may reduce agricultural water use. Using California's Department of Water Resources Analytical model, Christian-Smith, Cooley, and Gleick (2012) estimate

---

[2] Source: USDA NASS Irrigation and Water Management Survey, formerly called the Farm and Ranch Irrigation Survey. Multiple systems covering the same acreage are counted each time they cover the acres. This includes multiple systems of the same type.

Figure 1: Map of SWEEP Applicants by Treatment Status



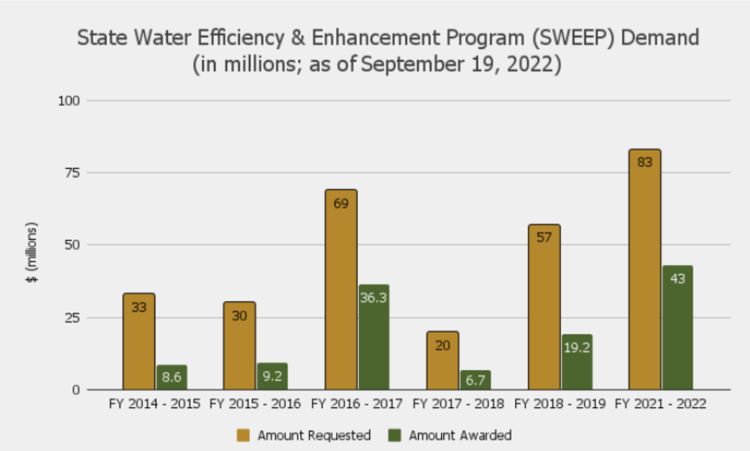
This figure presents a map of SWEEP applicants in California. Each dot represents a single application with dark blue representing applicants that were awarded funding and light blue applicants representing applicants that were not awarded funding.

agricultural water use by hydrologic region. They find that irrigation scheduling may reduce water use by 10% and switching from flood irrigation to more efficient irrigation systems may reduce water use by 3.5%.

Since 2014, SWEEP has awarded 1,111 projects, and more than $123 million in funding has been obligated with over $70 million in matching funds. CDFA estimates that the program generates benefits that reduces greenhouse gas emissions equivalent to taking 203,000 cars off the road per year and water savings totaling 144,545 acre-feet per year. In an average year, 9.6 million acres in California are irrigated using approximately 34 million acre-feet of water. These estimated water savings on less than 2% of California's irrigated cropland equate to a little less than half of a percentage decrease per year in total water used for irrigated agriculture in California.

Figure 1 present the location of SWEEP applicants by approval status. In general, applicants are concentrated in the Central Valley, consistent with the distribution of agricultural production in California. Since 2014, there have been nine Requests for Proposals (RFP) with requested funds

Figure 2: SWEEP Demand for Funding in Millions of Dollars from 2014 through 2022



This figure illustrates for each fiscal year, how much funding was requested by applicants and how much funding was subsequently awarded to approved applicants. A portion of the year-over-year variation has to do with changes in the maximum dispersal amount for the program as well as how many applicants the program received.
Source: `https://calclimateag.org/sweep/`

outpacing available funds each round.[3] Figure 2 illustrates, for each fiscal year, the requested funds and awarded funds. The annual variation in funds requested and awarded can be partially attributed to changes in the maximum dispersal amount as well as the total number of program applicants in that fiscal year. In general, eligible applicants are funded until the available funds run out for that RFP round.

Due to data limitations in the 2014 and 2015 funding rounds, I use applicant data from funding rounds 2016 through 2021, excluding 2020.[4] The RFP for each funding round outlines the eligibility requirements, funding information, approved practices and projects, and a description of the application process. To be eligible for funds, the project must be applied to an irrigation project on a commercial California agricultural operation. An operation is defined as row, vineyard, field and tree crops, commercial nurseries, nursery stock production, and greenhouse operations producing food crops or flowers. The applicant is responsible for providing documentation related to actual, on-farm water consumption and GHG emissions from the prior growing season.

Table 1 presents the total number of applications to SWEEP by funding year and application status. The same agricultural operation can apply for additional funding in future application years if they already have applied, however the new project cannot build upon a previous SWEEP project that

---

[3] Source: `https://calclimateag.org/sweep/`

[4] In 2020, there was no RFP due to the COVID-19 pandemic.

Table 1: Number of Applications by Funding Year and Application Status

Funding Year	Awarded	Not Awarded	Declined	Disqualified	Total
2016*	284	125	17	140	566
2017	81	115	4	33	233
2018	106	200	14	23	343
2019	120	207	5	34	366
2021	70	247	0	15	332

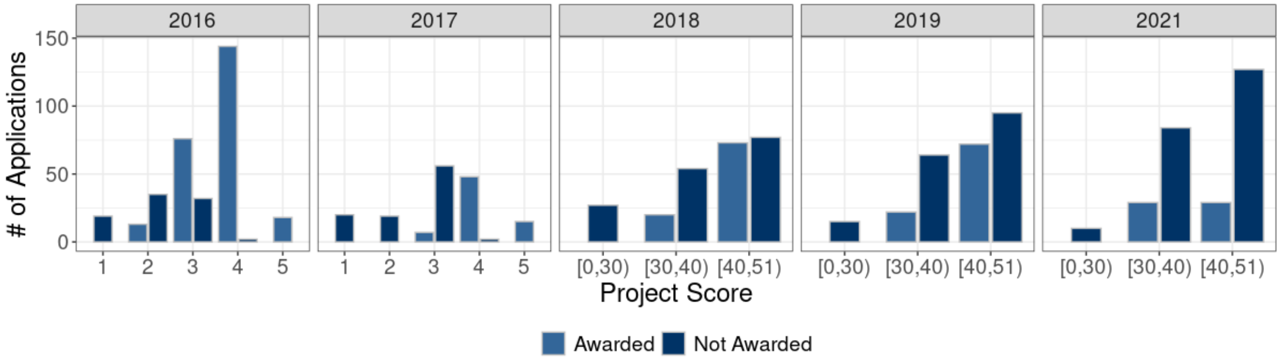
This table presents the total number of applications to SWEEP by funding year and application status. Awarded applicants received funding in that year, Not Awarded applicants did not receive funding in that year, Disqualified applicants were considered ineligible for ranking, and Declined applicants were offered funding but declined.

*There were two funding rounds conducted in 2016. Round I was awarded in March of 2016 and Round II was awarded in November of 2016.

affected the same assessor parcel number(s) (APN). If an applicant applied and was not awarded but re-applied and was awarded in a subsequent year, I remove them from the control group in their first year of application to account for potentially multiple instances of the same fields across treatment and control observations. SWEEP funding cannot be combined with USDA NRCS EQIP financial assistance, meaning that applicants may not accept funding from both entities for the same project components and APNs. Each agricultural operation is also restricted such that the cumulative SWEEP funding received across all approved projects does not exceed $600,000. Funds may not be used to expand existing agricultural operations, install new groundwater wells or increase well depth, test new technology or perform research.

Applications are reviewed and scored by CDFA in a multi-step process. The first step is an administrative review to determine whether application requirements were met and, if applicable, assess an applicant's past grant performance. If an application is disqualified during the first stage review, I exclude it from the potential control group. The second step is a technical review to evaluate the merits of the application and expected success of the project. Technical reviewers are comprised of agricultural irrigation water system specialists and experts affiliated with the University of California and California State University systems. The specific criteria that goes into application scores has changed over time. For funding years 2016 and 2017 applications were scored on an integer scale between zero and five. The RFP did not specify a minimum score cut off, but evaluation of the data indicates that approval was more likely with a score of at least three. An applicant can receive additional considerations which increase the likelihood of being funded if they meet any of the following criteria: completion of or commitment to complete irrigation training, being located within a critical overdrafted groundwater

Figure 3: Application Approval Status by Technical Reviewer Score



This figure presents the total number of applications by project score. Prior to 2018, applications were scored on a one through five scale. Since 2018, projects are scored on a zero to fifty scale and a minimum score of 30 is required for funding eligibility. Scores in 2018, 2019, and 2021 are binned from 0-30, 30-40, and 40-50 for simplicity.

basin, applying soil management practices, matching funds of at least 50%, and being a new SWEEP recipient. In 2016 round I, priority was also given to projects that provided benefits to disadvantaged communities.

In funding rounds 2018, 2019, and 2021, applications were scored on an integer scale between zero and 50. Approval was contingent on a score of 30 points or higher. For these rounds, the technical review includes the following criteria: Merit and Feasibility (12 points), Water Savings and Calculations (12 points), GHG Reductions and Calculations (12 points), Budget (8 points), Applicant Not Previously Awarded (3 points), and Additional Considerations (3 points). Additional considerations include irrigation training, being located within a critically overdrafted basin, and soil management practices. In funding rounds 2018 and 2019 additional considerations also included using recycled water or storm water capture, and including matching funds and/or in-kind contributions.

The distribution of application scores by funding year is presented in Figure 3, which provides a count of the number of applications that were awarded funding or not awarded funding based on their project score. Outcomes are plotted separately by funding year. Table 2 presents the components that determine application score by funding year. Beginning in 2018, priority funding was given to severely disadvantaged communities (DAC) and socially disadvantaged farmers and ranchers (SDFR).[5]

---

[5] A severely disadvantaged community is defined as a community with a median household income less than 60 percent of the statewide average. Socially disadvantaged farmers and ranchers are defined as being part of the following groups by the 2017 Farmer Equity Act (AB1348): African Americans, Native Indians, Alaskan Natives, Hispanics, Asian Americans, Native Hawaiians and Pacific Islanders.

Table 2: Scoring Criteria by Funding Year

Years	Score	Criteria
2016	1-5	Additional considerations[a], matching funds, new recipient
2017	1-5	Additional considerations[a], matching funds, new recipient
2018/19	0-50*	Merit and feasibility (12), water saving (12), GHG saving (12), budget (8), new recipient (3), additional considerations[a] (3), SDAC[c], DAC[b], recycled water/stormwater capture, matching funds
2021	0-50*	Merit and feasibility (12), water saving (12), GHG saving (12), budget (8), new recipient (3), additional considerations[a] (3), SDAC[c], DAC[b], sub-surface drip for applying dairy effluent

a Irrigation training, critically overdrafted basin, soil management practices

b Project serves disadvantaged communities

c Socially disadvantaged farmers and ranchers

∗ Minimum score of 30 to be eligible

In the 2018 and 2019 funding rounds, SWEEP funding was provided by Proposition 68, which requires that CDFA award 20% of the appropriated $20 million to serve severely disadvantaged communities. Qualifying projects received priority funding if they obtain a minimum score of 30 points during the technical review. Similarly, socially disadvantaged farmers and ranchers received priority funding during these years if they obtain a minimum score of 30 points during the technical review.

In the 2021 funding round at least 25% of the appropriated $50 million was reserved for socially disadvantaged farmers and ranchers and priority populations provided they obtained a minimum score of 30 points during the technical review.[6] An additional $2 million was set aside by CDFA in funding round 2021 for projects that used sub-surface drip irrigation to apply dairy effluent to field crops. Table 3 presents the maximum dispersal amount for the program, the maximum grant award, the maximum grant duration, and the final deadline by which the project had to be operational for each funding round. The maximum amount dispersed varied from a maximum of $43 to $45 million in 2021 to a minimum of $4.5 million in 2017. The maximum grant award was either $200,000 or $100,000 depending on the funding year. Over time, the maximum duration of a single grant increased from 12 months in 2016 and 2017 to 18 months in 2018 and 2019, and to 24 months in 2021.

Due to the temporal variation in program aspects such as the scoring criteria, program funding, and total available funds, there is concern with using standard methods of program evaluation such as difference-in-differences (DiD) and two-way fixed effects (TWFE). The applicants across funding

---

[6] Priority populations include disadvantaged communities, low-income communities, and low-income households. Projects which benefit priority populations must be located within an area designated as a priority population on the following map: `https://webmaps.arb.ca.gov/PriorityPopulations/`

Table 3: SWEEP Program Summary by Funding Round

Funding Round	Max Dispersal	Max Grant Award	Max Duration	Final Deadline
2016 Round I	$16 million	$200,000	12 months	March 31, 2017
2016 Round II	$18 million	$200,000	12 months	November 30, 2017
2017	$4.5 million	$100,000	12 months	May 31, 2018
2018	$9.5 million	$100,000	18 months	March 1, 2021
2019	$7 million	$100,000	18 months	December 15, 2021
2021	$43 to $45 million	$200,000	24 months	June 30 2024

In this table are, for each funding round, the maximum dispersal amount, the maximum grant award for a single project, the maximum duration of a grant, and the deadline for the project to be completed and operational. Due to the COVID-19 pandemic, there was no 2020 funding round.

round may be heterogeneous due to variation in the eligibility criteria or external factors such as annual variation in climate or market trends. The specific criteria used in each round is posted in the RFP and is visible to applicants in that funding round. To account for the heterogeneity associated with staggered treatment timing and variation in application criteria, I apply a non-parametric generalization of the DiD estimator proposed by Imai, Kim, and Wang (2021). This method uses existing matching and weighting techniques combined with the DiD estimator to estimate the treatment effect.

2.1 Irrigation Technologies Funded by SWEEP

Applicants may install technologies such as soil moisture or plant sensors, weather stations or ET-based irrigation scheduling, flow meters, pressure sensors, conversion from diesel to electric pump, installation of renewable energy, variable frequency drives, low pressure systems, and improved irrigation scheduling.[7] Different technologies may incentivize different types of adjustments along the intensive margin. Irrigation systems that require lower water pressure to irrigate and reduce evaporative loss may incentivize producers to irrigate more frequently and more intensely. More efficient irrigation has been shown to increase crop transpiration and yields, which may lead to adjustments to more water intensive or higher value crops and an increase in consumptive water use.

A number of previous studies have used simulations and analytical models to estimate how adoption of various irrigation efficiency technologies and practices influence water use. Medellín-Azuara, Howitt, and Harou (2012) use the UC Davis SWAP model to calibrate policy effects based on regional averages in the Tulare basin of California.[8] They estimate that irrigation subsidies increase capital investments

---

[7] Source: CDFA Office of Environmental Farming & Innovation, State Water Efficiency and Enhancement Program Request for Grant Applications, December 5, 2023.

[8] UC Davis SWAP Model Documentation, (Howitt et al., 2012)

in irrigation but without water restrictions, subsidies will not change total applied water. Using an agronomic and microeconomic model to analyze an expected rebound effect, Berbel and Mateos (2014) estimate that adoption of sprinkler or drip irrigation can reduce water use unless irrigated area is expanded. Christian-Smith, Cooley, and Gleick (2012) use an analytical model by the California Department of Water Resources (DWR) to estimate changes in water use under three scenarios: (1) conversion of flood irrigation to sprinkler or drip irrigation, (2) improved irrigation scheduling, e.g., soil moisture monitoring or weather stations, and (3) regulated deficit irrigation, i.e., applying less water to drought tolerant crops during certain growing stages. They estimate that more efficient irrigation and regulated deficit irrigation scheduling saves water by approximately 3.5-4% and that improved irrigation scheduling saves water by approximately 10%. These studies illustrate the potential heterogeneity in water saving benefits across irrigation technology improvements.

Since proposed irrigation technology improvements are contained within each project description, I construct an algorithm using NLP methods to categorize irrigation technologies by project. In the first step, I use the GPT-4 model, available through the OpenAI API, to extract irrigation technology names and descriptions from each project description using a standardized prompt that contextualizes the task.[9] In the second step of the algorithm, I use an embedding model, also available through the OpenAI API, to convert each technology description into a numerical vector.[10] Embedding text or phrases allow machine learning models to assign a numerical value based on the semantic relationships between words or phrases and perform clustering. Since each project description is written by the applicant, similar or identical technologies may have different names. For example, there are almost 200 unique descriptions for soil moisture related sensors and probes such as "soil moisture stations", "soil moisture sensor", "soil moisture probes", or "soil moisture monitoring system" and although these are all technically "distinct" names, they all refer to technologies with the same objectives that would result in similar changes to water and energy use. Generating an embedding for each of these phrases allows for the clustering of comparable technologies. I obtain an embedding for every distinct technology description from the first step. I then use the K-Means clustering algorithm to group each embedding into similar clusters. K-Means is a machine learning algorithm that clusters each data point

---

[9] GPT-4 prompt: "You will be provided with unstructured text that describes a proposed project on a farm. Your task is to extract the technology products that are going to be installed in this project. Do not extract keywords of existing technology products and do not extract product brands. Separate each technology with a comma. Here is the text:"

[10] The embedding model is called "text-embedding-3-small".

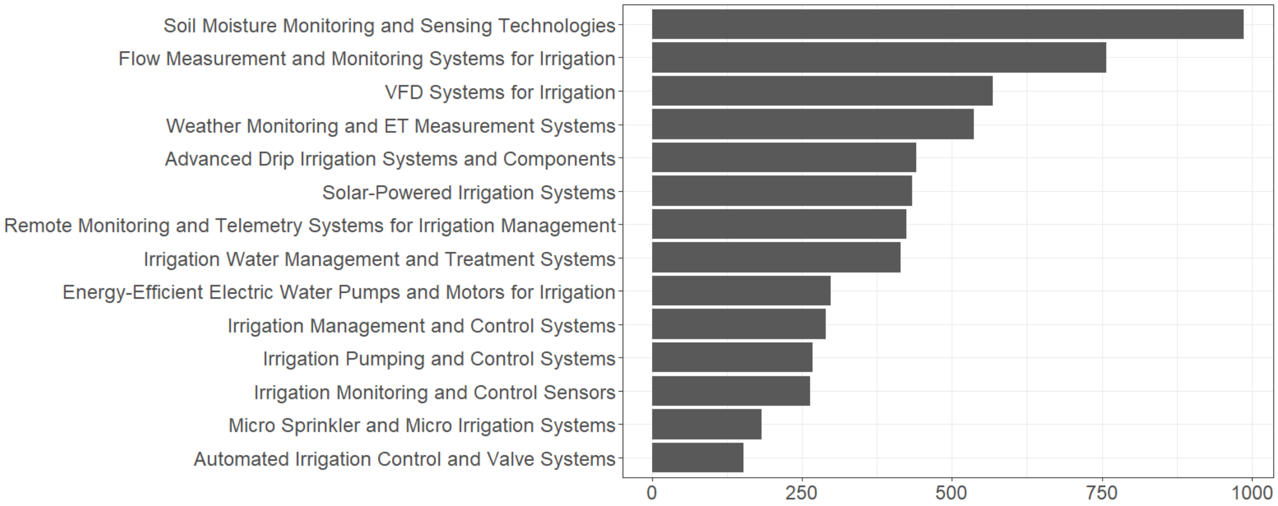
based on K number of centroids. In the final step, I cluster the different embedded technologies into 14 categories and assign common names and descriptions using GPT-4. Additional tables and figures related to the results of the embedding and clustering algorithm are provided in the appendix section A.

An example project description is: “This project proposes to replace the current flood irrigation system on an alfalfa field with a 2-set drip irrigation system. Part of this system will include a flowmeter to measure water use, a new 50 HP booster pump with a 50 HP VFD and a soil moisture probe, which will help with establishing irrigation schedules. These components will help aim the goal of decreasing the amount of water usage, increase uniformity, regulate pumping time and reduce GHG emissions.” In the first step of the algorithm, the GPT-4 model extracted the keywords: drip irrigation system, flowmeter, booster pump, VFD, and soil moisture probe. The embedding and clustering procedure then assigned the following clusters and standardized names to the extracted keywords: Soil Moisture Monitoring and Sensing Technologies, Flow Measurement and Monitoring Systems, VFD Systems, Irrigation Pumping and Control Systems, and Advanced Drip Irrigation Systems and Components.

The results of this procedure are presented in Figure 4. A full description of each technology category is provided in Table 8 in the Appendix. The most common technology category is Soil Moisture Monitoring and Sensing Technologies. Three quarters of all projects install a technology from this category. More than half of all projects install technologies from the Flow Measurement and Monitoring Systems for Irrigation category, followed by 44% of projects installing Variable Frequency Drives (VFD) Systems for Irrigation.

An interesting feature of the data is that projects install technologies from a variety of categories. On average, each application installs four to five technology categories, ranging from a minimum of one category to a maximum of eleven categories. Some of these technologies, such as weather monitoring or ET measurement and soil moisture monitoring are information technologies. They improve irrigation efficiency by increasing the amount of information irrigators have to make better informed decisions about the timing or location of irrigation. Alternatively, micro irrigation technologies or flow measurement and monitoring systems actually improve the precision with which water is applied or reduce the energy used by an irrigation system. This variation in the type of technologies funded by SWEEP illustrates the importance in how different applications potentially face differing incentives related to production decisions. There are 737 unique technology bundles across applications, with some

Figure 4: Frequency of Technology Categories



This figure presents the number of applications with each technology category. The total number of applications used in the analysis is 1,294. A project can install multiple technologies and therefore each technology category is not mutually exclusive.

Table 4: Count of Applications with Each Technology Category by Treatment Status

Technology Category	Awarded	Not Awarded	Total
Automated Irrigation Control and Valve Systems	65	88	153
Micro Sprinkler and Micro Irrigation Systems	71	112	183
Irrigation Pumping and Control Systems	97	171	268
Irrigation Monitoring and Control Sensors	130	133	263
Irrigation Management and Control Systems	132	157	289
Energy-Efficient Electric Water Pumps and Motors for Irrigation	136	162	298
Irrigation Water Management and Treatment Systems	152	262	414
Solar-Powered Irrigation Systems	180	253	433
Remote Monitoring and Telemetry Systems for Irrigation Management	192	232	424
Advanced Drip Irrigation Systems and Components	198	243	441
VFD Systems for Irrigation	222	346	568
Weather Monitoring and ET Measurement Systems	252	285	537
Flow Measurement and Monitoring Systems for Irrigation	346	412	758
Soil Moisture Monitoring and Sensing Technologies	443	544	987

This table presents the count of applications installing each technology category split between applicants that were awarded funded and not awarded funding. The groups are not mutually exclusive, a single application can be in multiple totals.

bundles appearing only once. Some technology categories may also only appear on a small number of applications, for example, Automated Irrigation Control and Valve Systems only appear on a total of 154 applications and only 66 treated applications. Table 4 presents the number of applications that install a technology from each group.

Table 5: Description of Data Sources

Data Source	Description	Level of Obs.	Years
SWEEP Application	Provided by CDFA, Public Records Act Request	APN-Applicant-Year	2016-2019, 2021
Assessor Parcel Boundaries	Geospatial file of parcel boundaries	APN	-
CA Statewide Crop Map	Spatial land use with field boundaries, CA DWR and LandIQ	Field-Year	2014, 2016, 2018-2023
Crop Choice	NASS Cropland Data Layer: 30m x 30m pixel of land use	Pixel-Year	2015, 2017
Evapotranspiration (ET)	OpenET: 30m x 30m pixel of ET (mm)	Pixel-Month	2014-2023
Precipitation	PRISM: 4km x 4km pixel of precipitation (mm)	Pixel-Month	2014-2023
Soil Water Properties	Soil Properties: 800m x 800m pixel of available water holding capacity	Pixel-Month	-
County-average Water Demand	CA DWR annual estimates of seasonal irrigation water requirements by crop group, 2016-2020	County-Year-Crop	2016-2020
California Public Utilities Data (Pending Receipt)	Agricultural operation level electricity data	Applicant-Year	2014-2023

This table presents each data source used, data description, and the level of observation.

3 Data

I use a combination of applicant-level data on SWEEP and geospatial crop and climate data to construct a panel data set for each applicant across a seven-year period, two pre years and four post years, where year zero corresponds to the year of application. Table 5 lists all data sources and their level of observation. A full description of the data cleaning and construction process is provided in the Appendix Section B.

The main data source is the SWEEP application data provided by CDFA through a Public Records Act request. The application data contains the project description, Assessor Parcel Number (APN), County, application score, number of acres impacted by the project, and all of the criteria that go into the application score such as whether or not the application falls within a critically overdrafted basin, the estimated water and GHG savings by the project, and the budget requested. I geographically locate each applicant using the reported APN(s). An APN is a series of fourteen numbers that are used as a file number to inventory or identify the property. These parcels are developed and maintained for assessment purposes by County Assessor Offices. Spatial files including parcel boundaries indexed by APN are available for each county in California.

Summary statistics for SWEEP applicant data and pre-treatment outcome variables in period $t-1$ are presented in Table 6. The statistics are presented separately for awarded applicants and not awarded applicants, and the p-value for the difference in group means is calculated for each variable. Because certain variables are only available in certain years, the last column of the table indicates which years the variable exists in the data. Additionally, since the score changed between 2017 and 2018 funding rounds the score is shown separately for before and after 2018. There are statistically significant differences in the score of awarded versus not awarded applicants in both sets of years. However, even in the post-2018 period when a minimum score for eligibility is required, the mean score of not awarded applicants is 38.5, indicating that many eligible applicants were unfunded in those years.

Impacted acres is a self-reported measure of the total acres that will be impacted by the project. The average between awarded (152 acres) and not awarded applicants (141 acres) is not statistically significant. Irrigation training, critically overdrafted basin, and soil management practices are binary variables indicating whether an application satisfies each consideration. A larger proportion of projects that are within critically overdrafted groundwater basins are funded, and the difference in the proportion of applications in a critically overdrafted basin between awarded and not awarded applicants is significant at the 1% level. Due to funding initiatives in certain years, a larger proportion of awarded applicants qualify as SDFR and benefiting priority populations. The source of water used for irrigation across awarded and not awarded applicants is consistent, with 64% of awarded applicants reporting only using groundwater, 6.6% only using surface water, and the remaining using a mixture of both groundwater and surface water. Using CDFA-provided water saving and GHG saving calculators, awarded applicants estimate water savings of 1.07 acre feet per acre per year and estimate GHG savings of of 1.89 tonnes of $C0_2$ equivalent per acre per year. The applicant estimated water savings is equal to approximately one-third of average seasonal irrigation crop water demand in the pre-period and 46% of average consumptive water use in the pre-period.

Parcel boundaries were created by County Assessor Offices for property assessment purposes, as such, these boundaries are generally larger than field boundaries and do not necessarily correspond to the level of planting decisions. Since assessor parcels do not correspond to field-level agricultural boundaries that correlate with the level that planting decisions are made, I convert parcel boundaries into field boundaries using California's Statewide Crop Maps published by the California Department of Water Resources (DWR). Figure 5 illustrates how a parcel for a single APN is deconstructed into

Table 6: Summary Statistics – SWEEP Application Data and Outcome Variables

Variable	Awarded		Not Awarded			
	Mean	N	Mean	N	p.value	years
Score (2016-2017)	3.76	321	2.31	185	0.000	2016-2017
Score (2018-2021)	42	245	38.5	553	0.000	2018-2021
Impacted Acres	152	566	141	733	0.398	All
Irrigation Training[a]	0.913	566	0.897	737	0.31	All
Critically Overdrafted Basin[b]	0.663	564	0.566	737	0.000	All
Soil Management Practices[c]	0.595	566	0.621	737	0.341	All
SDFR*	0.471	244	0.12	552	0.000	2018-2021
Priority Population**	0.428	404	0.177	627	0.000	2017-2021
Water Source: Only Groundwater	0.644	565	0.616	737	0.295	All
Estimated Water Savings (ac-ft/year/ac)	1.07	566	0.955	737	0.0351	All
Estimated GHG Savings (Tonnes of CO_2 equiv/ac/year)	1.89	566	0.836	738	0.395	All
Budget Requested	113,000	563	107,000	738	0.0661	All
Proportion Water Intensive (t-1)	0.768	560	0.791	732	0.268	All
Average Crop Water Demand (ac-ft/ac) (t-1)	3.15	560	3.16	732	0.782	All
Average Consumptive Water Use (ac-ft/ac) (t-1)	2.35	560	2.3	730	0.313	All

[a] Irrigation training is a binary variable that indicates whether the applicant agrees to attend irrigation training.

[b] Critically overdrafted Basin is a binary variable that indicates whether the project falls within a critically overdrafted groundwater basin.

[c] Soil Management Practices is a binary variable that indicates whether the applicant will also implement soil management practices.

* Socially Disadvantaged Farmers and Ranchers (SDFR) is a binary variable that indicates whether the applicant belongs to a socially disadvantaged group as defined by the 2017 Farmer Equity Act.

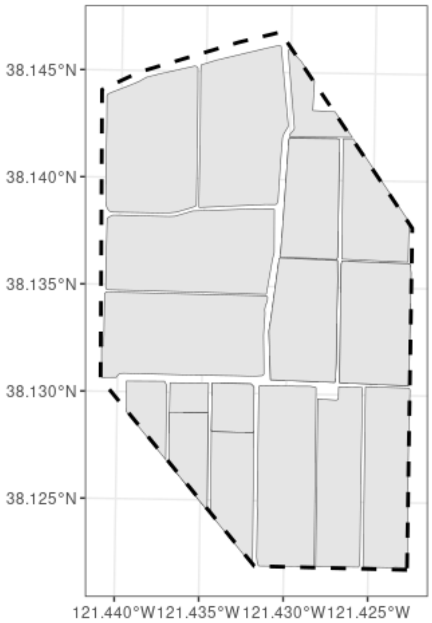
** Priority Population is a binary variable that indicates whether the proposed project benefits disadvantaged communities, low-income communities, or low-income households. Source: Cal Climate Investments

multiple fields as defined by the Statewide Crop Mapping dataset. The dashed black line corresponds to a single parcel boundary and the shaded gray boundaries are the individual fields that fall within that parcel. For years 2014, 2016, and 2018 through 2022 land use is classified using the Statewide Crop Mapping datasets. Land IQ worked in cooperation with DWR to categorize nearly 15.4 million acres into crop and land use types with accuracy exceeding 98%.[11] For the missing years, 2015 and 2017, crop choice by field is assigned using the U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) Cropland Data Layer (CDL). The CDL provides a geo-referenced raster of crop choice at the 30 meter by 30 meter pixel level. To reduce the impact of potential measurement error in CDL classification, the CDL is aggregated to the field-level using the modal crop on each field. To further increase the accuracy of the CDL classification, I implement two rules prior to calculating the modal crop: (1) if the prior year and the following year in Land IQ is classified as the same perennial

---

[11] Land IQ Land Use Mapping

Figure 5: Example: Single Assessor Parcel to Many Fields



This figure illustrates how a parcel for a single APN is deconstructed into multiple fields. The dashed black line corresponds to the single parcel boundary and the shaded gray boundaries are the individual fields that fall within that parcel.

(e.g., almonds), the CDL year must equal that perennial and (2) if the prior year and the following year in Land IQ is classified as an annual, the CDL year must not equal a perennial.

Consumptive water use is constructed at the field-year level using monthly mean ET, monthly mean precipitation, and soil water holding capacity. ET is equal to water that evaporates from land and water surfaces or transpires from plants and re-enters the atmosphere. It is the second-largest component of the water cycle, following precipitation, and is an important component in effective water management strategies. ET is often used to approximate consumptive water use, as it makes up the majority of water that is consumed, or in other words the portion of water that cannot be recovered or reused. OpenET is an open source data provider that uses satellite-based ET data for improved water management across the western United States. The primary inputs to calculate ET by OpenET come from satellites including Landsat, Sentinel-2, GEOS, and others, weather station networks and models, and field boundary and crop-type datasets.

Precipitation is constructed at the field-year level using data from PRISM Climate Group, which is available at the four kilometer by four kilometer pixel level. Soil water storage is constructed at the field level from from the UC Davis California Soil Resource Lab and UC ANR in collaboration with the USDA Natural Resource Conservation Service (Walkinshaw, O'Geen, and Beaudette, 2023). I combine these variables to construct a panel of crop choice, ET, precipitation, and soil water storage at the applicant-field level for the two years prior to application year and four years after application year.

3.1 Outcome Variable Construction

The two main responses I am interested in measuring are changes in crop composition and changes in water application. To estimate these, I construct three outcome variables to measure the program effect on crop choice and water use at the applicant level. Because cropping decisions are decided at the applicant level and there is spatial correlation across fields operated by the same applicant, I construct acre-weighted averages of all field-level outcomes to aggregate field-level decisions to the applicant-year level.

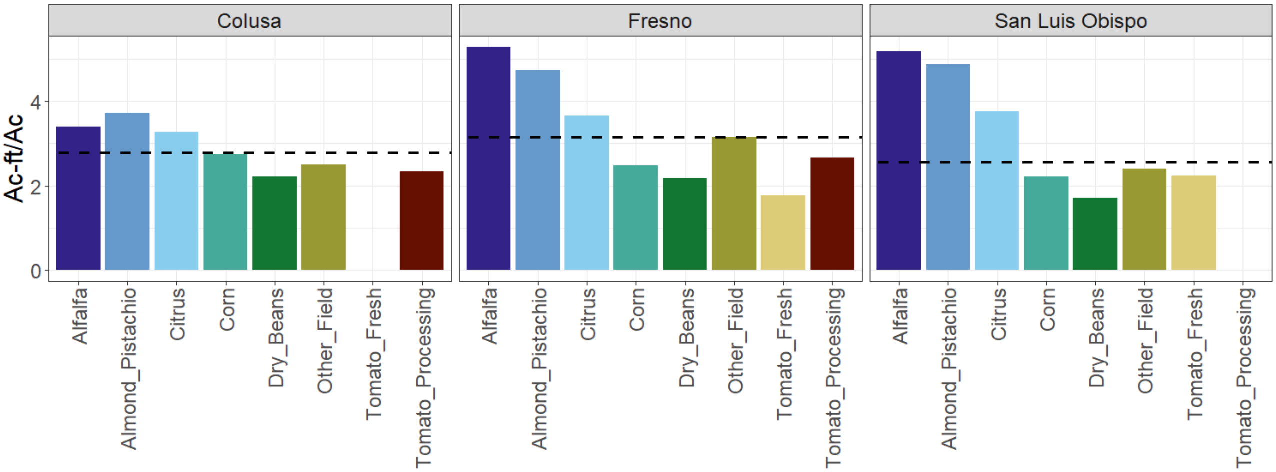
The first two outcome variables are used to estimate changes in crop composition. Since crop choice is a categorical variable, I index each crop to a measure representing relative water intensities using California DWR Agricultural Land & Water Use Estimates. DWR estimates a variety of water use variables, including applied water, for 20 crop categories between 1998 and 2020 at the state, hydrologic region, and county level.[12] Appendix Table 9 lists all 20 crop groups and their included crops. Applied water is estimated in both unit value (acre feet per acre) and volume (acre feet), and is estimated as the net amount of irrigation water needed to produce a crop divided by the mean seasonal irrigation system application efficiency. Soil characteristics data and crop information with precipitation and crop ET are used to generate hypothetical water balance irrigation schedules to determine the net amount of irrigation water needed to produce a crop.

I use the DWR estimates between 2016 and 2020 to calculate an average irrigation water demand in acre feet per acre for each crop group in each county. Therefore, crop water intensities are constant over time and within the same county. Water intensities can vary across counties depending on soil type, average precipitation during the growing season, and average irrigation efficiency. I calculate the median water intensity within each county and any crop within that county that has a water intensity above the mean is considered "water-intensive." This results in two variables at the applicant-year level that measure changes in crop composition over time: (1) Proportion of Acres Water Intensive

---

[12] DWR Land & Water Use Estimates

Figure 6: Crop Water Intensities by Crop and County (2016-2020)



County-level average crop group water demand for years 2016-2020 for three example counties and eight example crop groups. Dashed line in each panel is equal to county specific median.
Source: CA DWR Agricultural Land & Water Use Estimates, USDA NASS FSA-to-CDL Crosswalk

(%), and (2) Average Crop Water Intensity (Ac-ft/Ac).

Figure 6 presents an example including three counties and eight crop groups, to illustrate the difference in estimated water intensities across counties. For example, estimated water intensity for alfalfa in Colusa County is 3.55 ac-ft/ac, 4.16 ac-ft/ac in San Luis Obispo County, and 4.64 ac-ft/ac in Fresno County. The dashed line in each panel of the figure is equal to the county specific median.

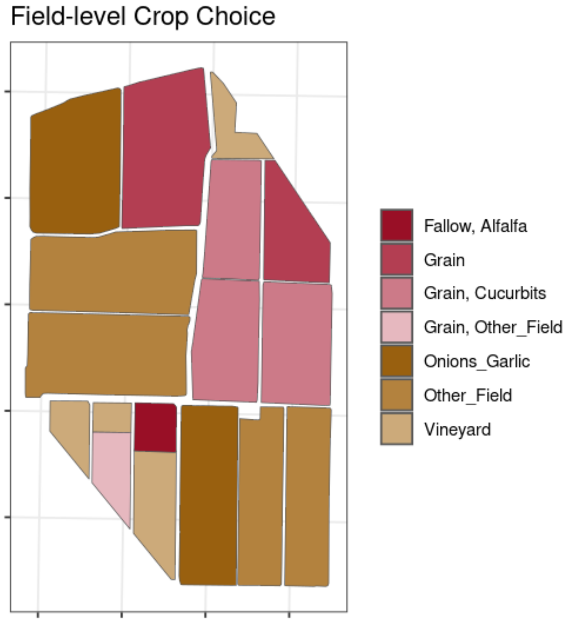
For each applicant I match the field-level crop choice to its corresponding DWR water intensity index and assign a binary indicator for whether it is considered water intensive using the relevant county median. Equations (1)-(3) illustrate the aggregation process to the applicant-level. Where f indexes a specific field that falls within SWEEP applicant i's parcel boundary, c indexes county, g indexes crop group, and t indexes year. Crop Water Intensity$_{fcg}$ is the estimated required water on field f for crop group g, and Median Crop Water Intensity$_{c}$ is the median required water across all crop groups in County c.

$$\text{Water Intensive}_{fcg} = \mathbb{1}\{\text{Crop Water Intensity}_{fcg} > \text{Median Crop Water Intensity}_{c}\} \tag{1}$$

$$\text{Proportion Water Intensive}_{it} = \frac{\sum_{f\in i} \text{Acres}_{if} * \text{Water Intensive}_{fcg}}{\sum_{f\in i} \text{Acres}_{if}} \tag{2}$$

$$\text{Average Crop Water Intensity}_{it} = \frac{\sum_{f\in i} \text{Acres}_{if} * \text{Crop Water Intensity}_{fcg}}{\sum_{f\in i} \text{Acres}_{if}} \tag{3}$$

Figure 7: Example Fields for a Single Applicant in a Single Year



This figure presents the crops planted across a group of fields associated with a single applicant in a single year. The county-level crop water intensities are used to calculate acre-weighted averages to calculate this applicants average crop water intensity (Equation 3) and proportion water intensive (Equation 2). Fields that have more than one crop associated with them correspond to that field being multi-cropped during the year.

Figure 7 presents an example of the distribution of crop groups across multiple fields for a single applicant in a single year. To construct a single year observation for this applicant, crop water intensities are averaged across all fields using the field acre-weighted average from Equation (3).

To evaluate changes in water use, I construct a third variable that estimates consumptive water use: (3) Consumptive Water Use (Ac-ft/Ac). This variable relies on remotely sensed weather data to measure water consumption at the individual field level. Monthly water consumption is calculated by first subtracting monthly effective precipitation from monthly ET. Consumption is then summed across the water year.[13] A few definitions in addition to ET are important for understanding consumptive water use. Effective precipitation is defined as the portion of rainfall that can be used to meet the ET of growing crops, excluding the surface runoff or percolation below the crop root zone (Soil Conservation Service, 1993). Using this component of precipitation is important because it represents the portion of measured ET that was not applied by a farmer to a specific field and instead was consumed from pre-

---

[13] A water year, also known as the hydrological year, is a 12-month period that runs from October 1 to September 30 and is used to track water movement and precipitation. The 2022 water year, for example, is calculated from October 1, 2021 to September 30, 2022.

cipitation. Subtracting total precipitation from ET will overstate the portion of ET that is attributable to rainfall so it is important to account for just the portion of rainfall that is considered effective. Using 50 years of rainfall records at locations throughout the United States the USDA Soil Conservation Service (SCS) developed a technique to predict monthly effective precipitation using mean precipitation, average crop ET, and soil water storage factor (Soil Conservation Service, 1993; Walkinshaw, O'Geen, and Beaudette, 2023). The soil water storage factor translates to how much water a soil can hold for crops to use. Equations 4 and 5 define how effective precipitation is calculated for each month in the sample based on Soil Conservation Service (1993).

$$\mathrm{P}_e = \mathrm{SF}\left(0.70917 \times \mathrm{P}_t^{0.82416} - 0.11556\right)\left(10^{0.02426 \times \mathrm{ET}_c}\right) \tag{4}$$

$$\mathrm{SF} = 0.531747 + 0.295164 \times \mathrm{D} - 0.057697 \times \mathrm{D}^2 + 0.003804 \times \mathrm{D}^3 \tag{5}$$

Where P_e is the average monthly effective precipitation in inches, P_t is the monthly mean precipitation in inches, ET_t is the average monthly crop transpiration from OpenET in inches, SF is the soil water storage factor calculated based on the usable soil water storage, D, in inches. The average monthly effective precipitation calculated cannot exceed either the average monthly rainfall or average monthly ET.

Average consumptive water use at the applicant level is then calculated as the acre-weighted average of field level consumptive water use. Appendix Figure 17 illustrates the flow of water applied to an agricultural field to its various destinations in return flows or consumption. Equations (6) and (7) illustrate the aggregation process from field-level consumption to average consumptive water use by water year, where m is month in water year t.

$$\text{Consumption}_{ift} = \sum_{m \in t} \mathrm{ET}_{ifm} - \text{Effective Precipitation}_{ifm} \tag{6}$$

$$\text{Average Consumptive Water Use}_{it} = \frac{\sum_{f \in i} \text{Acres}_{if} * \text{Consumption}_{ift}}{\sum_{f \in i} \text{Acres}_{if}} \tag{7}$$

4 Empirical Method

Due to the temporal heterogeneity in SWEEP application criteria and potential heterogeneity in treatment effect depending on the year due to common economic and climate characteristics, I use an empirical strategy that controls for staggered treatment timing and accounts for heterogeneity across units. I leverage observable applicant characteristics from the SWEEP application data, geographic characteristics such as hydrologic growing region and soil type, and pre-treatment outcome variables in a matched DiD estimator that also incorporates time-series cross-sectional data (Imai, Kim, and Wang, 2021). Treated observations are individuals that applied to SWEEP and received funding for their proposed project and control observations are individuals that applied to SWEEP and did not receive funding. There are two major challenges in evaluating conservation programs. First, due to the voluntary nature of participation, it can be difficult to identify a valid control group due to selection bias into the program. I leverage two characteristics of SWEEP to overcome the major challenges in evaluating conservation programs. California state budget constraints mean that the program is consistently over-subscribed in each fiscal year, there are more eligible applicants than available funding. Since requested funds exceeded available funds in all years and many unfunded applications had scores that made them eligible for funding, it is plausible that there exist control applicants (i.e., applicants that did not receive funding) that are comparable to treated applicants. Second, applications are evaluated based on specific criteria that are frequently not observable to the researcher. I use granular data on SWEEP application components, geographic characteristics, and pre-treatment outcome variables to refine the set of control applicants that are most similar to treated observations from the same funding year.

The estimation procedure is conducted in three steps.[14] In step one, a set of control units is defined for each treated observation. The set of control units is equal to all control applicants that have the same treatment history as the current treated observation, i.e., each control unit is an applicant from the same funding year that did not receive funding.

In the second step, the control set for each treated unit is refined for similar covariate values using matching. I use the Covariate Balancing Propoensity Score (CBPS) weighting method (Imai and Ratkovic, 2014). The CBPS method models treatment assignment while optimizing the covariate balance, which improves the performance over alternative propensity score matching and weighting methods. In the Appendix Section C.1 I report the main results using alternative matching methods: CBPS matching, Mahalanobis distance matching, propensity score matching, and propensity score weighting. The results are not sensitive to the choice of matching method. The main restriction

---

[14] This estimation procedure is available in an open source R package PanelMatch.

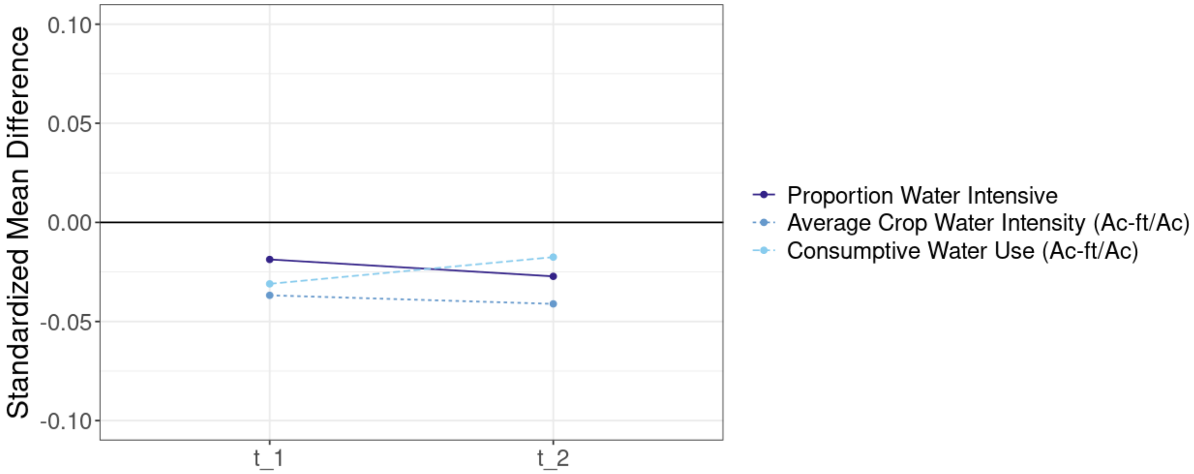
I impose during the matching procedure is an exact match on funding year and hydrologic region. This ensures that treated individuals are only matched with untreated individuals from the same funding year and are within the same hydrologic region. Matching within the same funding year controls for potential unobservable variables that are correlated with time and matching within the same hydrologic region controls for potential unobservable variables that are correlated with growing region characteristics. The other covariates used to match on are: lagged outcome values, application score, applicant-estimated water savings, agreement to attend irrigation training, project falling within a critically overdrafted basin, agreeing to apply soil management practices, irrigation water source, project impacted acres, project requested budget, and soil type. The result of the second step assigns, to each individual treated observation, a weight to each control observation within the matched control set based on the estimated propensity score.

The third and final step applies the DiD estimator to each treated observation using the weighted average of all matched control units. When using the CBPS weighting method, each control unit in the refined matched set is assigned a weight of ω_{it}. Estimating equation (8) presents how the matched DiD is aggregated in each period to estimate the treatment effect. $f \in F$ is equal to the current lead, where $F = \{0, 1, 2, 3, 4\}$, and L is the number of lags that the treatment history must be equal, in my setting I use two lags. Equal treatment history, regardless of the number of lags chosen, is automatically satisfied by construction. $D_{it} = X_{it}(1 - X_{i,t-1}) * \mathbb{1}\{|\mathcal{M}_{it}| > 0\}$ is an indicator function equal to one if an observation is treated (i.e., $X_{it} = 1$), and has more than one unit in its matched set, and zero otherwise, where $\mathcal{M}_{it}$ is the matched set for i. $\omega_{it}^{i'}$ are CBPS weights for each matched control unit i'. Y_{it} is the outcome variable for treated unit i in period t. I estimate treatment effects from t up to period T. In essence, this equation is a weighted average of the difference in the treated observation treatment effect and the weighted average of the control sets treatment effects. Standard errors are calculated using 1,000 weighted bootstrap samples to incorporate the uncertainty from matching.

$$\hat{\delta}(f, L) = \frac{1}{\sum_{i=1}^{N} \sum_{t=L+1}^{T-f} D_{it}} \sum_{i=1}^{N} \sum_{t=L+1}^{T-f} D_{it} \left((Y_{i,t+f} - Y_{i,t-1}) - \sum_{i' \in \mathcal{M}_{it}} \omega_{it}^{i'} (Y_{i',t+f} - Y_{i',t-1}) \right) \tag{8}$$

Similar to standard DiD methods, identification relies on a conditional parallel trend assumption. Conditional on the treatment, outcome, and covariate history up to time $t - L$, the treatment assignment is unconfounded, i.e., the trend in mean untreated outcomes is independent of observed treatment

Figure 8: Lagged Outcome Balance - Standardized Mean Difference

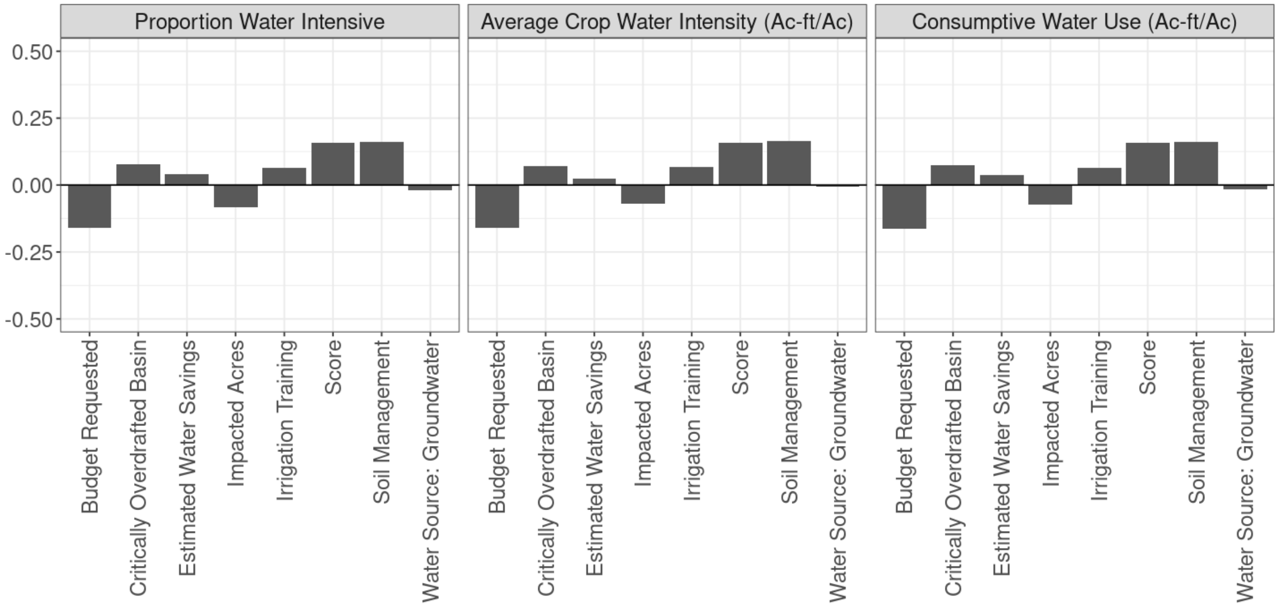


This figure presents the balance in the lagged outcome variables between treated and control observations. This standardized mean difference is calculated as the mean difference of each lagged outcome between each treated observation and its matched control observations and standardized by the standard deviation of the outcome variable. This figure suggests that treated and control groups are well-balanced after matching since the standardized mean difference is less than 0.05 for all outcomes.

status. To evaluate the validity of the conditional parallel trend assumption for each outcome variable, I calculate the mean difference of each outcome between a treated observation and it's matched control observations. This difference is then standardized by the standard deviation of each outcome across all treated observations such that the mean difference is measured in terms of standard deviation units (Imai, Kim, and Wang, 2021). Figure 8 presents the results of this analysis and suggests that treated and control groups are well-balanced after matching and do not violate the parallel trend assumption; the standardized mean difference is less than 0.05 standard deviations in all outcomes.

To evaluate post-matching covariate balance, the same method is applied to each matched covariate. Figure 9 presents the standardized mean difference for following main covariates used in matching: budget requested, critically overdrafted basin, estimated water savings, impacted acres, irrigation training, score, soil management practices, access to surface water. This figure provides further evidence that treated and control groups are well-balanced after matching and do not violate the parallel trend assumption; the standardized mean difference is less than 0.25 standard deviations in all outcomes.

Figure 9: Lagged Covariate Balance - Standardized Mean Difference

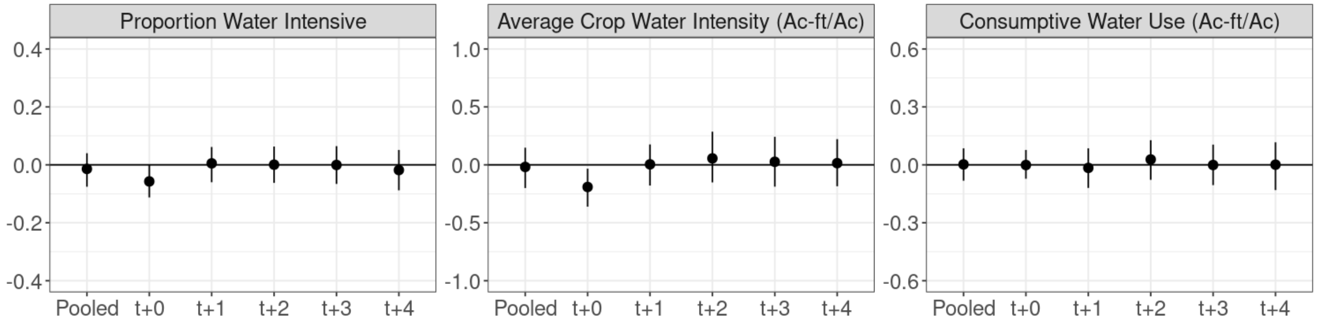


This figure presents the balance of the lagged covariate variables between treated and control observations. The standardized mean difference is calculated as the mean difference of each lagged outcome between each treated observation and its matched control observations and standardized by the standard deviation of the covariate. This figure suggests that treated and control groups are well-balanced after matching since the standardized mean difference is less than 0.25 for all outcomes.

5 Results

Table 7 presents the main results for the crop choice and consumptive use variables for all matched funded and unfunded applicants. These results are also shown graphically in Figure 10. These results show that for all applicants, on average, I find no significant effect of SWEEP participation on crop choice or consumptive water use and I can reject potentially large changes in crop composition or water use. The first two panels in Figure 10 provide information about whether the program is causing a shift in cropping decisions. The far left panel, Proportion Water Intensive, shows in each period the average change in the proportion of total acres that are planted to a water-intensive crop between funded and unfunded applicants. The middle panel, Average Crop Water Intensity (Ac-ft/Ac), presents the average change in the estimated seasonal crop irrigation water demand, measured in acre feet per acre. If funded applicants shifted towards crops that are more water-intensive following irrigation efficiency improvements we would expect to see a positive coefficient in these two panels. I find no significant changes in crop composition between funded and unfunded applicants and can reject large changes in

Figure 10: Results: Crop Choice and Consumptive Water Use



This figure presents the point estimates and 95% confidence intervals for each outcome variable. Standard errors are calculated using 1,000 weighted bootstrap samples.

the proportion of crops that are water-intensive by more than 8.8 percentage points.

The third panel, Consumptive Water Use (Ac-ft/Ac), provides information about changes in actual consumptive water use based on the remotely sensed ET and precipitation data. The point estimates in each period equal the average change in consumptive water use between funded and unfunded applicants. If funded applicants changed their irrigation decisions following efficiency improvements such that they irrigate more intensively, we would expect to see a positive coefficient. Alternatively, if improved irrigation scheduling led to changes in irrigation water management that reduce water use, we would expect to see a negative coefficient. I find no significant changes in consumptive water use and can reject large changes in consumptive water use of more than six percent.

Taken together, these results suggest that funded SWEEP applicants are not changing their behavior in ways that significantly impact crop water demand or consumptive water use compared to unfunded SWEEP applicants. However, given the variation in project types and applicants, it may be that there are certain applications that are more or less effective. I conduct three heterogeneity analyses based on the type of water used for irrigation, the regional groundwater regulatory standards, and the type of technology proposed.

5.1 Heterogeneity Analysis

Treatment effects may be concentrated within certain sub-groups of SWEEP applicants. To evaluate other potential mechanisms that may affect crop choice and water use, I estimate the effects separately for three types of potential heterogeneity: (1) location-specific groundwater regulatory standards, (2)

Table 7: Crop Choice and Consumptive Water Use Results

Outcome Variable	Estimate	Standard Error	2.5%	97.5%	Period
Proportion Water Intensive	-0.057	0.028	-0.113	0	t+0
	0.005	0.03	-0.06	0.062	t+1
	0.0005	0.033	-0.062	0.063	t+2
	-0.0005	0.033	-0.066	0.064	t+3
	-0.018	0.036	-0.088	0.051	t+4
	-0.014	0.029	-0.075	0.04	Pooled
Average Crop Water Intensity (ac-ft/ac)	-0.191	0.086	-0.361	-0.033	t+0
	0.004	0.092	-0.178	0.176	t+1
	0.055	0.114	-0.152	0.286	t+2
	0.026	0.11	-0.188	0.242	t+3
	0.014	0.104	-0.185	0.222	t+4
	-0.018	0.089	-0.2	0.148	Pooled
Consumptive Water Use (ac-ft/ac)	-0.001	0.039	-0.072	0.077	t+0
	-0.016	0.052	-0.12	0.086	t+1
	0.028	0.052	-0.078	0.128	t+2
	-0.001	0.056	-0.105	0.105	t+3
	0.001	0.065	-0.131	0.116	t+4
	0.002	0.044	-0.082	0.085	Pooled

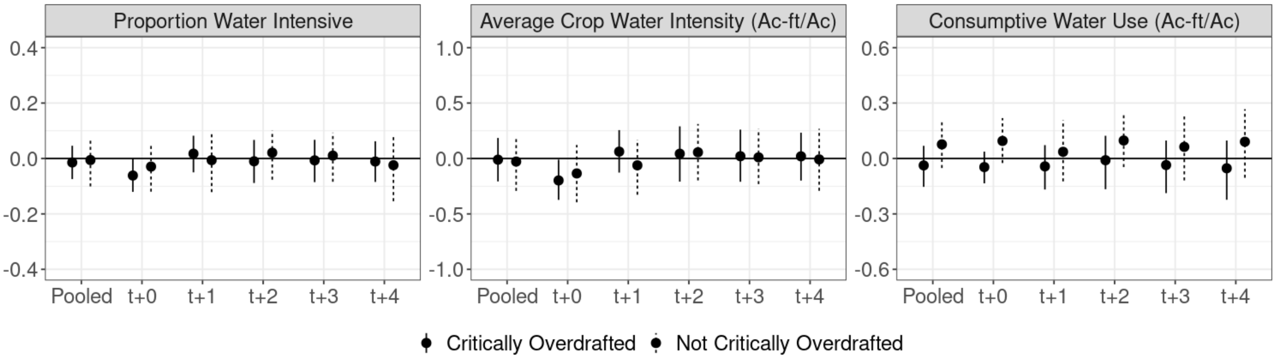
This table presents the results from the matched DiD of all SWEEP applicants for each outcome variable. Each treated observation is matched to control units using an exact match on funding year and hydrologic region and then matched on lagged outcome values, application score, applicant-estimated water savings, agreement to attend irrigation training, project falling within a critically overdrafted basin, agreeing to apply soil management practices, irrigation water source, project impacted acres, project requested budget, and soil type. Included are the point estimates and 95% confidence intervals. Standard errors are calculated using 1,000 weighted bootstrap samples.

whether the applicant has access to surface water for irrigation, and (3) the types of technologies installed by project. To estimate these results, I first split the sample based on the relevant moderating variable, and then perform the covariate balancing propensity score matching and difference-in-differences estimation on each sub-sample.

5.1.1 Heterogeneity: Groundwater Regulatory Standards

Each SWEEP proposal must indicate whether or not the project region is within a critically overdrafted groundwater basin. Projects within critically overdrafted basins are subject to more stringent groundwater regulatory standards through SGMA. Basins designated as medium priority, high priority, and critically overdrafted will have increased regulation and oversight. Although implementation of Groundwater Sustainability Plans (GSPs) is not required for the majority of the study period, deadlines are in 2020 for critically overdrafted basins and 2022 for medium and high priority basins, the timeline and designation of basin priority was completed before the study period. I use an indicator for

Figure 11: Heterogeneity Results: Groundwater Regulatory Standards



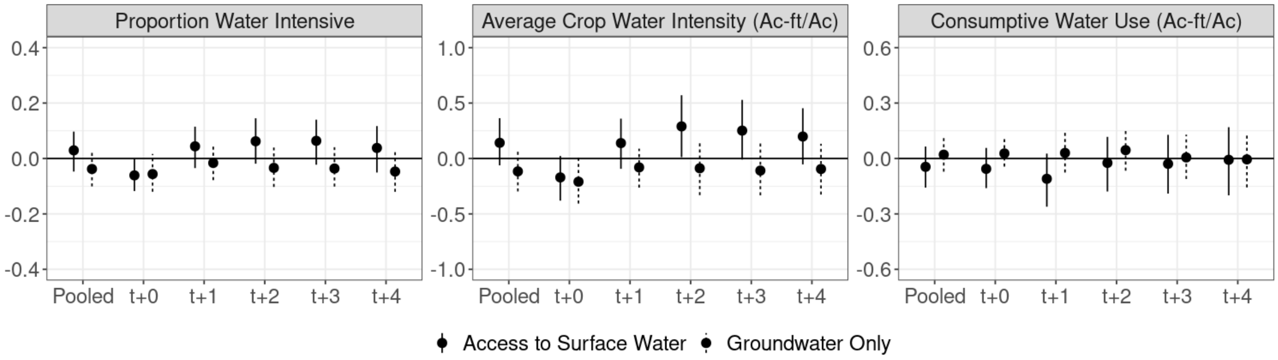
This figure presents the point estimates and 95% confidence intervals based on groundwater regulatory standards. Standard errors are calculated using 1,000 weighted bootstrap samples. The top row presents estimates for projects that are not within a critically overdrafted groundwater basin. The bottom row includes estimates for projects that are in a critically overdrafted basin.

whether a project lies within a critically overdrafted basin or not as a proxy for groundwater regulatory strength.

Intuitively, the effect of more stringent groundwater regulation on crop choice and consumptive water use is ambiguous. On the one hand, applicants within a critically overdrafted basin may have a greater incentive to curtail groundwater use in preparation for compliance with SGMA compared to applicants that do not fall in a critically overdrafted basin. They also may have less ability to increase groundwater use in response to irrigation technology efficiency improvements that may otherwise incentivize adjustments to water use from crop choice or consumptive water use. On the other hand, agricultural producers may have an incentive to increase their baseline groundwater use before GSP are implemented so they will have to curtail less water when SGMA takes effect.

Figure 11 present the point estimates and 95% confidence intervals comparing applicants within and outside of a critically overdrafted groundwater basin. The solid confidence band corresponds to those within a critically overdrafted basin and the dotted confidence band corresponds to projects outside of a critically overdrafted basin. The results are consistent with the main regression, I find no significant difference in behavior related to crop choice or water use for applicants depending on groundwater regulatory standards.

Figure 12: Heterogeneity Results: Irrigation Water Source



This figure presents the point estimates and 95% confidence intervals based on irrigation water source. Standard errors are calculated using 1,000 weighted bootstrap samples. The top row presents estimates for applicants that report using only groundwater for irrigation. The bottom row includes estimates for applicants that have access to surface water for irrigation.

5.1.2 Heterogeneity Analysis: Irrigation Water Source

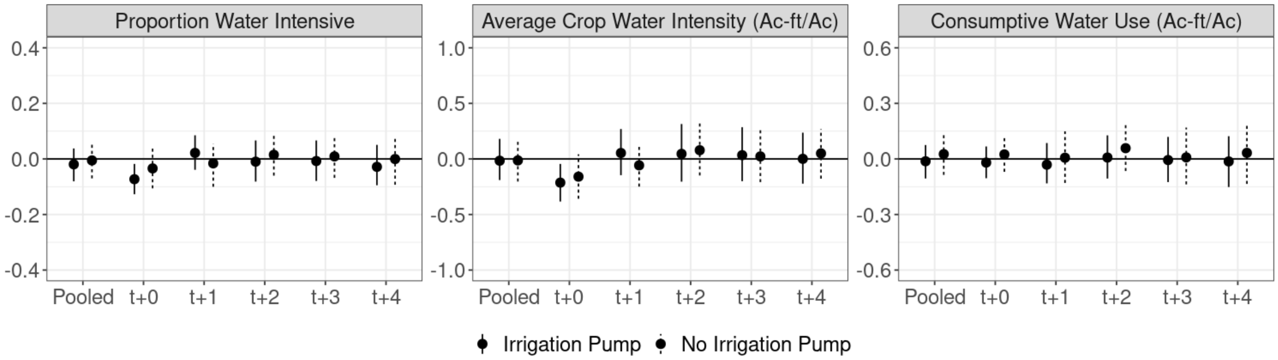
California's "beneficial use" water rights system may incentivize funded water users to switch to more water-intensive crops compared to their unfunded counterparts. If this is the case, we would expect to see a positive coefficient on the crop choice variable for applicants that have access to surface water only. Groundwater users may also have an incentive to switch to more water-intensive crops if increased efficiency reduces the marginal cost of water.

Figure 12 presents the point estimates and 95% confidence intervals for the water source heterogeneity analysis. The solid confidence band corresponds to applicants that report having access to surface water. The dotted confidence band corresponds to applicants that report using only groundwater for irrigation. I find that access to surface water increases the average crop water intensity for funded applicants compared to unfunded applicants. These results suggest that these funded applicants switch their cropping decisions to crops that, on average, have a higher estimated irrigation water demand.

5.2 Heterogeneity Analysis: Technology Categories

As discussed in Section 2.1, applicants install a wide variety of technology categories for each project and some categories are only installed on a small number of projects. These technologies also have a range of potential impacts on water and energy use. Some technologies provide additional information about irrigation scheduling and water management while other technologies improve the precision with which

Figure 13: Heterogeneity Results: Irrigation Pump Systems



This figure presents the point estimates and 95% confidence intervals by irrigation pump technology. Standard errors are calculated using 1,000 weighted bootstrap samples. The top row presents estimates for projects that do not install irrigation pump technologies that improve pump efficiency. The bottom row includes estimates for projects that install irrigation pump systems.

water is applied to a crop and reduce the energy required to irrigate a field. To evaluate whether certain technologies exhibit heterogeneous treatment effects from the program, I constructed two aggregated technology groups to focus on irrigation pump system improvements and micro irrigation technologies.[15]

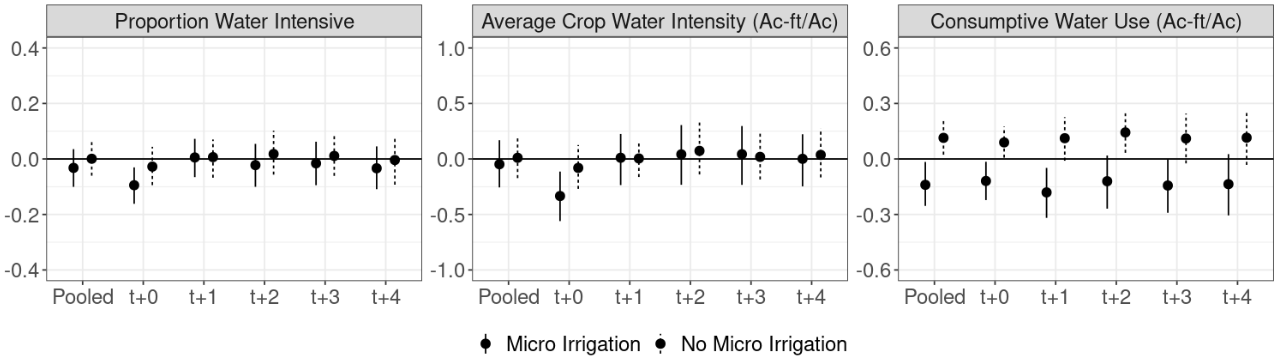
Irrigation pump system improvements includes three of the fourteen technology categories: Variable Frequency Drive Systems, Irrigation Pumping and Control Systems, and Energy-Efficient Electric Water Pumps and Motors. All three of these technology categories include components that improve efficiency of water delivery and reductions in energy use. Figure 13 presents the point estimates and 95% confidence intervals by irrigation pump technology. The dotted confidence band corresponds to projects that do not install irrigation pump system technologies and the solid confidence band corresponds to projects that do install irrigation pump system technologies. I find results consistent with the main results, and there is no difference in program response for applicants installing irrigation pump versus no irrigation pumps.

Micro irrigation technologies includes two of the fourteen technology categories: Micro Sprinkler and Micro Irrigation Systems, Advanced Drip Irrigation Systems and Components. These technologies improve the precision with which water is applied to crops, facilitate irrigation monitoring and scheduling, and reducing evaporation and run-off. Figure 14 presents the point estimates and 95% confidence intervals by application type. The solid confidence band corresponds to projects that install

---

[15] These groups were constructed based on personal communication with a UC Davis Professor of Water Management and Irrigation Engineer based on expected efficiency improvements and potential producer responses.

Figure 14: Heterogeneity Results: Micro Irrigation Technologies



This figure presents the point estimates and 95% confidence intervals by micro irrigation technologies. Standard errors are calculated using 1,000 weighted bootstrap samples. The top row presents estimates for projects that do not install micro irrigation technologies. The bottom row includes estimates for projects that install micro irrigation technologies.

micro irrigation technologies and the dotted confidence band corresponds to projects that do not install micro irrigation technologies. I find that funded projects that install micro irrigation technologies are effective at reducing consumptive water use, with up to 0.18 acre feet per acre average reduction for funded applicants compared to their unfunded counterparts. This is about an 8% decrease from the pre-treatment average consumptive water use. However, these reductions are almost entirely offset by those projects that do not install micro irrigation technologies. These funded applicants increase their consumptive water use by up to 0.14 acre feed per acre.

5.3 Program Additionality

The main results of this paper illustrate how participation in irrigation subsidy programs affect crop choice and consumptive water use by comparing funded applicants to their similar unfunded counterparts. I find that on average, funded applicants do not behave differently than unfunded applicants in terms of these outcomes, but that certain types of applicants do respond depending on irrigation water source or technology bundles. To further understand the impacts of these types of programs I have two ongoing extensions. The first extension relates to the additionality of conservation programs. A conservation program is additional if applicants would not install the proposed technologies without funding. If SWEEP is non-additional, and applicants to the program that do not receive funding install the proposed irrigation efficiency improvements anyway, then any efficiency-related incentives to adjust crop choice or consumptive water use are the same for both the treatment and control group. This would result in a true treatment effect of the program being zero, since applicant behavior regardless of the treatment (i.e., being funded) is the same. To evaluate the additionality of SWEEP, I construct a control group of farms that never applied to SWEEP. Using the California Parcel Boundary shapefiles, I exclude any APNs that correspond to a SWEEP application. I disaggregate the remaining parcels to the field-level using the CA Statewide Crop Mapping dataset, grouping fields that fall within the same parcel. Each outcome variable for farms that never applied to SWEEP is then aggregated to the parcel-year level. I define the treatment group as all SWEEP applicants, regardless of funding status. I then use the same empirical method to match control parcels to SWEEP applicants based on the lagged outcome variable, being within a critically overdrafted basin, and soil type. Treated and control units are exactly matched within the same County to account for potential unobservable variables that are correlated with regional characteristics.

5.4 Future Work: Energy Use

The second extension estimates the potential energy savings from irrigation efficiency improvements. I am using California electricity data from public utility companies to estimate how SWEEP enrollment impacts energy use at the farm-level. Using electricity data, I will also estimate changes in groundwater extraction following Martindill, Good, and Loge (2021). The change in return flow, or the water that is not consumed by ET, can be calculated as the difference between total groundwater extraction and consumptive water use. Increased irrigation efficiency may affect water use through both consumptive use as well as changes in return flows.

6 Conclusion

This paper uses NLP methods and field-level geospatial data to evaluate the effect of irrigation efficiency subsidy programs on agricultural production decisions. I use NLP methods to systematically categorize individual project components into technology groups and find large variation in project types and technology bundles. There are 737 unique technology combinations across applications, with an average of four to five technology groups per application. The types of technologies installed on each project range from information technologies, such as soil moisture monitoring and sensing technologies, or technologies that control water or energy efficiency such as micro irrigation or irrigation pumps and motors.

I use a non-parametric generalization of the DiD estimator to control for treatment effect heterogeneity and staggered treatment timing (Imai, Kim, and Wang, 2021). I estimate treatment effects comparing SWEEP applicants that received funding to SWEEP applicants that did not receive funding. This is the first paper, to my knowledge, that evaluates the impact of SWEEP on crop choice and ex post consumptive water use.

On average, I find no significant effect of SWEEP participation on crop choice or consumptive water use, however, I find important heterogeneities in program impacts. Having access to surface water for irrigation induces a shift in crop choice to crops with a higher estimated water requirement but no increase in consumptive water use for awarded applicants compared to not awarded applicants, reflecting that improved irrigation efficiencies may allow producers to shift to more water-intensive crops and adjust irrigation scheduling to meet crop consumptive water use more efficiently and not increase consumptive use. Applications that install micro irrigation technologies are effective at reducing consumptive water use, however these effects are almost entirely offset by projects that do not install micro irrigation and increase their consumptive water use.

The results of this paper highlight key issues that arise from a lack of targeting in conservation subsidy programs. SWEEP, and related programs, fund a broad range of potential technologies, some of which reduce water use, but ultimately when funding decisions are not made along the margins that most significantly decrease water use then these programs have minimal impact on conservation goals.

References

Berbel, J., and L. Mateos. 2014. "Does investment in irrigation technology necessarily generate rebound effects? A simulation analysis based on an agro-economic model." *Agricultural Systems* 128:25–34.

Binswanger, M. 2001. "Technological progress and sustainable development: what about the rebound effect?" *Ecological economics* 36:119–132.

Blumberg, J., C. Goemans, and D. Manning. 2022. "Producer Beliefs and Conservation: The Impact of Perceived Water Scarcity on Irrigation Technology Adoption." Working paper, National Bureau of Economic Research.

Börner, J., K. Baylis, E. Corbera, D. Ezzine-de Blas, J. Honey-Rosés, U.M. Persson, and S. Wunder. 2017. "The effectiveness of payments for environmental services." *World development* 96:359–374.

Brännlund, R., T. Ghalwash, and J. Nordström. 2007. "Increased energy efficiency and the rebound effect: Effects on consumption and emissions." *Energy economics* 29:1–17.

Callaway, B., and P.H. Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of econometrics* 225:200–230.

Cameron-Harp, M.V., and N.P. Hendricks. 2024. "Efficiency and water use: Dynamic effects of irrigation technology adoption." *Journal of the Association of Environmental and Resource Economists*, pp. .

Christian-Smith, J., H. Cooley, and P.H. Gleick. 2012. "Potential water savings associated with agricultural water efficiency improvements: a case study of California, USA." *Water Policy* 14:194–213.

Evans, R.G., and E.J. Sadler. 2008. "Methods and technologies to improve efficiency of water use." *Water resources research* 44.

Freire-González, J. 2019. "Does water efficiency reduce water consumption? The economy-wide water rebound effect." *Water Resources Management* 33:2191–2202.

Fuglie, K.O., and C.A. Kascak. 2001. "Adoption and diffusion of natural-resource-conserving agricultural technology." *Applied Economic Perspectives and Policy* 23:386–403.

Gillingham, K., D. Rapson, and G. Wagner. 2016. “The rebound effect and energy efficiency policy.” *Review of Environmental Economics and Policy*, pp. .

Gómez, C.M., and C.D. Pérez-Blanco. 2014. “Simple myths and basic maths about greening irrigation.” *Water Resources Management* 28:4035–4044.

Greening, L.A., D.L. Greene, and C. Difiglio. 2000. “Energy efficiency and consumption—the rebound effect—a survey.” *Energy policy* 28:389–401.

Hagerty, N. 2021. “Adaptation to surface water scarcity in irrigated agriculture.” *Working paper*, pp. .

Howitt, R.E., J. Medellín-Azuara, D. MacEwan, and J.R. Lund. 2012. “Calibrating disaggregate economic models of agricultural production and water management.” *Environmental Modelling & Software* 38:244–258.

Imai, K., I.S. Kim, and E.H. Wang. 2021. “Matching methods for causal inference with time-series cross-sectional data.” *American Journal of Political Science*, pp. .

Imai, K., and M. Ratkovic. 2014. “Covariate balancing propensity score.” *Journal of the Royal Statistical Society Series B: Statistical Methodology* 76:243–263.

Jevons, W.S. 1865. “The coal question: Can Britain survive.” *First published in* 914.

Karwowski, N., and M. Skidmore. 2023. “Nature’s Kidneys: the Role of the Wetland Reserve Program in Restoring Water Quality.”, pp. .

Li, H., and J. Zhao. 2018. “Rebound effects of new irrigation technologies: The role of water rights.” *American Journal of Agricultural Economics* 100:786–808.

Liu, P., Y. Wang, and W. Zhang. 2022. “The influence of the Environmental Quality Incentives Program on local water quality.” *American Journal of Agricultural Economics*, pp. .

Liu, P.W., J.S. Famiglietti, A.J. Purdy, K.H. Adams, A.L. McEvoy, J.T. Reager, R. Bindlish, D.N. Wiese, C.H. David, and M. Rodell. 2022. “Groundwater depletion in California’s Central Valley accelerates during megadrought.” *Nature Communications* 13:7825.

Martindill, J.R., R.T. Good, and F.J. Loge. 2021. “Estimating agricultural groundwater withdrawals with energy data.” *Journal of Water Resources Planning and Management* 147:04021018.

Medellín-Azuara, J., R. Howitt, and J. Harou. 2012. “Predicting farmer responses to water pricing, rationing and subsidies assuming profit maximizing investment in irrigation technology.” *Agricultural water management* 108:73–82.

Mount, J., and E. Hanak. 2016. “Water use in California.” *Public Policy Institute of California, San Francisco, California* 2.

Newman, C., R.E. Howitt, and D. MacEwan. 2018. “How are Western water districts managing groundwater basins?” *California Agriculture* 72.

Olen, B., J. Wu, and C. Langpap. 2016. “Irrigation decisions for major west coast crops: water scarcity and climatic determinants.” *American Journal of Agricultural Economics* 98:254–275.

Pfeiffer, L., and C.Y.C. Lin. 2014. “Does efficient irrigation technology lead to reduced groundwater extraction? Empirical evidence.” *Journal of Environmental Economics and Management* 67(2):189–208.

Schaible, G.D., A.K. Mishra, D.M. Lambert, and G. Panterov. 2015. “Factors influencing environmental stewardship in US agriculture: Conservation program participants vs. non-participants.” *Land use policy* 46:125–141.

Shi, J., J. Wu, and B. Olen. 2022. “Impacts of climate and weather on irrigation technology adoption and agricultural water use in the US pacific northwest.” *Agricultural Economics*, pp. .

Skidmore, M., T. Andarge, and J. Foltz. 2023. “Effectiveness of local regulations on nonpoint source pollution: Evidence from Wisconsin dairy farms.” *American Journal of Agricultural Economics* 105:1333–1364.

Soil Conservation Service. 1993. *National Engineering Handbook*, USDA, Washington, D.C., chap. 2: Irrigation Water Requirements.

Sorrell, S., and J. Dimitropoulos. 2008. “The rebound effect: Microeconomic definitions, limitations and extensions.” *Ecological Economics* 65:636–649.

USDA. 2013, 2018. “Irrigation and Water Management Survey.”

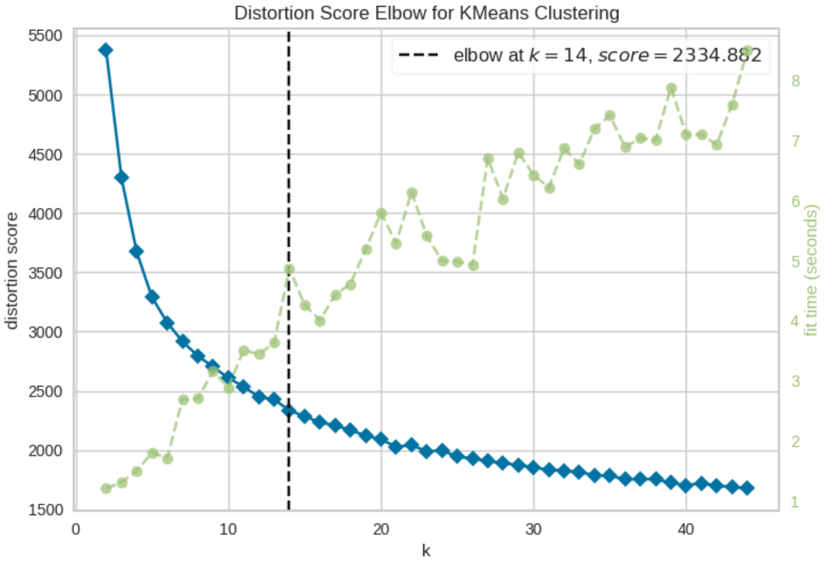
Walkinshaw, M., A. O'Geen, and D. Beaudette. 2023. "Soil properties." *Digital, 800 x 800 m pixel*, pp. , casoilresource.lawr.ucdavis.edu/soil-properties.

Wallander, S., M. Aillery, D. Hellerstein, and M. Hand. 2013. "The role of conservation programs in drought risk adaptation." *Economic Research Service ERR* 148.

Wallander, S., and M.S. Hand. 2011. "Measuring the impact of the Environmental Quality Incentives Program (EQIP) on irrigation efficiency and water conservation." Working paper.

Wilkening, E., D.M. Heeren, D. Hallum, J. Schellpeper, and D.L. Martin. 2021. "Impact of irrigation technologies on withdrawals and consumptive use of water." In *2021 ASABE Annual International Virtual Meeting*. American Society of Agricultural and Biological Engineers, p. 1.

Figure 15: Distortion Score Elbow for KMeans Clustering



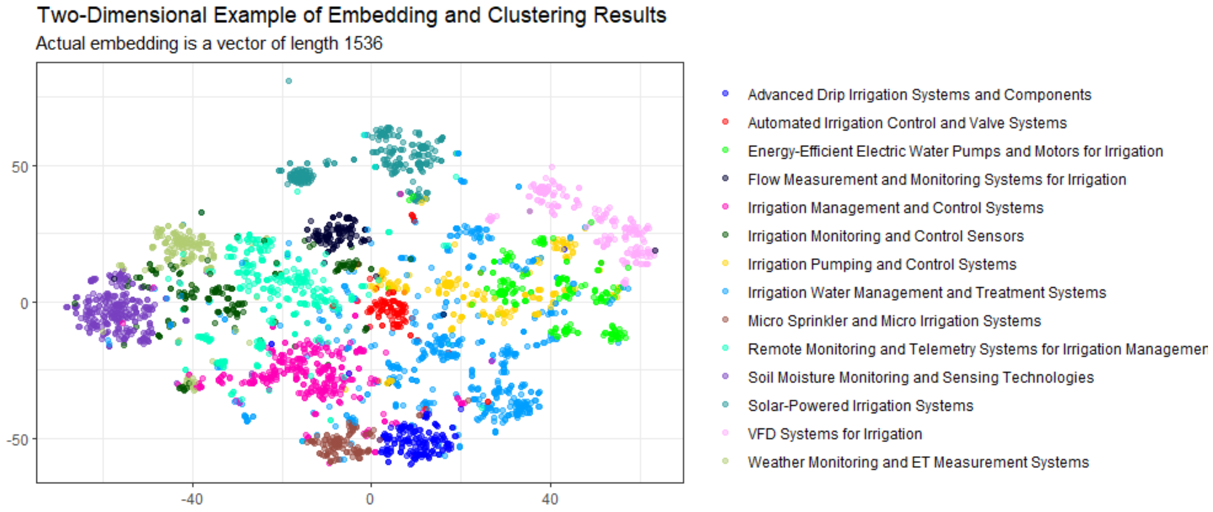
This figure presents the result of the K-Means clustering algorithm on the embedded technology descriptions.

Appendix A Project Technology Categories - Additional Tables and Figures

Figure 15 presents the results of the K-Means clustering algorithm. K-Means is a machine learning algorithm that clusters each data point based on K number of centroids. The centroid is the center of the cluster and each data point is assigned to a cluster by reducing the within-cluster sum of squares. I find the optimal number K using the “elbow” method. This method helps select the optimal number of clusters by plotting the within-cluster sum of squares to a range of possible K values. The “elbow”, or the point of inflection, on the curve indicates where the underlying model fits best. Figure 15 presents the distortion score elbow using the embedded technology descriptions. The number of optimal clusters estimated by the algorithm ranges from 14-18. I cluster the project components into 14 categories and assign common category names and descriptions using the GPT-4 LLM.

Table 8 presents the category names and descriptions resulting from the embedding and clustering exercise. Finally, Figure 16 illustrates a two-dimensional example of how the K-Means clustering model worked to group the project components into their respective categories. The true length of the embedded vectors is 1536, so the figure is an over-simplification of the true vector space. The true vector space is converted into two dimensions using T-distributed stochastic neighbor embedding,

Figure 16: Two-Dimensional Representation of Cluster Embeddings



This figure presents a two-dimensional representation of the embedding exercise and K-Means clustering algorithm. The clustering algorithm grouped technology descriptions and this figure illustrates the similarity of the embedded technology descriptions. Descriptions that were clustered together are plotted in different colors and it can be observed that descriptions that fell into the same technology category are close in space to one another.

this method converts similarities between data points to joint probabilities and minimizes the distance between the joint probabilities of the low-dimensional embedding and the high-dimensional data.[16] However, it provides a useful illustration of how closely related the technologies are that are ultimately clustered together.

Appendix B Data Cleaning and Construction Procedure

B.1 Fuzzy Matching of Assessor Parcel Numbers

Assessor Parcel Numbers (APNs) are fourteen digit codes that index parcels of land used by County Assessor Offices, generally in the form: 000-000-0000-0000. Since APN(s) are entered manually, frequently without necessary leading and trailing zeros, some do not match perfectly with the county APN file. In addition, APNs are not unique across counties, so both the APN and County is needed to obtain a unique combination. To address this problem, I used a machine learning algorithm that performs fuzzy matching to allow for imperfect matches. There are three key possibilities that occur that would result in non-matches: (1) parcel numbers are entered in the SWEEP application without

---

[16] T-distributed Stochastic Neighbor Embedding API Reference

Table 8: Irrigation Technology K-Means Clustered Categories

Category Name	Category Description
1. Irrigation Pumping and Control Systems	This category encompasses a wide range of technologies related to the operation, control, and efficiency of pumps used in irrigation systems. It includes various types of pumps, pump control and automation technologies, as well as systems for monitoring and improving pump performance and efficiency.
2. Irrigation Water Management and Treatment Systems	This category encompasses a wide range of technologies related to the storage, treatment, control, and efficient use of water in irrigation systems. It includes various types of filtration systems, water treatment equipment, pipelines, storage tanks, and efficiency-enhancing technologies. It also includes systems for managing and treating soil amendments and fertilizers.
3. Variable Frequency Drive (VFD) Systems for Irrigation	This category encompasses a wide range of technologies related to the control and efficiency of irrigation systems using variable frequency drives. It includes various types of VFD pumps, controllers, motors, and panels, as well as specific applications of VFD technology in different types of irrigation equipment.
4. Irrigation Monitoring and Control Sensors	This category encompasses a wide range of sensor technologies used for monitoring various parameters (like pressure, temperature, humidity, water level, etc.) and controlling irrigation systems. They include various types of sensors, transducers, dataloggers, and monitoring systems.
5. Solar-Powered Irrigation Systems	This category includes various types of solar systems used to power irrigation technologies. These technologies range from solar panels, photovoltaic systems, solar arrays, to solar-powered pumps and motors.
6. Automated Irrigation Control and Valve Systems	This category includes various types of valves, control systems, and automation technologies used in irrigation. These systems can vary in type (pressure release, reduction, compensating), control (remote, automated, electronic), power source (battery, solar), and specific applications (backflow prevention, pressure regulation, chemigation).
7. Weather Monitoring and Evapotranspiration (ET) Measurement Systems	This category includes various types of weather stations, sensors, and data collection systems used in irrigation. These systems can vary in type (on-farm, remote, web-based), data collected (temperature, humidity, ET), and specific applications (climate monitoring, soil monitoring, data transmission).
8. Irrigation Management and Control Systems	This category includes various types of irrigation systems, controllers, and scheduling tools that are used to manage and optimize the application of water in agricultural settings. These systems can vary in type (automated, remote, solar-powered), data used (ET, soil moisture), and specific applications (irrigation scheduling, water management).
9. Advanced Drip Irrigation Systems and Component	This category includes a wide range of technologies that are used to manage, monitor, control, and optimize the application of water in agricultural settings through drip irrigation.
10. Soil Moisture Monitoring and Sensing Technologies	This category includes a wide range of technologies that are used to measure, monitor, and control the moisture levels in the soil.
11. Micro Sprinkler and Micro Irrigation Systems	This category includes a variety of technologies designed for efficient and precise water distribution in agricultural settings. These systems are characterized by their ability to deliver water directly to the root zone of plants, minimizing evaporation and runoff. They include various types of sprinklers, sprayers, and nozzles, many of which are designed to operate under low pressure.
12. Remote Monitoring and Telemetry Systems for Irrigation Management	This category includes a variety of technologies designed for efficient and precise water and energy management in agricultural settings. These systems are characterized by their ability to collect, transmit, and analyze data remotely, enabling real-time monitoring and control of irrigation systems. They include various types of telemetry units, data loggers, automation controllers, and software for data management and processing.
13. Flow Measurement and Monitoring Systems for Irrigation	This category encompasses a range of technologies designed to measure, monitor, and control the flow of water in irrigation systems. These technologies include various types of flow meters, such as magnetic, digital, and pulse output flow meters, as well as systems for monitoring these flow meters remotely.
14. Energy-Efficient Electric Water Pumps and Motors	This category includes a variety of electrically powered pumps and motors designed for irrigation purposes. These technologies range from high-efficiency motors and pumps, electric turbines, submersible pumps, to variable speed pumps and motors.

(with) zeros or without (with) hyphens, when the parcel file APNs are entered with (without) zeros or with (without) hyphens; (2) the SWEEP applicant listed APNs in multiple counties, then searching for an APN-County match in the reported county may result in no match when in fact that APN is in a neighboring county, since it is not uncommon for agricultural operations to span across county borders; and (3) the applicant entered the APN incorrectly.

I construct an algorithm for the APN matching procedure to address these issues to obtain as many matches as possible. By default, the fuzzy matching algorithm in Python calculates a similarity score based on the Levenshtein distance between two strings and returns the five closest matches.[17] The basic steps of the algorithm are described below:

1. Step 1: Look for matches within the reported county and with zeros or hyphens. Any perfect matches, i.e., with a match score of 100, are removed at this step.
2. Step 2: Look for matches within the reported county with leading or trailing zeros and hyphens removed from the APN. Any perfect matches are removed at this step.
3. Step 3: Look for matches in any neighboring counties, i.e., counties that share a boundary, with any hyphens removed since some counties contain hyphens and some counties do not. Any perfect matches are removed at this step.
4. Step 4: Look for matches within neighboring counties with both leading or trailing zeros and hyphens removed. Any perfect matches are removed at this step.
5. Step 5: For the remaining imperfect matches, keep the highest scoring APN match. Throw out any matches that have a similarity score less than 80.

B.2 Asessor Parcel Boundary and Field Boundary Construction

I converted parcel boundaries into field boundaries using the CA SCM shapefile. Field boundaries in the CA SCM shapefile were developed to correspond more closely with cropped area as opposed to legal parcel boundaries and as such, should more accurately reflect homogeneously cropped areas. In each year that the Statewide Crop Mapping shapefile is available from Land IQ and DWR, I dis-aggregated each applicant-parcel observation to the applicant-field observation by overlaying the field boundary

---

[17] Python TheFuzz Package Documentation

shapefile on the parcel boundary shapefile. I maintained the field within the parcel intersection if is at least 20% of the original field area. However the results are not sensitive to this choice. For the years that the Statewide Crop Mapping shapefile is unavailable, 2015 and 2017, I use the field unions of the 2014 and 2018 or 2016 and 2019 files, respectively.

B.3 California Department of Water Resources Agricultural Land & Water Use Estimates

To map crops to their relative water intensities, I use California Department of Water Resources (DWR) Agricultural Land and Water Use Estimates. DWR provides estimates of irrigated crop acreages, crop evapotranspiration, evapotranspiration of applied water, effective precipitation, and applied water for 20 crop categories each year. To estimate water intensity, I use the estimated applied water per acre for each crop group. The applied water estimates incorporate adjustments for irrigation efficiencies and the amount of water required for cultural practices such as ponding of water in rice fields or extra water applied to leach accumulated salts from the soil. DWR publishes estimates at the state, hydrologic region, and county. I use estimates published at the county level to get the smallest study area available. Table 9 presents the breakdown of the 20 crop groups that DWR publishes estimates for. I aggregate all individual crops in my data into these 20 groups and assign the county level crop water intensity by group.

Two other adjustments were made to calculate county-level crop water intensities. First, DWR reports processing tomatoes and fresh market tomatoes separately however the Statewide Crop Mapping Data and the CDL do not report these categories separately. To calculate a single tomato water-intensity for each county I use the county-level irrigated crop acreage published by DWR Agricultural Land & Water Use Estimates 2016-2020.

The second adjustment is made to crops that are classified in the Statewide Crop Mapping dataset as young perennials. For any fields classified as young perennials If later in the sample period the classification changes from young perennial to a perennial crop, e.g., Almonds, then I change each of the years that the field is classified as young perennial to it's true crop classification and it's resulting water-intensity for that crop. If the field is classified as young perennial for the entire time series, then I assign an average crop water intensity that is equal to the weighted average of all perennial crops in that specific county, weighted by irrigated crop acreage published by DWR Agricultural Land & Water

Table 9: CA DWR Agricultural Land & Water Use Estimates: 20 Crop Group Definitions

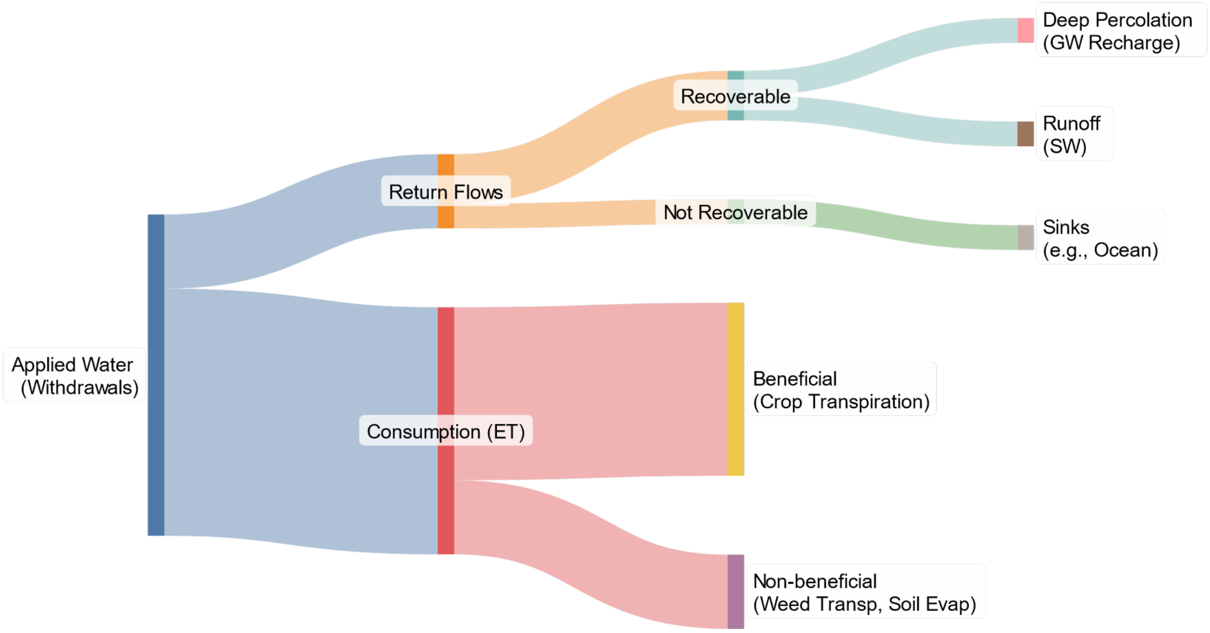
Crop Group	Crops
Grain	Wheat, winter wheat, spring wheat, barley, oats, miscellaneous grain and hay
Rice	Rice, wild rice, flooded rice, upland rice
Cotton	
Sugar beet	
Corn	
Dry beans	
Safflower	
Other field crops	Flax, hops, sorghum, sudan, castor beans, miscellaneous field crops, sunflower, millet, sugarcane
Alfalfa	
Pasture	Pasture, clover, miscellaneous grasses, rye
Tomato processing	
Tomato fresh	
Cucurbits	Melons, squash, cucumbers, watermelon
Onions and garlic	
Potatoes	Potatoes, sweet potatoes
Truck crops	Artichokes, asparagus, green beans, celery, carrots, lettuce, peas, spinach, bush berries, strawberries, peppers, broccoli, cabbage, cauliflower
Almond and pistachios	
Other diciduous	Apples, apricots, walnuts, cherries, peaches, nectarines, pears, plums, prunes, figs, kiwis
Citrus and subtropical	Grapefruit, lemons, oranges, dates, avocados, olives, jojoba
Vineyards	Table grapes, raisin grapes, wine grapes

Use Estimates 2016-2020.

B.4 Consumptive Water Use

Figure 17 provides a basic illustration of how water moves through the system in agriculture beginning with total water withdrawals, or total water application. A portion of applied water is consumed through evapotranspiration (ET) either through beneficial crop transpiration or non-beneficial weed transpiration and non-beneficial soil evaporation. The remaining water returns to the system through return flows that are either (1) recoverable through deep percolation to recharge groundwater aquifers or runoff to surface water to be used downstream or (2) non recoverable through sinks such as the ocean.

Figure 17: Diagram of Water Flows



This diagram illustrates a simplified flow of applied water on an agricultural field, separated between return flows and consumption (ET).

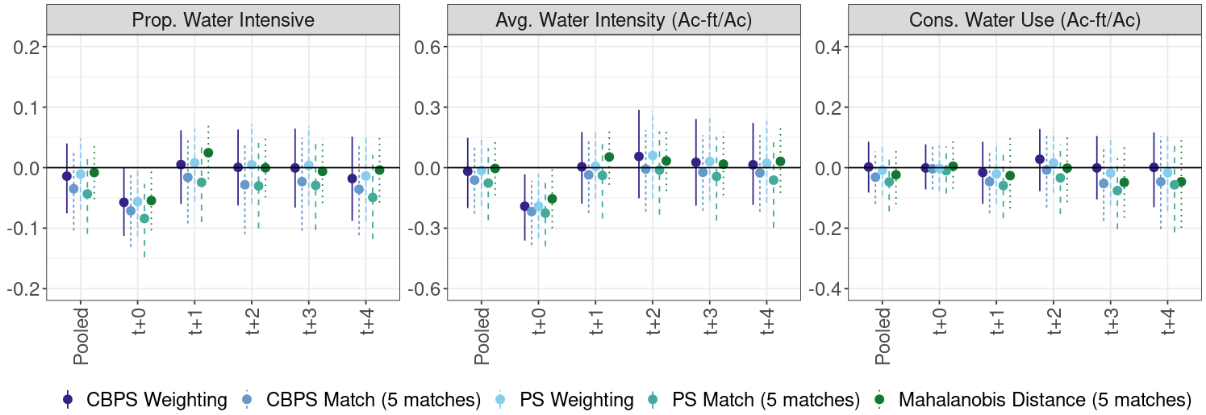
Appendix C Robustness Checks

The following section presents the results from a number of empirical robustness checks. I evaluate the sensitivity of the results to alternative matching and weighting methods in the main estimating procedure (Imai, Kim, and Wang, 2021) and the sensitivity of the results to an alternative estimator that is robust to staggered treatment timing and heterogeneous treatment effects (Callaway and Sant'Anna, 2021). I find that the main results are not sensitive to the matching and weighting method used and the results are not sensitive to using an alternative estimator.

C.1 Alternative Matching Procedures

The matched DiD estimator proposed by Imai, Kim, and Wang (2021) allows the researcher to choose which matching or weighting method to use to estimate the results. In the main results of this paper I use the Covariate Balancing Propensity Score (CBPS) weighting method (Imai and Ratkovic, 2014). The CBPS uses a single model to determine the treatment assignment mechanism and the covariate balancing weights, which estimates the propensity score while also optimizing the resulting covariate balance. This mitigates the effect of parametric propensity score model misspecification. I estimate all three main outcome variables using three alternative matching and weighting methods: (1) CBPS matching with five matches, (2) Mahalanobis distance measure with five matches, and (3) propensity score weighting. The results are presented in Figure 18. Each panel corresponds to a different method.

Figure 18: Alternative Matching and Weighting Methods



C.2 Alternative Estimators

To evaluate the robustness of my chosen estimator, which is a variation of a matched DiD estimator, I also run my main regression results using Callaway and Sant'Anna (2021) heteroskedasticity and staggered treatment timing robust DiD estimator. The main results are presented in the following figures. The results are consistent with those found using the Panel Match estimator.

Figure 19: Alternative Estimator: Callaway and Sant'Anna (2021) DiD

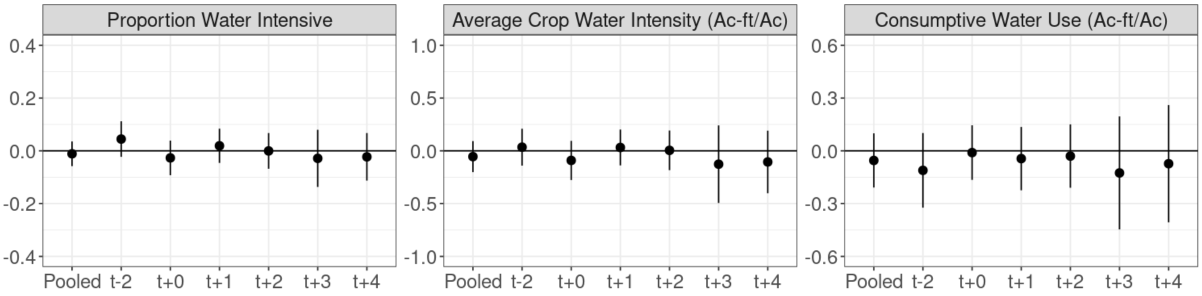


Figure 20: Alternative Estimator: Irrigation Water Source

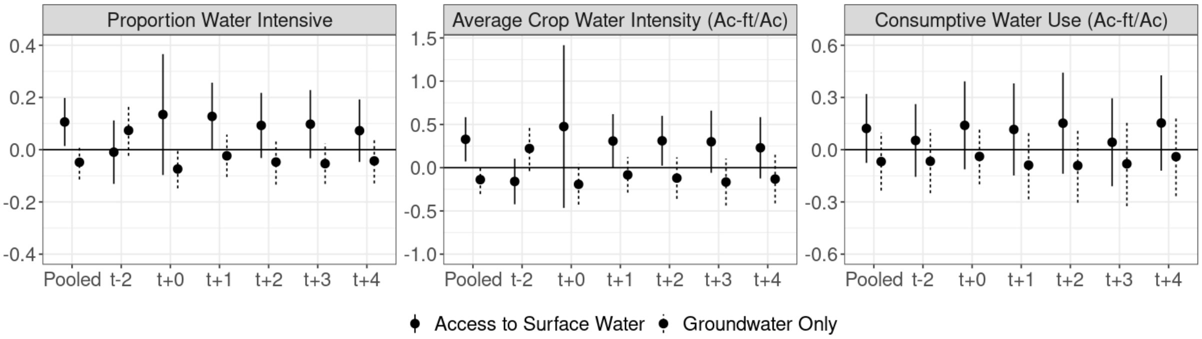


Figure 21: Alternative Estimator: Groundwater Regulatory Strength

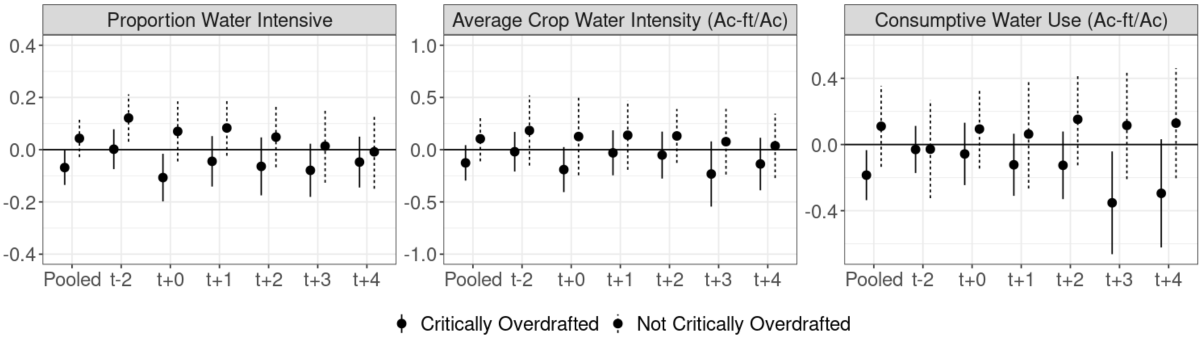


Figure 22: Alternative Estimator: Micro Irrigation Technologies

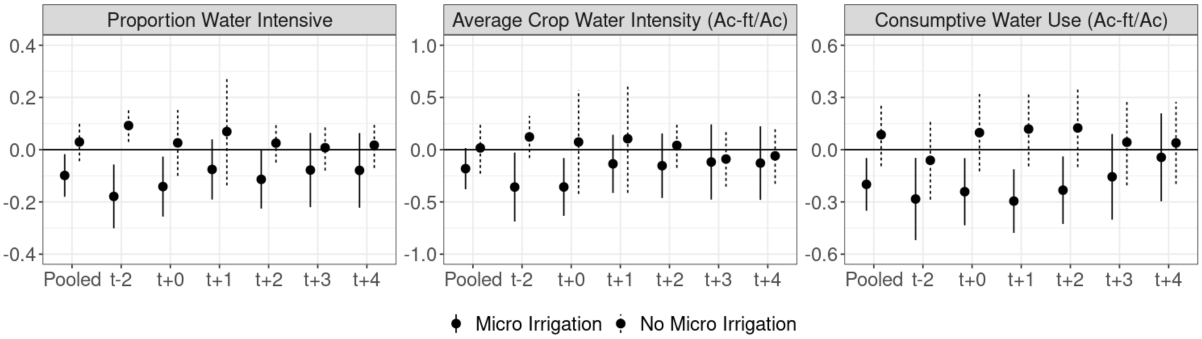


Figure 23: Alternative Estimator: Irrigation Pump Technologies