ESSAYS ON THE ECONOMICS OF WATER SCARCITY

By

Rania Lachhab

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food, and Resource Economics—Doctor of Philosophy

2024

ABSTRACT

Understanding how to manage water scarcity effectively is critical for current and future generations, as competition for this vital resource is only expected to intensify. This dissertation explores two key aspects of water scarcity: the potential of cloud seeding as a tool to increase water availability and the decision-making processes of smallholder farmers facing water constraints.

In the first chapter, I examine the economic impact of cloud seeding programs on both target and downwind areas, through a theoretical framework of welfare economics and spatial externalities. I use state-of-the-art causal inference methods to estimate the effects of 12 long-running cloud seeding projects in California on precipitation. I find that cloud seeding increases monthly precipitation totals by 57% in the target areas, with significant differences in effectiveness between projects, ranging from 30% to 140%. The additional surface water costs between \$3.7 to \$24.5 per ac-ft, depending on the project. Perhaps more importantly, I find that cloud seeding has a significant negative effect on downwind areas. I estimate an 8% decrease in monthly precipitation, generating a possible loss of \$49 million in surface water for Nevada. Results provide guidance on the contexts in which cloud seeding programs may be viable, along with estimating potential deleterious impacts to neighboring communities.

In the second chapter, I conduct a case study of cloud seeding in Santa Barbara County, California. This is the first paper that examines the effects of cloud seeding methods separately. Utilizing high-resolution data, I show that cloud seeding operations increase precipitation by 26-49% per operation, with an average cost of \$6.5 per acre-foot. I further differentiate between ground-based and aerial seeding methods, providing valuable insights into their relative effectiveness within the same program. This underscores the importance of operational details when considering cloud seeding as a water management strategy.

In the third chapter, I shift focus to agricultural decision-making in developing countries facing resource constraints. I examine the effects of agricultural extension and irrigation schemes on the joint choices of irrigation adoption and crop selection by smallholder farmers in Ethiopia. I develop a theoretical framework that incorporates both profit maximization and subsistence needs. Using

household survey data, I find that subsistence needs significantly influence crop choice, with farmers prioritizing crops for household consumption. While extension services encourage cash crops, they may discourage riskier options like fruits. Irrigation adoption is limited, despite its potential to mitigate rainfall variability. These findings highlight the complex decision-making processes of smallholder farmers in water-scarce environments and emphasize the need for policies that consider both food security and economic opportunities.

This dissertation is dedicated to Mama and Baba.
Your countless sacrifices made my success possible, and your unwavering love and support have been the foundation of everything I have achieved.

ACKNOWLEDGEMENTS

Completing this dissertation has been both a rewarding and challenging journey, and I am incredibly grateful to the many individuals who supported me along the way.

First and foremost, I would like to express my deepest gratitude to Dr. Molly Sears. Her guidance and support were instrumental in shaping this dissertation and seeing it through to completion. Her expertise, insightful feedback, and genuine care for my success have been invaluable. Beyond her excellent academic guidance, Molly also extended her kindness and warmth by welcoming me into her home and helping me through many personal challenges.

I would like to extend my sincere appreciation to Dr. Scott Swinton, who served as my advisor and my biggest advocate. His meticulous guidance inspired me to strive for excellence. I fondly remember our walks in the park during the pandemic, which provided much needed breaks and insightful conversations.

I am grateful to Dr. Rick Horan and Dr. Mark Skidmore for their excellent feedback and insights throughout the dissertation process. Their expertise significantly improved the quality of my work. I am grateful to Dr. Matthew Gammans, who served as my co-advisor throughout my graduate studies. I am also deeply indebted to Dr. Rachid Doukkali, who instilled in me a passion for research and economics. His guidance steered me onto this path, and I would not have been here without him.

A special thank you goes to Thomas Keene, my best friend since meeting him in math camp during our first week of graduate school. Thomas and his parents, Dave and Erin, became my family and a constant source of strength. I would not have survived the highs and lows of this journey without this precious friendship.

My heartfelt thanks go to my friends and fellow graduate students in the department. Their camaraderie made this journey not only academically enriching, but also personally fulfilling.

Most importantly, I am incredibly grateful to my parents and my brother whom unconditional love and support served as a constant anchor throughout this journey.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

U.S.	United States
NOAA	National Oceanic and Atmospheric Administration
PRISM	Parameter-elevation Relationships on Independent Slopes Model
HUC	Hydrologic Unit Code
NAWMC	North American Weather Modification Council
CPR	Common-Pool Resources
USGS	United States Geological Survey
WGISC	Wyoming Geographic Information Sciences Center
DiD	Difference-in-Differences
ATT	Average Treatment effect on the Treated
TWFE	Two-Way Fixed Effects
QMLE	Quasi-Maximum Likelihood Estimation
LEF	Linear Exponential Family
BCA	Benefit-Cost Analysis
CEC	California Energy Commission
NAWC	North American Weather Consultants, Incorporated
AHOGS	Automated High Output Ground Seeding
WMA	Weather Modification Association
SBCWA	Santa Barbara County Water Agency
SDiD	Synthetic Difference-in-Differences
ICE	Ice Crystal Engineering
HCF	Hundred Cubic Feet
LSMS	Living Standards and Measurement Study
ISA	Integrated Survey on Agriculture

GDP	Gross Domestic Product
FAO	Food and Agriculture Organization
ADF	African Development Bank
DA	Development Agent
NL	Nested Logit model
MNL	Multinomial Logit model
IIA	Independence of Irrelevant Alternatives
GEV	Generalized Extreme Value
IV	Inclusive Value
PWI	Potential Wetness Index
CSA	Ethiopian Central Statistical Agency

CHAPTER 1

SILVER BULLETS: CLOUD SEEDING AND WATER RESOURCES IN CALIFORNIA

1.1 Introduction

Climate change is a pressing global issue that has devastating impacts on the environment, agricultural production, and water resources. One significant consequence of climate change is an increase in water scarcity in arid and semi-arid regions, exacerbated by rising temperatures and increased variability in precipitation patterns. This has led to a growing need for efficient water resources management strategies. One potential solution to mitigate the impacts of water scarcity from the supply side is to implement precipitation enhancement programs, which aim to increase the amount of rainfall using cloud seeding over a specific target area.

Cloud seeding involves injecting particles into clouds, encouraging microscopic water droplets to form snowflakes large enough to exit clouds as precipitation. While this is a commonly adopted practice worldwide, the effects of cloud seeding on increasing water supply and managing water scarcity are unclear. Although early studies of cloud seeding made progress in understanding the conditions under which it could enhance precipitation, they were unable to determine the magnitude of that enhancement. The National Research Council stated that "there is still no convincing scientific proof of the efficacy of intentional weather modification efforts" (Research Council et al., 2004). Some cloud seeding projects have been halted due to the great uncertainty surrounding their effectiveness. For example, the Israel Water Authority decided to stop a 38-year-long operational seeding program after updated experimental results suggesting only a statistically insignificant 1.8% increase in precipitation. This randomized experiment was conducted from 2013 to 2020 to reassess the effect of cloud seeding on rainfall and provide a basis for evaluating its utility, after Levin et al. (2010) suggested that cloud seeding has been ineffective in Israel. Despite the uncertainty, weather modification has continued in many arid regions in an effort to meet the increasing demand for water and due to the large potential cost benefit of producing additional water by cloud seeding, especially compared to the cost of other water supply management techniques such as groundwater banking and desalination plants (Rauber et al., 2019).

More broadly, cloud seeding is part of a large suite of approaches commonly used to address water scarcity and ensure sustainable water management. These can generally be summarized in two categories: (1) increasing water supply and (2) managing water demand. Freshwater supply can be increased by investing in water infrastructure, such as dams, reservoirs, groundwater banking, desalination plants, and recycling systems for the treatment and reuse of wastewater. Managing water demand consists of integrated management of water resources through comprehensive planning, monitoring, and governance mechanisms to optimize water allocation across different users. It also involves the implementation of water conservation measures, such as promoting efficient irrigation practices and encouraging water-saving technologies. The current economic literature has explored many of these topics in significant detail, but cloud seeding has been largely ignored in the economic literature.

Estimating the effect of cloud seeding programs on precipitation and runoff is relevant to economists and policy makers as it provides valuable insights into the economic implications of these programs, allowing decision making related to resource allocation and water resource management. By estimating the impact of these programs on precipitation, economists can help identify and assess the most effective strategies for addressing water scarcity and mitigating the impacts of climate change on water resources. However, little research exists on the economic impacts of cloud seeding or on designing cloud seeding applications that reduce negative externalities to neighboring agricultural and/or urban areas upon seeding. More research exists on the meteorology and technical feasibility of cloud seeding projects (Boyle et al., 2006; Griffith et al., 2005; DeFelice et al., 2014; Friedrich et al., 2020; Griffith et al., 2015). The increased interest of water management agencies in precipitation enhancement opportunities has not been met with more economic research on the topic, creating a gap in the literature on the economic impacts of cloud seeding (Rauber et al., 2019). The empirical application in this paper focuses on California, where cloud seeding has been used since the 1950s.

While experimental results are important to understand the feasibility of cloud seeding as a technology to increase precipitation, they have many shortcomings in informing policymakers on the effectiveness of the implementation of cloud seeding as a long-term water management strategy. Direct observational evidence for the cloud seeding effects is limited, and statistical proof of its effectiveness is lacking (Malik et al., 2018), since experiments are limited in time and space, and the operational stage of projects is too costly to allow for randomization of seeding events. Therefore, while experiment-based research exists on the meteorology and technical feasibility of cloud seeding projects (Boyle et al., 2006; Griffith et al., 2005; DeFelice et al., 2014; Friedrich et al., 2020; Griffith et al., 2015), little research exists on the economic impacts of cloud seeding or in designing cloud seeding applications that reduce negative externalities on neighboring agricultural and/or urban areas. The increased interest of water management agencies in precipitation enhancement opportunities has not been met with more economic research on the topic, creating a gap in the literature about the economic impacts of cloud seeding (Rauber et al., 2019). Analyzing the impact of cloud seeding programs on precipitation and runoff is particularly relevant for economists and policymakers, as it offers valuable insights into the economic implications of such programs. This analysis enables informed decision-making regarding resource allocation and the management of water resources.

California provides a context where a careful analysis of cloud seeding programs has the opportunity to provide valuable implications to arid regions worldwide. California is the world's fourth largest economy and the most agriculturally productive state in the United States, with a heavy reliance on water and a susceptibility to drought. Considered a "Mediterranean climate", most precipitation occurs during the winter months. In northern California, precipitation falls as snow over higher elevations along coastal ranges, and at nearly all elevations over interior mountain ranges (Rauber et al., 2019). In southern coastal California, precipitation falls mostly as rain or snow that melts before reaching the ground. The ensuing snowmelt and runoff in spring and summer provide annual water supplies. An opportunity to increase snowpack and water storage in targeted areas could have a dramatic impact on productivity and water sustainability in California and other drought-prone regions.

In this paper, I conduct a comprehensive analysis of 12 precipitation enhancement programs in California. I leverage Public Law 92-205 (1972) that requires all non-Federal weather modification activities to be reported to the United States (U.S.) Secretary of Commerce, via the National Oceanic and Atmospheric Administration (NOAA) Weather Program Office, to retrieve monthly cloud seeding operations data. The reports provide data on the number of cloud seeding events carried out each month for precipitation augmentation purposes, as well as the total quantity of seeding agent used each month. Monthly precipitation totals are obtained from PRISM[1], for target and control areas defined as eight-digit Hydrologic Unit Codes (HUC8).

The results suggest that in general, weather modification programs in California increase precipitation in the target areas by 57% on average per seeding event, thus producing additional water supply at a cost of $7 per acre-foot. The project-specific results show significant heterogeneity between the effectiveness of cloud seeding in different project areas. Estimates range between 22% and 140% increases in precipitation in target areas. However, the analysis of downwind effects shows that cloud seeding decreases precipitation by 8% in the leeway of the Sierra Mountains.

This paper provides a step forward in evaluating the effectiveness of cloud seeding in increasing water supply and managing water scarcity. My work fills a gap in the economic literature by evaluating the heterogeneity in the costs and benefits of cloud seeding programs which have the potential to be considered a more feasible option than other water management strategies. It also provides value to the scientific literature by providing a credible causal analysis with a higher resolution time scale than previous studies, which focus on annual precipitation at an individual water basin. Finally, this paper provides one of the first estimates of downwind effects of cloud seeding on neighboring areas. The results of this study have important implications for policymakers and water management agencies in developing effective strategies for addressing water scarcity in arid regions.

The remainder of the manuscript is organized as follows. Section 2 briefly overviews cloud seeding in the U.S. and research efforts surrounding the effectiveness of the technique. Section 3 presents the theoretical framework, followed by the data and empirical strategy in Section 4. The results are summarized in Section 5, followed by a discussion on the implications of the results in

[1]PRISM stands for Parameter-elevation Regressions on Independent Slopes Model.

Section 6.

1.2 Background on Cloud Seeding

The science underlying cloud seeding as a weather modification technique has been extensively studied and used to implement multiple projects across the world. Shortly after the pioneering discoveries of Schaefer (1946) and Vonnegut (1947) about cloud seeding, both scientists and water resource managers began exploring the potential of seeding wintertime orographic cloud systems to increase water supplies in arid and semi-arid regions. The process of cloud seeding typically involves injecting silver iodide into the clouds to artificially stimulate them to produce more rainfall or snowfall than they would produce naturally. This additional precipitation could be used to increase the amount of water available in the targeted areas.

Cloud seeding is designed to "build" snowflakes by aggregating microscopic drops of water vapor that would otherwise not have enough mass to fall to the ground. This process is explained in Figure 1A.1. In cold conditions, especially at temperatures cooler than -39 degrees Celsius, water vapor cannot form snowflakes on its own, requiring condensation nuclei to exist. Condensation nuclei are microscopic particles of dust in the atmosphere. Ice crystals and water vapor attach to these particles, which then accumulate enough mass that they fall to the ground. When silver iodide is injected into clouds via aerial or ground flares, the tiny silver iodide particles become condensation nuclei. Cloud seeding is thought to be effective in cold weather conditions, and in the convective bands of clouds that have a lot of water but not enough condensation nuclei.

There have been several randomized experimental projects on cloud seeding in winter environments with the goal of increasing rainfall, using ground-based seeding with Silver Iodide (AgI). These include Santa Barbara I and II (Neyman et al., 1960; Bradley et al., 1978), Tasmania I, II, and III (Ryan and King, 1997; Morrison et al., 2010), and Israel I,II, and IV (Gagin and Neumann, 1981; Benjamini et al., 2023). A comprehensive review of these experiments reveals inconsistent statistical results across the projects, as well as faults in the experimental designs and erroneous statistical methods (Rauber et al., 2019). In an effort to improve experimental design and statistical analysis to study natural and seeded cloud systems and to verify basic components of the hypothesis (Gabriel, 2000; Manton and Warren, 2011; Breed et al., 2014) where programs typically do not employ randomized seeding techniques, the approach to evaluating the effectiveness of cloud seeding has been to use a target-control approach where one basin is seeded throughout the winter season, while a neighboring basin remains unseeded as the control.

One major concern of cloud seeding programs is their effect on downwind areas. The term "rainshadow" refers to the area on the leeward side of a topographic barrier where precipitation is less than on the windward side. Although a rainshadow effect is often a naturally occurring phenomenon, such as the rainshadow of the Sierra Nevada in California, in theory weather modification efforts to increase precipitation in one area could create a rainshadow downwind of that area even in the absence of an orographic barrier. Such efforts could also intensify a naturally occurring rainshadow effect. This theory rests on the assumption that, all other factors remaining constant, artificial removal of atmospheric moisture from a cloud or cloud system results in less atmospheric moisture available for precipitation as the cloud or cloud system moves downwind (Bulman and Shelton, 1991).

1.2.1 Cloud Seeding Programs in the United States

The first discoveries pertaining to the formation of ice crystals in cooled clouds are credited to Irving Langmuir, Vincent Schaefer, and Bernard Vonnegut, in 1946. Rogers and Gahan (2009) explain that according to Secretary of the Interior Stewart Udall, in a speech delivered before the American Meteorological Society, the Bureau of Reclamation took interest in weather modification in 1947, a year after Schaefer's seeding demonstrations. While the Bureau of Reclamation did not take part in cloud seeding experiments undertaken by General Electric, other government entities such as the Army Signal Corps, the Air Force, and the Navy coordinated efforts with Langmuir to initiate Project Cirrus. This early cloud seeding project produced demonstrations and statistical information that were used later by other public and private entities. Experimental phases were initiated in several arid places across the U.S., especially in the western states. In March 2023, the U.S. Bureau of Reclamation offered a $2.4 million grant to the Southern Nevada Water Authority to fund cloud seeding in the Upper Colorado River Basin, affecting numerous Western states. As cloud seeding became more popular, the North American Weather Modification Council (NAWMC) reports operational projects in the states of California, Colorado, Idaho, Nevada, Utah, Wyoming, Kansas, North Dakota, and Texas.

Cloud seeding has been conducted in California since the 1950s, one of the longest records of operational weather modification anywhere in the world. California's operational programs included the Southern Sierra Nevada, Santa Clara Valley, San Joaquin River Basin, San Luis Obispo County, and Santa Barbara County. The earliest program was at the Bishop Creek watershed in the eastern Sierras in 1948. Santa Barbara County first conducted cloud seeding trials in the 1950s, and experiments were introduced in the 1970s to determine if cloud seeding was beneficial for the region. Other programs have been implemented in the following years, covering additional watersheds in California.

While the economic literature on weather modification is scarce, meteorologists have contributed many estimates to the scientific literature on cloud seeding. However, the empirical methods used in this literature fail to provide an explanation for the differences in precipitation trends in areas targeted by cloud seeding, beyond simple correlation. Griffith et al. (2015) used linear regression equations to predict the amount of target area precipitation in the Santa Barbara project, based on precipitation observed in the control area for a historical period before the start of the project. They applied these linear regression equations to estimate what the target area precipitation should have been without seeding, after the start of the project. They used these estimates to compare the predicted seasonal target precipitation and the actual observed precipitation while cloud seeding is operational and attribute the differences to seeding activities. Silverman (2010) used a Monte Carlo permutation analysis of the regression ratio test statistic to compare water-year streamflow data from streamflow gauging stations inside basins that were targeted by cloud seeding with control stations. He attributed the statistical differences to the operational cloud seeding projects in the Sierra Nevada.

The same method was used by Silverman (2007), Silverman (2008), and Silverman (2009) to evaluate the effectiveness of the cloud seeding on streamflow for targets in the Kings River, Kern, and San Joaquin watersheds. While these previous studies have attempted to assess the effects of cloud seeding, their methodologies often have limitations, making it difficult to definitively establish causality between seeding activities and increased precipitation. Methods like those used by Griffith et al. (2015) rely heavily on the assumption that the historical relationship between target and control areas in terms of precipitation remains consistent even after cloud seeding operations begin. This is problematic as various factors, including natural climate variability and other environmental changes, can disrupt these historical patterns. Any change in precipitation cannot be solely attributed to seeding activities, leading to potentially inaccurate estimates, or overemphasizing the effects of cloud seeding. Similarly, approaches like those used by Silverman (2010) using permutation analysis focus on finding statistical differences in streamflow between targeted watersheds and control watersheds. However, observed differences could result from a multitude of factors beyond cloud seeding, including differing weather patterns, geological characteristics, or variations in vegetation cover. Without directly controlling for these factors, it remains difficult to isolate the true causal impact of cloud seeding on streamflow. The fundamental issue with these methods is their vulnerability to "omitted variable bias." This bias occurs when unobserved or unaccounted-for factors influence both cloud seeding operations and the outcomes of interest (precipitation or streamflow). If these omitted variables are not isolated, their effects can be falsely attributed to cloud seeding, undermining the ability to establish a true causal relationship.

1.2.2 Cloud Seeding Decision Criteria

To examine the effectiveness of cloud seeding, understanding the decision-making process behind when, where, and how clouds are seeded is crucial. While the final decision to seed clouds rests with the meteorologist on duty, individual differences between meteorologists are not a significant factor. This is because cloud seeding is not an arbitrary choice; it follows a strict set of pre-defined criteria. These criteria, encompassing factors like atmospheric temperature, wind patterns, and cloud characteristics, are well-established and documented in official project materials and contracts. They are similar across projects in the Sierra Nevada mountain range because these projects target a specific type of cloud formations, orographic clouds, which form due to the interaction of airflow and topography, specifically mountains and hills. This standardized approach ensures uniform decision-making across different meteorologists and projects, regardless of individual preferences or interpretations. Essentially, the criteria act as a decision-making framework, guiding the meteorologist on duty to objectively assess whether a particular cloud system is suitable for seeding. This consistency in approach helps to minimize the influence of individual bias and ensures that cloud seeding decisions are based solely on objective scientific principles.

The effectiveness of cloud seeding is inherently linked to the presence of suitable cloud formations. The decision to treat clouds for precipitation enhancement hinges on specific atmospheric conditions that favor successful ice nucleation and subsequent precipitation development. The ideal targets for cloud seeding are convective bands, which are elongated regions of rising air often associated with winter storms and cold fronts. These bands typically move from west to east and are oriented in a north-south direction. Their presence is a crucial first step in determining the suitability for cloud seeding. Several key atmospheric factors influence the decision to proceed with cloud seeding. Notably, the effectiveness of silver iodide, the most common seeding agent, is significantly impacted by temperature. Seeding is generally only attempted when the temperature at 700 mb (approximately 10,000 feet) is colder than -4°C to -5°C. Warmer temperatures might necessitate aircraft seeding for material delivery to colder regions aloft. The height of the -5°C level relative to ground-based seeding sites plays a significant role in the speed of the seeding effect. The closer this level is to the ground, the quicker the seeding material can be transported by convective elements within the bands to colder regions, where it can influence precipitation formation. The second factor that influences the decision to proceed with cloud seeding is wind, as favorable wind directions at different atmospheric levels are crucial for transporting the seeding material over the target area within the convective band and ensuring its interaction with the clouds. Both 700 mb wind direction and wind direction at 850 mb are evaluated to ensure effective material transport. The decision to seed also considers factors beyond the listed criteria. Suspension criteria, such as excessively strong winds or precipitation exceeding specific thresholds, might necessitate pausing seeding operations to ensure safety and avoid unintended consequences.

1.3 Theoretical Framework

1.3.1 Welfare Economics, Externalities, and Cloud Seeding

This analysis draws upon the theoretical foundations of neoclassical welfare economics and the theory of externalities to examine the economic implications of cloud seeding for water resource management. Pareto optimality, a cornerstone of welfare economics, serves as a normative benchmark for assessing the desirability of various economic outcomes. It posits that an allocation is Pareto optimal if there is no alternative allocation that can improve the well-being of at least one individual without making someone else worse off. The first and second fundamental theorems of welfare economics further emphasize the potential for markets to achieve Pareto optimality under specific conditions. However, in the presence of externalities, where the actions of one agent generate unintended costs or benefits for others, achieving Pareto optimality becomes more complex, as it requires a trade-off between the negative impact on agents and the productive benefits of the externality. Cloud seeding presents a case of externality, as its impact on precipitation patterns can have both positive and negative consequences for different agents.

From a positive perspective, this research aims to understand how existing institutional arrangements influence the decision-making of self-interested agents, potentially leading to unintended consequences for others. For instance, the economic benefits of cloud seeding for one region might come at the expense of reduced precipitation in another. Conversely, a normative lens seeks to identify policy interventions that mitigate these negative externalities and promote socially desirable outcomes. In either case, the core concern lies in the potential misallocation of resources, which impacts the overall level of well-being that members of society can obtain. Therefore, this research situates itself within the realm of welfare economics and the theory of externalities, specifically focusing on the market's failure to properly account for the environmental ramifications of economic activities like cloud seeding. Additionally, the study deals with the behavioral interactions between humans and the natural environment, establishing connections to broader fields of environmental science and public policy.

Building upon this theoretical foundation, the following section will argue that clouds constitute a common-pool resource and that cloud seeding activities generate unidirectional spatial externalities. This framework will guide the subsequent analysis of the economic implications of cloud seeding for water resource management.

1.3.2 Cloud Seeding as a Common-Pool Resource

From an economics perspective, natural resources are factors of production, akin to inputs that, when combined with labor and other forms of capital, produce goods and services. The accurate valuation of natural resources, as well as their products, hinges on the establishment of well-defined property rights. A property right is a bundle of characteristics that convey certain powers to the owner of the right. Natural resources are subject to various types of property rights. One important distinction between the types of rights that apply to natural resources is exclusivity. Using this characteristic, property rights can be put into two categories: private and common. A private property right gives the holder the power to the exclusive use of a natural resource, without having to share it with any other potential users. Conversely, a common property right is nonexclusive, which means that no agent can prevent another agent from using the natural resource and appropriating a share of the returns from the resource.

Common-pool resources (CPR) refer to natural or human-made resources that are collectively owned or used by a group of individuals. These resources are characterized by two key attributes: non-excludable and rivalry. Non-excludability implies that it is difficult to prevent individuals from accessing or using the resource. The feasibility, whether legal or economical, of excluding or restricting the use of resources by potential beneficiaries depends on both the physical characteristics of said resources and the institutions that exist in a specific jurisdiction. Rivalry, on the other hand, refers to the degree of subtractability of the benefit consumed by one individual from those available to others. This second attribute implies that one person's use diminishes the availability of the resource for others. Examples of common-pool resources include fisheries, grazing lands, water bodies, and the atmosphere. Clouds and rain can be considered common-pool resources within the broader context of the atmosphere, due to their characteristics of non-excludability and rivalry. In this particular context, non-excludability means that individuals cannot be easily excluded from using or benefiting from the atmosphere, and rivalry implies that one person's use of the atmosphere affects its availability for others.

Plott and Meyer (1975) called the process of withdrawing units of a resource "appropriation". Ostrom et al. (1994) followed by using the term "appropriator" to refer to all individuals who withdraw or appropriate resource units from a common-pool resource. This definition helps identify two categories of externalities faced by users of a common-pool resource: appropriation and provision. A common-pool resource can be thought of as a facility that creates the conditions for the existence of a stock of resource units such that this stock makes available a flow of resource units over time that are appropriable and subtractable in use. In appropriation problems, the flow aspect of the common-pool resource is what is problematic, whereas provision problems exist when the resource stock of the common-pool resource is problematic. Using this terminology, one can define the atmosphere as a common-pool resource and clouds as its resource units. Clouds, which are visible masses of water droplets or ice crystals suspended in the atmosphere, can also be considered a facility. Thus, water droplets are resource units of a cloud. The water droplets are subtractable from the common pool resource, the cloud.

Given the aforementioned definitions, cloud seeding, a technique employed to manipulate precipitation patterns, presents a unique case of externality generation within the context of common-pool resources. While the intended outcome of cloud seeding is often localized precipitation enhancement for a specific region, the atmospheric processes involved are inherently non-excludable and rivalrous. Non-excludability implies that the impacts of cloud seeding extend beyond the targeted area, influencing weather patterns and precipitation across vast geographical scales. This characteristic is evident as the consequences of altered precipitation patterns, such as droughts or floods, may be felt by communities far removed from the location where cloud seeding activities took place. Additionally, cloud seeding activities exhibit rivalry, meaning that increasing precipitation in one region can potentially reduce precipitation in another. This competition for a shared resource highlights the potential for tragedy of the commons scenarios, where individual actors, seeking to maximize their own benefits through cloud seeding, can inadvertently harm others by depleting the available moisture or disrupting established weather patterns.

It is crucial to acknowledge the uncertainty surrounding the production function of cloud seeding, meaning the clear link between effort/investment and the degree of precipitation enhancement remains unclear in existing literature. This ambiguity regarding the effectiveness and potential consequences of cloud seeding necessitates further research and cautious consideration before widespread implementation. The present essay aims to address these knowledge gaps by answering many critical questions about the effectiveness, the efficiency, and the externalities of cloud seeding.

1.3.3 The Spatial Externalities of Cloud Seeding

The utilization of common-pool resources often generates externalities, which are unintended and often unaccounted for effects of an economic activity on third parties. An externality exists when agent A's utility or production function depends directly on real variables chosen by another agent B, without an offer of compensation or other attention given to the effect on A's well-being (Baumol and Oates, 1988). One common externality associated with CPR is the problem of overuse, known as the tragedy of the commons, driven by individual actors seeking to maximize personal gain. Open access to a CPR incentivizes individuals to exploit the resource, often to the detriment of others. This behavior can lead to depletion or degradation of the resource, causing negative externalities such as overfishing, deforestation, or water pollution.

Cloud seeding involves the introduction of substances like silver iodide into clouds to stimulate precipitation, either as rain or snow. While intended to address water scarcity or agricultural needs, cloud seeding activities can generate unintended consequences for downwind communities. These spatial unidirectional externalities arise due to the inherent characteristics of the atmosphere. As air masses and precipitation patterns move, downwind areas may experience changes in precipitation levels, leading to both potential benefits and drawbacks. Downwind regions might receive more rainfall, increasing water availability for agriculture and other purposes. Conversely, excessive rainfall or unintended concentration of precipitation due to successful cloud seeding can result in flooding, infrastructure damage, and disruptions to agricultural activities. Another potential consequence could be the decrease in available rain to the downwind areas, which would exacerbate aridity conditions. The unidirectional flow of atmospheric processes makes cloud seeding a source of unidirectional spatial externalities.

The usual presumption is environmental policy is that solutions to environmental problems require legislative or executive rule making. However, within environmental economics, two distinct paradigms offer contrasting perspectives on addressing externalities: the Coase Theorem (Coase, 1960) and the Pigouvian paradigm (Pigou, 1920). The Coase Theorem proposes that, under certain conditions, efficient outcomes can be achieved even in the presence of externalities, via negotiation, and without government intervention. These conditions include (1) clearly defined property rights, such that the ownership and usage rights of the natural resource in question must be clearly established, (2) zero transaction costs, meaning that negotiation and bargaining between the involved parties must be costless and frictionless, and (3) rational actors, such that all parties involved must act rationally and in their own best interests. If these conditions are met, the theorem suggests that the parties will negotiate an agreement that maximizes the total social welfare, regardless of the initial allocation of property rights. This negotiated outcome is considered efficient because it allocates the resource in a way that maximizes the combined benefits for all parties involved.

However, the feasibility of the Coasean approach in real-world scenarios is often limited due to several challenges. Namely, negotiation and enforcement of agreements can be time-consuming, expensive, and complex. Furthermore, assigning clear property rights and accurately measuring the costs and benefits of externalities can be challenging, especially if numerous stakeholders with diverse interests are involved in the negotiation process. Therefore, while the Coase Theorem offers a valuable theoretical framework, its practical application often faces significant hurdles. In contrast, the Pigouvian paradigm advocates for government intervention through taxes or subsidies to address externalities. This approach aims to internalize the external costs associated with an activity by imposing a tax on the resource-depleting party. This tax, equal to the marginal social cost of the externality, incentivizes the party to reduce its activity to a level where the social costs and benefits are balanced. Choosing the most appropriate approach for a specific environmental issue depends on several factors, including the nature of the externality, as some externalities might be easier to address through Coasean bargaining than others. In the case of cloud seeding, the externality seems to be easily defined. The challenge remains to quantify it to have grounds for negotiations.

1.3.4 Framework for Efficient Cloud Seeding Investment

Allocating effort efficiently for cloud seeding across diverse regions necessitates a data-driven approach grounded in economic principles. This approach centers on the concepts of net benefit, marginal cost, and marginal benefit, guided by the efficiency equimarginal principle. The principle of efficiency emphasizes achieving the optimal outcome given available resources. In simpler terms, this principle posits that maximizing net benefits from resource allocation requires equating the marginal benefit gained from an investment with its marginal cost. Translated to the specific context of cloud seeding, this means directing effort towards regions where the additional economic value generated by increased precipitation, measured in dollars, is equal to the additional cost incurred for cloud seeding operations, conditional on there being sufficient clouds to seed.

Implementing this principle effectively requires various steps. The initial step is to quantify the impact of cloud seeding on precipitation by measuring the percentage increase in precipitation attributable to cloud seeding in each target region. The next step is to estimate the marginal cost of producing usable surface water. While a precipitation increase is crucial, the ultimate goal is to enhance readily available water resources. Therefore, it is useful to estimate the marginal cost associated with producing an additional unit of usable water, rather than focusing only on precipitation itself. This cost includes various factors such as infrastructure investments, operational expenses, and potential negative externalities. Lastly, where data permits, the analysis should be extended to assess the marginal benefit of each project. By comparing the dollar value of the additional precipitation benefit, based on its impact on sectors like agriculture and municipal water, with the marginal cost of producing usable water across different regions, conclusions could be drawn about the identification of areas where cloud seeding investments yield the highest return on investment. This rigorous and data-driven approach ensures that limited resources are directed toward the most impactful interventions.

1.3.5 Testable Hypotheses

Drawing upon the established theoretical framework, I formulate a set of three hypotheses to comprehensively evaluate the effectiveness and potential externalities associated with cloud seeding projects. The first hypothesis posits that cloud seeding is effective in increasing precipitation within the targeted regions. This aligns with the core objective of cloud seeding interventions, which is to manipulate atmospheric conditions to stimulate precipitation formation and increase water resources. The second hypothesis posits that cloud seeding effectiveness is not uniform across different projects. This variability is attributed to project-specific factors such as pre-existing weather conditions including temperature, humidity, wind patterns, as well as geographical features of the target area such as topography and proximity to water bodies. These factors could lead to heterogeneity in the observed increase in precipitation across different cloud seeding projects. The third hypothesis focuses on the potential consequences of cloud seeding beyond the targeted regions. Cloud seeding activities may generate either positive or negative externalities for downwind areas. These externalities would arise due to the large-scale nature of atmospheric circulation, where altered weather patterns triggered by cloud seeding can extend beyond the intended target zone. In particular, I hypothesize that downwind regions might experience decreased precipitation. This could exacerbate existing drought conditions in downwind areas, potentially leading to negative consequences. By testing these hypotheses, this study aims to gain a comprehensive understanding of the effectiveness and potential drawbacks of cloud seeding as a water management tool in California. The findings will contribute valuable insights for policymakers and water resource managers as they navigate the complexities of implementing and evaluating cloud seeding projects.

1.4 Empirical Methods

I evaluate the effects of the implementation of 12 precipitation enhancement programs using panel data in a setting where some locations, represented by 8-digit hydrologic units (HUC8), are exposed to cloud seeding events starting in 1999 but other units located outside of the targeted areas are not treated. In the following subsections, I present the empirical methodology and summarize the data used in the analysis.

1.4.1 Data

1.4.1.1 Cloud Seeding Operations

I obtain weather modification data from the National Oceanic and Atmospheric Administration (NOAA) Weather Program Office. As part of Public Law 92-205, commonly known as the Weather Modification Act of 1976, all non-federal weather modification activities must be reported to the U.S. Secretary of Commerce. These annual reports are made publicly available by the NOAA Weather Program Office. I obtain all the reports submitted by companies operating in the state of California available online from the NOAA Library. All weather modification activities in this subset involve cloud seeding for the purpose of increasing the local water supply through increases in rain or snow. The available reports range from 1999 to 2022, which encompasses the full extent of programs conducted in California, aside from operations conducted in Santa Barbara County, which have been in existence since 1974. Each annual report indicates the geographic delineation of the area targeted by cloud seeding activities, as well as the approximate location of the area that could be used as a reference when conducting comparison analysis of precipitation to quantify the effectiveness of the weather modification efforts. Cloud seeding is typically done by dispersing one of three seeding agents: silver iodide, sodium chloride, or potassium chloride, into the clouds using generators that can be either mounted on the ground or attached to the wings of an aircraft. For each month of the year reported, the lead meteorologist of the project reports the number of seeding events conducted, the type and quantity of each seeding agent used, the technology used to disperse the seeding agents into clouds, and the total number of hours that the generators were active.

1.4.1.2 Selecting Target and Control Units

For the 12 cloud seeding projects operational in California between 1999 and 2022, I identify a set of target and control units. Target units are areas that are targeted by cloud seeding projects to increase precipitation. They are identified in the NOAA reports either by hand-drawn maps or by the name of the sub-basin they cover. To standardize the unit of analysis, I use the boundaries of the 8-digit level hydrologic units of the United States Geological Survey (USGS) Watershed Boundary Dataset to identify the subwatershed hydrologic units that correspond to the target areas as determined by the cloud seeding reports. The subwatershed hydrologic unit boundaries provide a uniquely identified and uniform method of subdividing large drainage areas (WGISC, 2013). A hydrologic unit is a drainage area nested in a multi-level, hierarchical drainage system. Its boundaries are defined by hydrographic and topographic criteria that delineate an area of land upstream from a specific point on a river or a stream. This unit is useful in the context of this study not only because the HUC8 correspond to the shape and location of the cloud seeding projects areas, but also because precipitation can only be turned into surface water within a defined drainage area. Since the purpose of cloud seeding is to increase the surface water supply, it seems appropriate to use the HUC8 sub-basins as the unit of observation in this analysis.

Determining the best control units is a critical step in the analysis. When the control area is explicitly specified in the NOAA reports, I use the corresponding HUC8 as the unit of observation. I complement this set by using public projects description provided by RHS Consulting [2], when applicable. One caveat of using these private reports as the primary source of data for the selection of control units is that some areas are simultaneously considered target of one project and control for another adjacent project. This causes a problem of contamination when analyzing the effectiveness of cloud seeding for different projects independently. To overcome this, I also use typical wind direction during storms that impact a given sub-watershed to choose control units that are located upwind of the target area. This selection process results in a set of 15 target and 15 control sub-basins, as shown in Figure 1A.2.

To analyze downwind effects of cloud seeding projects on adjacent areas, I use the Eastern Sierra project located the closest to the Sierra Nevada mountains. Figure 1A.3 shows the location of control and target HUCs for this project, as well as the downwind HUCs located primarily in Nevada. While multiple projects were considered, the Eastern Sierra project emerged as the most appropriate choice due to its unique geographical context. Unlike other projects where downwind areas might be difficult to delineate due the proximity of different target areas from other projects causing some overlapping in the downwind areas of various projects, the Eastern Sierra project

[2]RHS Consulting is a privately owned cloud seeding company located in Nevada, that operates in parts of the Sierra Nevada.

offers a clear and distinct separation. This spatial clarity allows for unambiguous identification of downwind hydrologic units that are likely to be influenced by cloud seeding activities in the Eastern Sierra without the potential for contamination from other projects.

Different cloud seeding projects started operating in different years throughout the span of the panel dataset. The sample is composed of four cohorts, such that each cohort is a set of projects that were implemented at the same time. In terms of units of observation, a cohort includes all watershed units that start being targeted by a cloud seeding project in a given year, as well as units that are selected to be controls for said projects. For instance, the 1999 cohort is the set of projects that were implemented starting in 1999 in California. There is one project in this cohort - Sacramento, represented by five watersheds, two of which are treated by cloud seeding, and the other three are control units. Table 1B.1 summarizes the projects included in each cohort.

1.4.1.3 Weather variables

The purpose of cloud seeding in California is to increase surface water supply through an increase in precipitation. To analyze the efficacy of cloud seeding projects, one could look at changes in surface water or directly at changes in precipitation. Silverman (2010) uses full natural flow streamflow data to conduct a statistical comparison between the streamflows at the drainage area outlets of the sub-watersheds he studies. Many physical parameters can influence streamflow at any given geographical point, making it difficult to draw causal inference for any observed changes. In this paper, I use precipitation data directly and precipitation to inflow conversion ratios for each sub-watershed.

Weather variables are obtained from PRISM Climate Data Monthly Time Series. For each selected unit of observation, whether it is a target or control sub-basin, I extract monthly precipitation totals to use as the dependent variable, along with other variables such as temperature and vapor pressure deficit, averaged over all days in the month. Precipitation totals in PRISM are estimated using the total of rain and melted snow in each grid.

1.4.1.4 Descriptive statistics

As reported in Table 1B.2, for the period 1981-1999, the average monthly precipitation totals amount to approximately 49.5 mm in watersheds that never receive cloud seeding -control, and 67.7 mm in watersheds that are eventually treated. As expected, precipitation in both groups exhibit high variability, with standard deviations being 1.5 times higher than the means. The pooled -across watersheds- average monthly precipitation totals for the period 1999-2022 suggest a decreasing trend, with 43.1 mm in control watersheds and 41.4 mm in target areas during months where no cloud seeding occurs. This simple comparison shows that, without cloud seeding, target and control watersheds receive the same amount of monthly precipitation, on average. On the other hand, target watersheds receive 102 mm of average monthly precipitation during months of active cloud seeding operations. Unlike previous papers that systematically attribute this difference in trend to the cloud seeding projects (e.g. Griffith et al. (2015)), this paper tests the causal relationship between cloud seeding events and precipitation enhancement, using robust empirical methods explained in the next section.

1.4.2 Empirical strategy

1.4.2.1 Difference-in-differences

I hypothesize that cloud seeding operations increase precipitation in target areas. To investigate this hypothesis, I adopt a difference-in-differences (DiD) methodology and estimate the average treatment effects on the treated (ATTs). The identification assumption is that the difference in monthly precipitation between units that were treated by cloud seeding and those that were not would be constant over time in the absence of a precipitation enhancement program. A common approach to estimating a linear model in this context is to include unit fixed effects and time fixed effects in ordinary least squares estimation (Wooldridge, 2021). The resulting estimator is called the "two-way fixed effects" (TWFE) estimator. The DiD approach compares the changes in precipitation between seeded and non-seeded regions before and after cloud seeding implementation. This comparison isolates the causal effect of cloud seeding by controlling for time-invariant factors that might influence precipitation patterns in both regions, such as geography, topography, and long-term climate trends. Furthermore, the TWFE estimator accounts for unobserved, time-invariant characteristics specific to each region (e.g., average elevation, prevailing wind patterns) and year (e.g., overall weather patterns, large-scale climate fluctuations). By controlling for these fixed effects, the causal effect of cloud seeding is isolated from other potential confounding factors that might vary across regions or over time. There have been many contributions to the literature on DiD, the reader is referred to an excellent review of the most recent advancements by Roth et al. (2023).

To estimate the model, the target and control sub-watersheds, denoted $i \in N$, are uniquely identified panel units. Let y_{it} denote the total daily precipitation at time t in sub-watershed i. Let $CloudSeeding_{it}$ be the cloud seeding event indicator, such that $CloudSeeding_{it} = 1$ if unit i is treated in month t. The Two-Way Fixed-Effects estimator is obtained from estimating the following equation, where c_i and f_t are unit- and time fixed effects, respectively, and u_{it} is the error term:

$$y_{it} = \beta_0 + \tau \cdot CloudSeeding_{it} + c_i + f_t + u_{it} \tag{1.1}$$

Since the implementation of cloud seeding projects happens in different years for different projects, the estimating equation is modified to obtain cohort average treatments and the ATTs are estimated using the 'extended' TWFE. The first project is implemented at time $q \in 2, ..., T$ and then some additional units are treated for the first time in subsequent periods. The potential outcomes are denoted $y_t(r), r \in \{q, \ldots, T, \infty\}, t \in \{1, 2, \ldots, T\}$ such that r indicates the cohort and t is month of the year. $y_t(\infty)$ indicates the potential outcome at time t for units that are never treated. Thus, the ATTs of interest are:

$$\tau_{rt} = E\left[y_t(r) - y_t(\infty) \mid d_r = 1\right], r = q, \ldots, T; t = r, \ldots, T,$$

where the $\{d_r : r = q, \ldots, T\}$ are cohort indicators defining the first period a unit is subjected to the intervention. For each treated cohort r, τ_{rt}, $r = q, \ldots, T$ are the ATTs in all subsequent time periods. These average treatment effects are aggregated, post-estimation, to obtain one point estimate of the pooled effect of cloud seeding on precipitation. The extended TWFE is obtained from the following equation:

$$y_{it} = \sum_{r=q}^{T} \sum_{s=r}^{T} \tau_{rs} \cdot \underbrace{d_{ir}}_{\text{cohort}} \cdot \underbrace{f_{st}}_{\text{time}} + c_i + f_t + u_{it} \tag{1.2}$$

According to the hypothesis that cloud seeding increases precipitation in target areas, the aggregated ATT, $\tau = \sum_{r=q}^{T} \sum_{s=r}^{T} \tau_{rs}$, as well as cohort ATTs, are expected to be non-negative. Downwind effects are estimated using the same equation, such that target HUCs are replaced by downwind HUCs and the treatment remains the same. The ATTs could take either sign, indicating a spillover or contamination effect if the sign is positive, or adverse effect of cloud seeding on adjacent areas is the sign is negative.

There are two potential pitfalls to applying a TWFE model in this context. First, a TWFE model implicitly assumes homogeneity across cross-sectional units in the impact of observed determinants of the outcome variable, referred to as a slope parameter homogeneity assumption. In the context of sub-basin-level analysis of cloud seeding, a slope parameter homogeneity assumption would imply that changes in cloud seeding variables, such as the binary treatment, the technology used, or the amount of seeding agent would affect precipitation in all sub-basins in the same manner. Perhaps it would have been straightforward to justify this assumption if sub-basins were identical in terms of other characteristics not held constant in a regression model. This might, however, not be the case, and sub-basins could be different along many observed and unobserved dimensions. For instance, due to differences in the topographical characteristics of sub-basins (e.g., elevation, distance from coast), the marginal effect of cloud seeding on precipitation might not be uniform across all sub-basins. If this is the case, then assuming a common slope parameter could lead to biased and inconsistent estimates of the impact of treatment variables of interest (Sun and Shapiro, 2022).

Second, a TWFE model based on sub-basin-level panel data characterizes the impact of any time-specific, unobservable common shocks to the outcome using sub-basin-invariant time dummies, thereby implicitly assuming homogeneity in their effects across projects. This additional homogeneity assumption also poses a problem in my empirical application, since one of the objectives of this paper is to capture differences in the efficacy of cloud seeding programs across the state of California.

Additionally, the presence of such types of unobserved common shocks with potentially differential effects across sub-basins on precipitation patterns and water availability would create dependence across sub-basins in the error terms of a TWFE model. To see why this might be the case, consider drought episodes as an example of an unobserved common shock. Although droughts substantially decrease cloud systems in nearly every sub-basin, such effects might not spread equally across the state of California, and some sub-basins could experience more drought effects than others. The cross-sectional dependence in error terms could cause TWFE estimators to yield biased and inconsistent estimates if the source generating the dependence is also correlated with observed variables in the model (Chudik et al., 2011; Chudik and Pesaran, 2015). To alleviate concerns about time-varying unobserved heterogeneity, I explicitly model time-varying unobserved heterogeneity by adding sub-basin-specific linear time trends to the traditional TWFE models (Wooldridge, 2021).

1.4.2.2 Identification

Identification of the average treatment effects on the treated is based on random variables representing an underlying population of sub-basins present over time periods $t \in 1, 2, \ldots, T$. Under the "No Anticipation" assumption and the "Parallel Trends" assumption, the ATT is identified. The "No Anticipation" assumption states that the treatment, cloud seeding, has no effect on the outcome, precipitation, before it is actually implemented. For this analysis, the assumption would simply imply that precipitation patterns within the treated sub-basins are not influenced by any prior knowledge of the cloud seeding intervention. This means that any observed changes in precipitation after the intervention can be confidently attributed to the cloud seeding itself, not to any anticipation or preparatory actions within the sub-basin. Naturally, the physical processes involved in precipitation formation are not affected by any pre-existing knowledge of the intervention. Therefore, the "No Anticipation" assumption holds.

The parallel trends assumption guarantees that, absent the intervention, both groups would have experienced similar precipitation trends. This allows the observed change in the treated group, relative to the control group, to be confidently attributed to the cloud seeding intervention.

Verifying the parallel trends assumption can be tricky but is crucial for drawing reliable conclusions. I use pre-trend tests or ex- ante placebo tests as an assessment of parallel trends. Placebo tests introduce a hypothetical "fake treatment" applied to the real sample at time points before the actual cloud seeding intervention happens. These time points represent fictional interventions that never occurred. By statistically comparing the trends in precipitation after these "fake treatments" to the actual trends in the control group, we essentially create a synthetic control for the treated group before the actual intervention. This is one possible robustness check to assess whether the trends in both groups were truly parallel prior to the cloud seeding. If the trends following the "fake treatments" diverge significantly from the untreated control group, it raises concerns about the parallel trends assumption. This divergence suggests that the treated and control groups were already on different trajectories before the actual intervention, potentially leading to biased estimates of the cloud seeding effect. Conversely, if the trends in precipitation after the "fake treatments" in the control group remain parallel to the actual trends in the untreated control group, it strengthens the confidence in the parallel trends assumption.

Encouragingly, the ex-ante placebo tests I conducted yielded reassuring results regarding the parallel trends assumption in the cloud seeding analysis. Visual results of multiple placebo tests are presented in Figure 1A.4, showing that there is not enough statistical evidence to reject the null hypothesis that the fake treatment has no effect on precipitation. This suggests that the observed difference in precipitation between the treated and control groups after the "real" cloud seeding is less likely to be due to pre-existing divergence in their trajectories. While the parallel trends assumption cannot be definitively proven, the lack of significant effects in the placebo tests provides support for its validity.

While the established methodology offers a starting point to analyze the true impact of cloud seeding on target areas, it's crucial to acknowledge the potential presence of selection bias. Selection bias can arise in this context if the factors influencing the decision to conduct cloud seeding are also correlated with the natural precipitation patterns in the target area. This can lead to misleading comparisons between the target and control areas. The crux of the issue is that cloud seeding is intrinsically linked to a specific condition: the presence of clouds. This introduces a classic case of selection bias because the sample becomes inherently skewed towards days when precipitation is more likely to occur naturally. Since there are no observations of cloud seeding on dry days, the model is not getting a fully representative picture of the weather conditions.

Normally, a well-executed difference-in-differences strategy helps account for pre-existing differences between the treatment and control groups. However, in this case, the very act of cloud seeding means there is a fundamental pre-existing difference on seeded days: the presence of more clouds, making the treatment group likely to have higher baseline precipitation than the control group. So far, without correcting for this source of bias, the model is likely to compare days with natural precipitation plus cloud seeding to days with only natural precipitation. This comparison would exaggerate the apparent impact of cloud seeding, as natural precipitation would be present anyway during the seeded days. While eliminating selection bias entirely can be challenging, certain strategies can help mitigate its potential impact. Common methods include (1) propensity score matching, which is a pairing of observations from the target and control areas based on their similarity in factors influencing the decision to seed, leading to a more balanced comparison, and (2) instrumental variables - which consists of using a factor, the instrumental variable, that influences the decision to seed but is not directly correlated with precipitation, to estimate the causal effect of seeding. The application of the aforementioned methods in this case is limited due to the lack of data on relevant control variables.

To address the challenge of insufficient data on control variables, this study employs a thresholding approach to mitigate some of the potential bias arising from the inherent link between cloud seeding and the presence of clouds. This approach acknowledges the inherent selection bias but attempts to reduce its influence by restricting the sample to observations that meet a specific threshold. In this case, the threshold is set for positive precipitation values (precipitation > 0). By focusing solely on days with precipitation, the analysis operates within a more homogeneous subset where naturally occurring clouds are already present. This mitigates some of the bias because the comparison is now primarily between cloudy days with cloud seeding (treatment) and cloudy days without cloud seeding (control). This reduces the influence of the pre-existing difference in cloud cover, leading to a more comparable treatment and control group. It is also possible that the treatment effect of cloud seeding is genuine and not completely explained by selection bias. The key point here is that without addressing this bias, it is not possible to reliably distinguish the true effect of seeding from the effect that arises simply due to the differences between cloudy/dry days.

The potential outcomes framework implicitly assumes the stable unit treatment value assumption that unit i's outcome is independent of the treatment status of unit $j \neq i$, which rules out spillover effect. This is ensured by choosing appropriate control units that are located upwind of the target units, as explained in the data section.

1.4.2.3 Non-linear difference-in-differences: Poisson Quasi Maximum Likelihood Estimator (QMLE)

Assuming a constant treatment effect and using the TWFE estimator can lead to an estimated coefficient that is difficult to interpret (De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021). To obtain heterogeneous treatment effects, Sun and Abraham (2021) propose weighted estimators in event studies, Borusyak et al. (2021) derive imputation estimators based on TWFE, and Callaway and Sant'Anna (2021) apply treatment effects estimators to long differences, using never treated as well as already treated units as control groups. Wooldridge (2023) proposes an alternative, or at least a supplement, to the linear model when the observed outcome is nonnegative with a corner at zero: the exponential mean model estimated by Poisson QMLE.

Furthermore, consistent estimation of the ATTs by TWFE requires a linear parallel trends assumption. Since the outcome variable, monthly precipitation totals, is zero-inflated such that it takes on the value zero regularly as well as large values, the linear parallel trends assumption might fail. In this case, a better assumption would be that the parallel trends assumption holds in the ratio of means (Wooldridge, 2023). This is implied by an exponential conditional mean function, and this specification is sufficient to apply a pooled quasi-maximum likelihood estimation (QMLE). Using pooled QMLE in the class of linear exponential family (LEF) distributions ensures that the resulting estimators of the ATTs do not rely on other distributional assumptions, nor on restrictions in patterns of serial dependence across time. When $y_{it} \geq 0$ without a natural upper bound, the exponential mean makes sense as an alternative to a linear mean. With the exponential mean, the TWFE estimator does not suffer from an incidental parameters problem. The incidental parameters problem arises in statistical models where there is a substantial number of parameters that grow along with the sample size. A common scenario where this occurs is panel data. The problem is usually amplified in nonlinear models compared to simple linear regression. As the sample size increases and the number of "incidental" parameters grows along with it, traditional estimation methods like maximum likelihood estimation can become inconsistent.

TWFE models often assume a linear relationship. In panel data, this means adding a unit-specific intercept term (fixed effect) for each unit and a time fixed effect for each period, as this helps control for unobserved factors that might influence the outcome. Unlike linear models where unit-specific effects are added, the exponential mean model assumes a multiplicative relationship. This means that the fixed effects impact the outcome variable multiplicatively rather than through simple addition. Due to the nature of the exponential mean, the number of parameters needed to estimate no longer grows with the sample size. This is what effectively sidesteps the incidental parameters problem. As mentioned previously, the exponential mean is particularly well-suited for count data with high occurrences of zero. This aligns with how Poisson QMLE is utilized for estimation. The adoption of Poisson QMLE aligns with the evolving landscape of statistical methods in causal inference research (Wooldridge, 2023; Roth et al., 2023). This ultimately leads to a more reliable and insightful analysis of the impact of cloud seeding on precipitation.

For any time period t and random draw i,

$$\begin{aligned} E\left(y_{it} \mid d_{iq}, \ldots, d_{iT}, c_i\right) = \underbrace{c_i}_{\text{unit-FE}} \exp \Bigg[& \sum_{s=2}^{T} \theta_s f_{s_t} \\ & + \sum_{r=q}^{T} \sum_{s=r}^{T} \beta_{rs} \Bigg(\underbrace{w_{it}}_{\text{treatment}} \cdot \underbrace{d_{ir}}_{\text{cohort}} \cdot \underbrace{f_{s_t}}_{\text{time-FE}} \Bigg) \Bigg] \end{aligned} \tag{1.3}$$

Such that β_{rs} are the coefficients of interest as they represent the average effect of cloud seeding, considering the cohort and time period, w_{it} is a binary variable indicating the treatment status (1 if cloud seeding occurred at site i in period t, 0 otherwise), d_{ir} captures the cohort effect, f_{st} is a dummy variable indicating the time fixed effects, and c_i indicates site fixed effects.

1.4.2.4 Costs and Benefits

Benefit-cost analysis (BCA) is a cornerstone of applied welfare economics within the discipline of natural resource economics, particularly in the water sector. Its historical significance lies in the evaluation of water resource investments undertaken by federal agencies, primarily focusing on projects spearheaded by the US Bureau of Reclamation and the US Army Corps of Engineers. The main objective of BCA in this context was to provide a comprehensive picture of the costs and benefits associated with water development initiatives. The intellectual foundation of BCA can be traced back to Jules Dupuit, who, in 1844, introduced the concept of consumer surplus in his work "On the measure of the utility of public works" (Dupuit, 1844). This groundbreaking notion recognized that the true value of public works projects extends beyond their direct revenue generation, acknowledging the existence of intangible benefits for consumers. Beyond its initial application in evaluating water resource investments, BCA has witnessed a remarkable expansion in its scope, which includes the economic consequences of emerging technologies and regulatory programs. Although its initial use cases primarily involved relatively straightforward evaluations of water development projects, economists actively engaged in debates about the appropriate methods for addressing both empirical and conceptual challenges associated with the technique. These discussions focused on value and equity considerations, including whether to account for the distribution of benefits and costs among individuals or solely focus on aggregate values.

This analysis is tangent to a benefit-cost analysis, but it differs from it in various ways, such that the focus is on the marginal cost and benefit rather than the total costs and benefits. Due to data limitations, this paper does not employ a standard benefit-cost analysis. Rather, it builds upon the established principles of the technique to go beyond simply assessing the effectiveness of individual cloud seeding projects in enhancing precipitation, in an effort to offer a framework to optimize resource and effort allocation between cloud seeding projects, and within the broader context of water management.

I begin by leveraging the estimated treatment effect, β_{it} from the Poisson QMLE for each month and each project. This estimate represents the additional precipitation attributable to cloud seeding interventions. I use these estimates to construct two precipitation scenarios: (i) Precipitation with cloud seeding ($precip_{1,it}$), which is the observed precipitation data for each month and project; and (ii) precipitation without cloud seeding ($precip_{0,it}$), which represents a counterfactual scenario, constructed by assuming no cloud seeding intervention. I calculate this counterfactual precipitation as $precip_{0,it} = precip_{1,it}/(1 + \beta_{it})$. This calculation essentially adjusts the observed precipitation downwards by the estimated treatment effect, providing an estimate of what precipitation would have been without cloud seeding. Then, the estimated difference in precipitation that is attributed to cloud seeding is $\Delta precip_{it} = precip_{1,it} - precip_{0,it}$.

While assessing the impact of cloud seeding on precipitation is crucial, solely focusing on this metric provides an incomplete picture of its effectiveness as a water management tool. A more informative and practical approach lies in estimating the effect of cloud seeding on usable surface water. This shift in focus acknowledges the complexities involved in translating increased precipitation into readily available water resources. Not all precipitation translates directly into usable water. Factors like evaporation, infiltration, and runoff contribute to significant losses, particularly in arid and semi-arid regions. Focusing solely on precipitation overlooks these essential aspects, potentially overestimating the actual water gain from cloud seeding.

Runoff ratio is the runoff for each watershed divided by the precipitation for that watershed. It is the proportion of rainfall that does not infiltrate and is not taken up by evapotranspiration, and thus ends up as runoff. The conversion rate from precipitation to usable surface water is watershed-specific. Runoff is affected by (1) meteorological factors such as the type of precipitation, the intensity, amount, and duration of rainfall, the distribution of rainfall over the drainage basin, the direction of storm movement, and soil moisture; as well as (2) physical characteristics such as vegetation, soil type, basin shape, elevation, topography, and drainage network patterns (USGS, 2019). General regions behave differently based on these factors and runoff varies greatly between mountain and valley areas.

Runoff data is obtained from the USGS Water Watch for each hydrologic unit, i, and each month, t, for the period 1981-1998. The USGS generates 8-digit-HUC runoff by combining historical flow data collected at streamgages, the drainage basins of the streamgages, and the boundaries of the HUCs, for the period 1901-2009 [3]. Using this data, I estimate average HUC runoff ratios, denoted ρ_i, by taking the time average of the ratio precipitation/runoff for each HUC. The next step in estimating the quantity of additional surface water, w_{it}, obtained from cloud seeding is straightforward: $w_{it} = \Delta precip_{it} \cdot \rho_i \cdot area_i$ such that $\Delta precip_{it}$ is the difference in monthly precipitation totals with and without cloud seeding.

Since the only cost data available is the approximate yearly cost for each project, I employ a yearly aggregation approach to estimate the marginal cost. I sum up the monthly runoff differences across the year for each project to obtain the annual difference in surface water: $w_{iy} = \sum_{t=1}^{T} w_{it}$, where the subscript y denotes year. Then, I estimate the yearly marginal cost for each project by dividing the total project cost by the annual difference in runoff: $MC_i = Total\_Cost_{iy}/w_{iy}$. This calculation provides an initial estimate of the marginal cost per unit of additional surface water generated through cloud seeding for each project in a given year. To account for potential yearly variations and obtain a more robust estimate, we average the yearly marginal cost across all years for each project. This yields the average marginal cost of surface water obtained by cloud seeding for each project.

1.5 Results

1.5.1 Changes in monthly precipitation totals

Results from the two-way fixed effects models, summarized in Table 1B.3, suggest that cloud seeding increases monthly precipitation totals by 52 mm, on average. Using the restricted sample, the estimation, presented in Table 1B.4, shows an increase of 51 mm/month on average.This suggests that selection bias might not be a significant concern. Furthermore, using ground-based or aircraft-based generators results in the same effects, as shown in Table 1B.5. All results are highly statistically

[3]Runoff data is not used as the dependent variable in the analysis because it is not available for the time period of interest.

significant. When estimated separately, cloud seeding by aircraft increases precipitation by 52 mm/month while cloud seeding using ground generators increases precipitation by 55 mm/month. These estimates are not statistically different from each other. In fact, both methods are always used simultaneously, which means that the estimates in Column 2 and Column 3 mirror each other. I include both methods as treatment variables in the same estimating equation and find that the partial effects are such that cloud seeding by aircraft-based generators increases precipitation by 34 mm/month while cloud seeding by ground-based generators increases monthly precipitation totals by 40 mm, when considered simultaneously. These effects are complementary and suggest that combining both technologies might increase the overall effect of any given cloud seeding operation. Results using the restricted sample, as shown in Table 1B.6, reveal estimated increases of 27 mm/month and 29 mm/month for aerial seeding and ground seeding, respectively.

I explore the sensitivity of the findings to the inclusion of different fixed effects in the model, in Table 1B.7, shedding light on the potential influence of unobserved factors. The first specification includes only the treatment variable, indicating an average increase in precipitation of 52 mm/month due to cloud seeding. However, this estimate might be biased by unobserved characteristics specific to each project site that are correlated with both cloud seeding and precipitation. To address this, I introduce site fixed effects, which control for these time-invariant factors. This adjustment results in a slightly lower estimated effect of 44.20 mm/month, suggesting that some of the initial increase could be attributed to site-specific differences rather than cloud seeding itself. Finally, I add time fixed effects to control for any unobserved temporal trends in precipitation that might affect all projects equally. This additional control leads to an estimated effect of 52 mm/month, which is surprisingly close to the initial estimate. This suggests that time-varying factors not specific to individual projects may have played a role in the initial larger estimate. Controlling for both site and time variations strengthens the evidence for a positive effect of cloud seeding, with an estimated increase of 52 mm/month.

Having established the positive impact of cloud seeding on precipitation, I explore the influence of key weather variables on the outcome variable in Table 1B.8. First, I introduce the covariate $tdmean$, representing the average daily dew point temperature for the month. The negative coefficient of -7.316 for $tdmean$ indicates that higher dew point temperatures are associated with lower precipitation. This suggests that when dew point temperatures are already high, there may be less room for additional precipitation enhancement from cloud seeding, leading to the smaller estimated effect of 29.36 mm compared to the initial model. However, it's important to consider a potential endogeneity issue here. The decision to cloud seed might be correlated with dew point temperature since, as explained in the background section, the decision to proceed with cloud seeding requires low temperatures. Thus, cloud seeding might be more likely to occur during periods of lower dew point temperature (drier conditions) when the potential for precipitation increase is higher. Similarly, adding the explanatory variable $vpdmean$, the average monthly vapor pressure deficit, shows a negative coefficient of -6.656, which indicates that higher vapor pressure deficit (drier conditions) is associated with lower precipitation.

Finally, I estimate the full model with both $tdmean$ and $vpdmean$. Here, the cloud seeding effect remains positive at 18.74 mm, further confirming its influence. Interestingly, the coefficient for $tdmean$ becomes statistically insignificant, suggesting an issue with including both these variables. On the other hand, $vpdmean$ maintains its negative association with precipitation (-6.963). While including relevant covariates might seem like a good idea to improve the explanatory power of the model, it might be causing over-fitting problems. The inclusion of fixed effects in the analysis already controls for many unobserved factors that might be correlated with both cloud seeding and precipitation. Therefore, including additional covariates to mitigate spurious relationships is unnecessary.

Moving to the analysis of the second hypothesis on the effectiveness of different levels of cloud seeding intensity, I estimate the effect of the intensity of treatment using four distinct measures of cloud seeding intensity: (1) $NumberDays$ indicates the number of days that cloud seeding occurred within a month; (2) $HouseAirborne$ captures the total hours of cloud seeding conducted using aircraft during the month; (3) $HoursGround$ represents the cumulative hours of ground-based generators active in cloud seeding throughout the month; and (4) $AmountSilverIodide$ measures the quantity of seeding agent (silver iodide) used in cloud seeding operations for the month, expressed in kilograms. The results from the separate regressions are presented in Table 1B.9. Each additional day of cloud seeding is associated with an increase of 14 mm in monthly precipitation. This suggests a strong positive impact of prolonged cloud seeding activities. An extra hour of cloud seeding using aircraft translates to an average increase of 6 mm in precipitation. While positive, the effect is smaller compared to the number of days, indicating diminishing returns with longer airborne operations. Interestingly, an additional hour of ground-based generator activity has a marginal effect of only 0.18 mm on precipitation. This suggests a limited impact of this specific method compared to airborne seeding. Each additional kilogram of silver iodide used leads to an increase of 10 mm in precipitation. This finding highlights the potential dose-dependent relationship between the seeding agent and its effectiveness. While all measures show a positive impact on precipitation, the magnitude and rate of increase vary considerably. This suggests that the optimal approach might involve a combination of methods and durations, tailored to specific weather conditions and geographical contexts.

To address the challenge of potential time-varying unobserved heterogeneity, I introduce heterogeneous slopes into the model, allowing the effect of cloud seeding to vary across different sites. The results from this robustness check are presented in Table 1B.10. Reassuringly, allowing this site-specific flexibility yields an estimated effect of cloud seeding on precipitation of 51 mm/month. Compared to the initial estimate of 52 mm/month, the difference is negligible in the real-world context as these two estimates are statistically indistinguishable from each other. This suggests that accounting for the unique characteristics of each project site and their temporal variations does not substantially alter the overall conclusion about the positive impact of cloud seeding on precipitation.

1.5.2 Changes in percentages

Moving beyond the average treatment effect on the precipitation levels, and in order to understand the magnitude of these effects and be able to make comparisons between the different projects, I show the results of the Poisson QMLE estimation strategy, first for the complete sample in Table 1B.11, and then for the 12 subsamples. This approach is particularly suited for this setting as it acknowledges the non-linear nature of precipitation events. The Poisson QMLE results suggest that cloud seeding is associated with a 57% increase in monthly precipitation totals, on average. This substantial increase underscores the potential power of cloud seeding to augment precipitation levels. Again, using the restricted sample in Table 1B.12 suggests that there is no concern for selection bias as the estimate shows an increase of 56% per month in precipitation attributable to cloud seeding. The reader is reminded that this percentage change differs from the millimeter-based increases reported in previous sections. This shift reflects the focus of the Poisson QMLE on proportional changes, providing insights into the relative impact of cloud seeding across different precipitation levels. However, the results are reassuringly close in magnitude because the increased 52 mm/month corresponds to a 51% increase in precipitation on average.

Having established the overall impact of cloud seeding on precipitation, I test the third hypothesis on the difference of cloud seeding effectiveness across different projects. I show the differences across individual projects in Tables 1B.13 and 1B.14, to examine heterogeneity across projects. Using the same Poisson QMLE strategy, I estimate the project-specific effects on precipitation changes. The results reveal significant variation in the impact of cloud seeding across the 12 projects. The estimated increases in monthly precipitation range from 30% in San Joaquin Valley to 140% in the Kings River watershed. I use the graphic in Figure 1A.5 as a visual representation of these effects. This variation highlights the potential influence of local factors on cloud seeding effectiveness. Interestingly, a spatial pattern emerges. Projects located closer to the Eastern Sierra mountain range exhibit higher increases in precipitation compared to those further away. This observation aligns with the general understanding of orographic lifting, where rising air masses cool and condense, leading to enhanced precipitation. Changes in the rainfall amounts between the coast and inland during the rainy season occur due to the differences between sea surface and land temperatures during the winter months and the effects of orographic lifting inland (Levin et al., 2010). A hypsographic map of California showing elevation levels is presented in Figure 1A.6.

1.5.2.1 Downwind Effects: The Eastern Sierra Project

The effects of cloud seeding on downwind areas are estimated using data from the Eastern Sierra project, and shown in Table 1B.15. Focusing on the Eastern Sierra project, I find that cloud seeding results in a 22% increase in average monthly precipitation totals within the targeted HUCs. However, results also indicate a negative impact of cloud seeding on precipitation in downwind HUCs, with an 8% decrease in average monthly totals. This result, statistically significant at the 1% level, suggests that cloud seeding might have unintended consequences beyond the immediate target area. While the exact mechanisms remain subject to further investigation, several hypotheses could explain this downwind effect. Changes in cloud droplet size due to seeding could influence precipitation formation and distribution downwind, potentially leading to decreased precipitation totals. Cloud seeding might also trigger competing convective processes in downwind areas, thereby reducing the overall moisture available for precipitation. The physical explanation of these processes is beyond the scope of this essay as the focus is on highlighting the importance of considering potential downwind effects when evaluating the economic viability of cloud seeding projects. In particular, cloud seeding in the Eastern Sierra, while potentially benefiting the targeted area, seems to generate a negative externality for downwind communities. The 8% decrease in precipitation imposes economic costs on these communities, potentially impacting agricultural yields, water availability, and even ecosystem health.

1.5.3 Economic viability of cloud seeding

Following the estimation of the average treatment effect of cloud seeding on precipitation, this section explains the economic implications by calculating the marginal cost of producing additional surface water through cloud seeding for each of the 12 projects under study.

Overall, I find that cloud seeding increases surface water at the cost of $7 per ac-ft, across all projects in the sample. Due to the heterogeneity in the effectiveness of cloud seeding, as well as project characteristics such as the target area, the frequency of seedable cloud systems, and the total investment, the cost of an additional ac-ft of water varies between $3.7 to $24.5 per ac-ft, as shown in Figure 1A.7. This visual representation allows for a quick comparison of the relative costs associated with producing additional surface water through cloud seeding in different project contexts. By following these steps, I translate the estimated impact of cloud seeding on precipitation into an economically relevant metric - the marginal cost of producing additional usable surface water. This analysis provides valuable insights into the economic feasibility and cost-effectiveness of cloud seeding interventions for water resource management in the specific contexts of each project.

While the previous section established a framework for estimating the marginal cost of producing additional surface water through cloud seeding for each project, this analysis remains incomplete without considering the marginal benefits. Simply comparing costs across projects is insufficient if the value of the additional water varies significantly due to differing contexts and uses. The value of water is not uniform across regions or applications. In drought-stricken areas, where water scarcity is acute, the value of additional water will likely be much higher compared to regions with abundant water resources. Additionally, water can be used for various purposes, each with its own associated economic value (e.g., agriculture, recreation, hydropower generation). These diverse uses further highlight the potential heterogeneity in the marginal benefit of cloud-seeding generated water. The scarcity of water varies significantly across different regions. For instance, the southern part of California generally faces greater water scarcity compared to the northern regions. This implies that the marginal benefit of additional water generated through cloud seeding would likely be higher in southern projects due to the increased value placed on scarce resources. Additionally, project-specific factors, such as the intended use of the water and its impact on local economic activities, can further influence the marginal benefit. Ultimately, comparing the estimated marginal benefits of additional water with the corresponding marginal costs for each project can provide a more complete picture of the economic feasibility and potential net benefits of cloud seeding interventions.

California State law stipulates that water gained from cloud seeding is treated the same as natural supply in regard to water rights (California Department of Water Resources, 2016). Therefore, the benefits of additional water are directly extrapolated from the different values of water. Agricultural water in California is valued from $40 to $50 per ac-ft and up to $175 per ac-ft during drought

(Hunter, 2007). The value of water for hydroelectric use by PG&E is $100 per ac-ft, while Municipal and Industrial values range from $300 to $600 per ac-ft, according to a 2007 California Energy Commission (CEC) report, that was made publicly available in 2018. While these established values offer a starting point, they may not fully capture the specific context of each cloud seeding project. Water values can fluctuate based on factors like drought conditions, location, and intended use. Recognizing these limitations, this section attempts to conduct a more thorough estimation of marginal benefits for each cloud seeding project, considering project-specific details and employing alternative valuation approaches.

First, utilizing the U.S. Department of Agriculture's "cultivated land" layer, I identify the proportion of agricultural land within each project's target area (HUC). However, the analysis reveals limited agricultural land coverage in most HUCs, necessitating an alternative approach for benefit estimation. Since agricultural water value wasn't directly applicable, I adopt water rates from the "California small water systems rates dashboard" as a proxy for the value of water. Matching HUCs with corresponding water systems within California counties, I extract residential water rates for each project. The estimated marginal costs for each project, as discussed in the previous section, as well as the water rates are presented in Table 1B.16. Water rates may not perfectly capture the economic value of increased water availability, particularly for agriculture. Moreover, the analysis relies on publicly available data, which might have limitations in accuracy or granularity. However, while the analysis acknowledges data limitations, it provides a valuable comparative perspective. Table 1B.16 summarizes the net benefits of cloud seeding by project. Perhaps the most striking observation from this table is that the estimated benefits exceed the costs by at least one order of magnitude in all cases. In fact, benefits are evaluated at 10 to 70 times higher than the costs, resulting in very high net benefits.

The analysis of project-specific costs, benefits, and net benefits generated by cloud seeding provides valuable insights for optimizing effort allocation based on efficiency principles. It highlights the heterogeneity in project performance and allows for targeted allocation of resources, ensuring that efforts are directed towards the most impactful investments. The projects with the highest net benefits per unit cost should be prioritized for effort allocation. Looking at Table 1B.16, or Figure 1A.8 for a visual representation, Kern and San Joaquin emerge as potential frontrunners, boasting net benefits of $444 and $687 per acre-foot, respectively. This suggests that focusing efforts on these projects could maximize the economic return on investment for cloud seeding initiatives. While prioritizing high net benefits is crucial, other factors might influence allocation decisions. For instance, projects with lower net benefits but serving critical water-scarce regions might warrant continued operation due to social equity concerns. Additionally, factors like political feasibility and environmental impact also require consideration.

The estimated 8% decrease in precipitation due to cloud seeding activities over the Eastern Sierra Nevada raises important economic considerations related to spatial externalities and equity issues. Cloud seeding, while generating localized benefits through increased precipitation, can also produce unintended consequences for downwind areas, highlighting the complex economic landscape of water management. Cloud seeding primarily impacts the target area by increasing snowfall, potentially leading to benefits like higher reservoir levels, improved water availability for agriculture and hydropower generation, and enhanced recreational opportunities. However, these benefits come at the potential expense of downwind regions. Reduced cloud moisture due to seeding upstream can translate to decreased precipitation downwind, creating negative externalities for those reliant on natural snowfall patterns. In the case of the Eastern Sierra Nevada, the downwind areas are HUCs located in Nevada, already experiencing limited precipitation due to the rain shadow effect of the mountain range. An 8% decrease in precipitation might exacerbate this existing water scarcity. While the potential 8% decrease in precipitation due to cloud seeding could translate to an estimated loss of $49 million per year if the water were directed to municipal use in a high-value location like North Las Vegas (based on their water rate of $2.34/1000 gallons[4] (Vegas, 2024) and an estimated additional volume of 64,000 acre-feet/year[5]), it's important to remember that this is likely an overestimation. Since the downwind HUCs in Nevada experience minimal agricultural

[4]The water rate of $2.34/1000 gallons is converted to $763/acre-foot.

[5]$Volume = \Delta precip_i \cdot area_i \cdot \rho_i$ such that $\Delta precip_i = |precip_{1,i}(\beta_i/1+\beta_i)|$. The total area of the HUCs considered in this analysis is 3332,819 acres, located on the leeway of the Sierra Mountains, in Nevada. The runoff conversion ratio is assumed to be $\rho_i = 0.4$, and $\beta_i = 0.08$.

activity and have no nearby cities with significant water demands, the actual economic impact is likely much lower. The lack of readily quantifiable economic losses does not negate the potential environmental and social impacts. Exploring environmental health valuation methods could be valuable in understanding the broader consequences of reduced water availability in these downwind areas, by assessing the impact on ecosystem health, biodiversity, and the well-being of communities that rely on these natural resources.

1.6 Discussion

The effectiveness of cloud seeding has been a topic of research and debate. Some studies have shown that cloud seeding can increase precipitation and activate convection, leading to enhanced rainfall and snowfall in experimental settings (Huggins, 2008; Zheng et al., 2020; Saito et al., 2017). Kulkarni shows that cloud seeding was effective in increasing rainfall in Karnataka State, with an average enhancement of 27.9% above natural rainfall (Kulkarni et al., 2019). However, the direct observational evidence for the cloud seeding effect is limited, and statistical proof of its effectiveness is lacking (Malik et al., 2018). In this paper, I use state-of-the-art causal inference methods to estimate the effects of 12 long-running cloud seeding projects in California on precipitation. I find that overall, cloud seeding increases monthly precipitation by 57% in the target areas, with significant differences in effectiveness between projects, ranging from 30% to 140%. These differences are partially attributed to topographical and climatological characteristics (Levin et al., 2010), as supported by the spatial analysis conducted in this paper. Early in the winter season, the rain falls more along the coast due to the convective activity produced by convergence resulting from the temperature gradient between the sea and the land. Later on, as this coastal gradient decreases, the clouds precipitate less over the coast and the orographic enhancement of the rainfall is more pronounced (Goldreich, 2003; Khain et al., 1993; Khain and Sednev, 1996). Therefore, differences in timing of storm days could contribute to differences in the estimated effects of cloud seeding.

Furthermore, the comparison between aircraft- and ground-based cloud seeding suggests that these two technologies, although highly cost-differentiated, offer the same effectiveness on average. The difference in performance between the two has been reduced due to technological advances in generators and flares, making ground generators highly effective, and reducing the need for aircraft interventions. However, depending on operational constraints, different projects might need to use either technology to reach maximum effectiveness. Research shows that there is still room for improvement of cloud seeding technology. Rosenfeld and authors showed the limitations and uncertainties of hygroscopic flares as a method of cloud seeding and proposes an alternative method using salt powder (Rosenfeld et al., 2010). In 2021, the first negative ion-based cloud seeding trial in China was effective in increasing rainfall in the trial area by 20% (Zheng et al., 2021). Cloud seeding using the shell structured TiO_2/NaCl aerosol has also been shown to enhance surface precipitation by more than 15% compared to traditional seeding with pure NaCl aerosol (Lompar et al., 2018). These advancements in technology could make cloud seeding more effective and reduce the costs of running precipitation enhancement programs.

Beyond the immediate effect of increasing precipitation, the benefits of cloud seeding include potential increases in surface water supplies, potential increases in the volume of groundwater stored, and potential increases in environmental and economic benefits (Batisha, 2012). In agriculture, cloud seeding can be modeled as a supply-increasing technology that is regulated by a State agency. When this technology increases precipitation in a region, it can potentially lead to an increase in production for farmers located in the affected region. Since irrigation water is a commodity with inelastic demand, the total revenue to farmers in this region will increase with an increase in the supply of water if their share of the total market is less than the coefficient of elasticity of demand (Buller et al., 1980). Moreover, the price effects of increased production can be distributed to all farmers in the market. Knowles and Skidmore find empirical evidence that supports the effectiveness of cloud seeding in increasing crop yields and reducing crop losses (Knowles and Skidmore, 2021). I contribute to this literature by estimating the average costs and benefits of providing additional usable water through cloud seeding. While the costs vary significantly between projects, a common conclusion is that they remain low compared to other technologies aimed at increasing supply, and well below the estimated benefits.

More importantly, this is the first paper to quantify the potential negative effects of cloud seeding on downwind areas. Cloud seeding affects the downwind area by redistributing precipitation (Zhen and Heng-Chi, 2010). Results from the Poisson QMLE suggest a statistically significant decrease of 8% in monthly precipitation for the Eastern Sierra project, which increases monthly precipitation in the target area by 22%. This opens a debate on the cost of externalities generated by precipitation enhancement programs. As it relates to the theoretical framework that motivates the empirical analysis, this result still suggests that cloud seeding could increase total welfare by relaxing the constraint on water supply, since the estimated increase outweighs the negative effect on downwind areas. However, this result only holds assuming the social planner optimizes welfare for both target and downwind areas. When different economic agents are making decisions for these areas, other questions arise. The empirical results of this study answer a pressing question about managing competition for limited water resources. In particular, this study shows that there is an opportunity for target HUCs in California to internalize the negative externalities they generate for downwind HUCs in Nevada by compensating them, financially, for the estimated $49 million loss due to decreased surface water every year.

Since clouds represent a limited common-pool resource, altering them to extract more precipitation by one user ought to generate externalities for other users. While this paper provides evidence that cloud seeding could have negative impacts on neighboring communities, other studies found this technology to have positive or no effects on the downwind areas. Wang and authors found that cloud seeding increased overall seasonal average rainfall in the downwind area by 21.67% and that this effect was detected as far as 120 km away from the target area (Wang et al., 2019). Miao and Geerts find that cloud seeding has a strong low-level reflectivity enhancement on the windward side of the mountain, but no effect in the lee of the mountain or downwind of the mountain crest (Miao and Geerts, 2013).

While concerns about externalities are crucial in water management, in this specific case, the potential negative impacts on downwind areas in Nevada might be less significant than initially anticipated. The HUCs in question experience limited agricultural activity due to the rain shadow effect, meaning that an 8% decrease in monthly precipitation due to cloud seeding might not translate to substantial economic losses in that sector. Additionally, the smaller population size in these areas suggests that municipal water demands might be lower, potentially making them less vulnerable to the decrease in water availability. However, it is crucial to emphasize that this does not diminish the importance of considering externalities entirely. Even seemingly minor reductions in water resources can have ripple effects on ecosystems, cultural values, and community well-being. Therefore, a comprehensive assessment that considers both the upstream benefits and potential downwind impacts, even if seemingly minimal, remains essential for making informed decisions about cloud seeding practices.

In conclusion, this paper evaluates 12 long-running cloud seeding programs in California using two estimation methods: extended TWFE with varying slopes for the estimation of changes in levels of precipitation, and Poisson QMLE for the estimation of changes in percentages, which allows for a comparison between the performance of different projects. With an average increase in monthly precipitation totals of 57%, this effectively increases the water supply at a cost of $8.65/ac-ft. Cloud seeding seems to be more effective in projects closest to the Sierra Nevada mountain rage and average costs of additional surface water seem to be higher in watersheds at lower altitudes. These findings have significant implications for policy and decision-making as they provide empirical evidence on many considerations to inform strategic investments in precipitation enhancement programs, including the potential benefits, costs, and externalities related to such projects. The region of southern California is at risk of severe drought if water supplies from northern California and the Colorado River were to diminish (Averyt et al., 2013). This paper suggests that cloud seeding could be used to mitigate such shortages. Consequently, this research contribution enriches the academic discourse on cloud seeding and furnishes pragmatic insights germane to sustainable, economically optimized, and regionally tailored climate adaptation and water resource management endeavors.

BIBLIOGRAPHY

Averyt, K., Meldrum, J., Caldwell, P., Sun, G., McNulty, S., Huber-Lee, A., and Madden, N. (2013). Sectoral Contributions to Surface Water Stress in the coterminous United States. *Environmental Research Letters*, 8(3):035046.

Batisha, A. F. (2012). Water resources sustainability in Egypt: Cloud seeding opportunity. *International Journal of Sustainability*.

Baumol, W. J. and Oates, W. E. (1988). *The theory of environmental policy*. Cambridge University Press.

Benjamini, Y., Givati, A., Khain, P., Levi, Y., Rosenfeld, D., Shamir, U., Siegel, A., Zipori, A., Ziv, B., and Steinberg, D. M. (2023). The Israel 4 cloud seeding experiment: Primary results. *Journal of Applied Meteorology and Climatology*, 62(3):317–327.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Boyle, D. P., Lamorey, G. W., and Huggins, A. W. (2006). Application of a hydrologic model to assess the effects of cloud seeding in the Walker River Basin of Nevada. *The Journal of Weather Modification*, 38(1):66–76.

Bradley, R. A., Srivastava, S. S., and Lanzdorf, A. (1978). *An Examination of the Effects of Cloud Seeding in Phase I of the Santa Barbara Convective Band Seeding Test Program*. Florida State University, Department of Statistics.

Breed, D., Rasmussen, R., Weeks, C., Boe, B., and Deshler, T. (2014). Evaluating winter orographic cloud seeding: Design of the Wyoming Weather Modification Pilot Project (WWMPP). *Journal of Applied Meteorology and Climatology*, 53(2):282–299.

Buller, O. H., Bark, D. L., and Vanderlip, R. (1980). Effect of weather modification on supply and total revenue of a region. *Western Journal of Agricultural Economics*, 6(1):57–68.

Bulman, T. L. and Shelton, M. L. (1991). Cloud seeding in California: An evaluation of the rainshadow effect of weather modification projects. *California Geographical Society*.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Chudik, A. and Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, 188(2):393–420.

Chudik, A., Pesaran, M. H., and Tosetti, E. (2011). Weak and strong cross-section dependence and estimation of large panels.

Coase, R. H. (1960). The problem of social cost. *The Journal of Law and Economics*, 3:1–44.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

DeFelice, T., Golden, J., Griffith, D., Woodley, W., Rosenfeld, D., Breed, D., Solak, M., and Boe, B. (2014). Extra area effects of cloud seeding—an updated assessment. *Atmospheric Research*, 135:193–203.

Dupuit, J. (1844). On the measurement of the utility of public works. *International Economic Papers*, 2(1952):83–110.

Friedrich, K., Ikeda, K., Tessendorf, S. A., French, J. R., Rauber, R. M., Geerts, B., Xue, L., Rasmussen, R. M., Blestrud, D. R., Kunkel, M. L., et al. (2020). Quantifying snowfall from orographic cloud seeding. *Proceedings of the National Academy of Sciences*, 117(10):5190–5195.

Gabriel, K. R. (2000). Parallels between statistical issues in medical and meteorological experimentation. *Journal of Applied Meteorology and Climatology*, 39(11):1822–1836.

Gagin, A. and Neumann, J. (1981). The second Israeli randomized cloud seeding experiment: Evaluation of the results. *Journal of Applied Meteorology and Climatology*, 20(11):1301–1311.

Goldreich, Y. (2003). *The Climate of Israel: Observation, Research and Application*. Springer.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Griffith, D., Yorty, D., and Beall, S. (2015). Target/control analyses for Santa Barbara County's operational winter cloud seeding program. *The Journal of Weather Modification*, 47(1):10–10.

Griffith, D. A., Solak, M. E., Almy, R. B., and Gibbs, D. (2005). The Santa Barbara cloud seeding project in coastal southern California, summary of results and their implications. *The Journal of Weather Modification*, 37(1):21–27.

Huggins, A. W. (2008). Summary of studies that document the effectiveness of cloud seeding for snowfall augmentation. *The Journal of Weather Modification*.

Khain, A. P., Rosenveld, D., and Sednev, I. (1993). Coastal effects in the Eastern Mediterranean as seen from experiments using a cloud ensemble model with detailed description of warm and ice microphysical processes. *Atmospheric research*, 30(4):295–319.

Khain, A. P. and Sednev, I. (1996). Simulation of precipitation formation in the Eastern Mediterranean coastal zone using a spectral microphysics cloud ensemble model. *Atmospheric Research*, 43(1):77–

110.

Knowles, S. and Skidmore, M. (2021). Cloud seeding and crop yields: Evaluation of the North Dakota Cloud Modification Project. *Weather, Climate, and Society*.

Kulkarni, J., Morwal, S., and Deshpande, N. (2019). Rainfall enhancement in Karnataka State Cloud Seeding Program "Varshadhare". *Atmospheric Research*.

Levin, Z., Halfon, N., and Alpert, P. (2010). Reassessment of rain enhancement experiments and operations in Israel including synoptic considerations. *Atmospheric Research*, 97(4):513–525.

Lompar, M., Ćurić, M., Romanic, D., Zou, L., and Liang, H. (2018). Precipitation enhancement by cloud seeding using the shell structured TiO2/NaCl aerosol as revealed by new model for cloud seeding experiments. *Atmospheric Research*, 212:202–212.

Malik, S., Bano, H., Rather, R. A., and Ahmad, S. (2018). Cloud seeding; its prospects and concerns in the modern world-a review. *Int. J. Pure App. Biosci*, 6(5):791–796.

Manton, M. J. and Warren, L. (2011). A confirmatory snowfall enhancement project in the Snowy Mountains of Australia. part ii: Primary and associated analyses. *Journal of Applied Meteorology and Climatology*, 50(7):1448–1458.

Miao, Q. and Geerts, B. (2013). Airborne measurements of the impact of ground-based glaciogenic cloud seeding on orographic precipitation. *Advances in Atmospheric Sciences*, 30:1025–1038.

Morrison, A. E., Siems, S. T., Manton, M. J., and Nazarov, A. (2010). A modeling case study of mixed-phase clouds over the Southern Ocean and Tasmania. *Monthly weather review*, 138(3):839–862.

Neyman, J., Scott, E. L., and Vasilevskis, M. (1960). Statistical evaluation of the Santa Barbara randomized cloud seeding experiment. *Bulletin of the American Meteorological Society*, 41(10):531–547.

Ostrom, E., Gardner, R., and Walker, J. (1994). *Rules, Games, and Common-Pool Resources*. University of Michigan press.

Pigou, A. C. (1920). *The Economics of Welfare*, volume 1. Palgrave Macmillan London.

Plott, C. and Meyer, R. (1975). The technology of public goods, externalities, and the exclusion principle. In *Economic Analysis of Environmental Problems*, pages 65–94. NBER.

Rauber, R. M., Geerts, B., Xue, L., French, J., Friedrich, K., Rasmussen, R. M., Tessendorf, S. A., Blestrud, D. R., Kunkel, M. L., and Parkinson, S. (2019). Wintertime orographic cloud seeding—a review. *Journal of Applied Meteorology and Climatology*, 58(10):2117–2140.

Research Council, N., on Earth, D., Studies, L., on Atmospheric Sciences, B., on the Status of, C., and in US Weather Modification Research, F. D. (2004). *Critical Issues in Weather Modification Research*. National Academies Press.

Rogers, J. S. and Gahan, A. H. (2009). Project Skywater. *Historic Reclamation Projects, Tehran*.

Rosenfeld, D., Axisa, D., Woodley, W. L., and Lahav, R. (2010). A quest for effective hygroscopic cloud seeding. *Journal of Applied Meteorology and Climatology*, 49(7):1548–1562.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*.

Ryan, B. F. and King, W. D. (1997). A critical review of the Australian experience in cloud seeding. *Bulletin of the American Meteorological Society*, 78(2):239–254.

Saito, T., Sashihara, R., Komatsuzaki, K., and Tomine, K. (2017). Numerical experiments on glaciogenic artificial cloud seeding conducted in a stable cloud. *African Journal of Engineering Research*, 5(3):54–82.

Schaefer, V. J. (1946). The production of ice crystals in a cloud of supercooled water droplets. *Science*, 104(2707):457–459.

Silverman, B. A. (2007). On the use of ratio statistics for the evaluation of operational cloud seeding programs. *The Journal of Weather Modification*, 39(1):50–60.

Silverman, B. A. (2008). A statistical evaluation of the Kern River Operational Cloud Seeding Program. *The Journal of Weather Modification*, 40(1):7–16.

Silverman, B. A. (2009). An evaluation of the San Joaquin Operational Cloud Seeding Program using Monte Carlo permutation statistics. *The Journal of Weather Modification*, 41(1):15–22.

Silverman, B. A. (2010). An evaluation of eleven operational cloud seeding programs in the watersheds of the Sierra Nevada Mountains. *Atmospheric Research*, 97(4):526–539.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Sun, L. and Shapiro, J. M. (2022). A linear panel model with heterogeneous coefficients and variation in exposure. *Journal of Economic Perspectives*, 36(4):193–204.

Vegas, C. N. L. (2024). Water and sewer rates in the City of North Las Vegas. Technical report.

Vonnegut, B. (1947). The nucleation of ice formation by silver iodide. *Journal of Applied Physics*, 18(7):593–595.

Wang, W., Yao, Z., Guo, J., Tan, C., Jia, S., Zhao, W., Zhang, P., and Gao, L. (2019). The extra-area effect in 71 cloud seeding operations during winters of 2008–14 over Jiangxi Province, East China. *Journal of Meteorological Research*, 33(3):528–539.

WGISC (2013). Subbasin Boundaries, California (8-digit HUC), 2012. Available at: http://purl.stanford.edu/sr396hp9621.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. Available at SSRN 3906345.

Wooldridge, J. M. (2023). Simple approaches to nonlinear difference-in-differences with panel data. *The Econometrics Journal*, 26(3):C31–C66.

Zhen, Z. and Heng-Chi, L. (2010). Numerical simulation of seeding extra-area effects of precipitation using a three-dimensional mesoscale model. *Atmospheric and Oceanic Science Letters*, 3(1):19–24.

Zheng, W., Fengming, X., Zhang, M., Qiqi, W., Zhou, Y., Shaoxiang, M., Haotian, L., Chuliang, W., Yuxing, W., Ai, X., Yong, Y., and Kexun, Y. (2020). Charged particle (negative ion)-based cloud seeding and rain enhancement trial design and implementation. *Water*, 12(6):1644.

Zheng, W., Hengben, M., Zhang, M., Fengming, X., Kexun, Y., Yong, Y., Shaoxiang, M., Chuliang, W., Yuan, P., Zhiliang, S., Jianhua, M., Wenqing, Y., and Xianzhi, Y. (2021). Evaluation of the first negative ion-based cloud seeding and rain enhancement trial in China. *Water*, 13(18):2473.

APPENDIX 1A

FIGURES

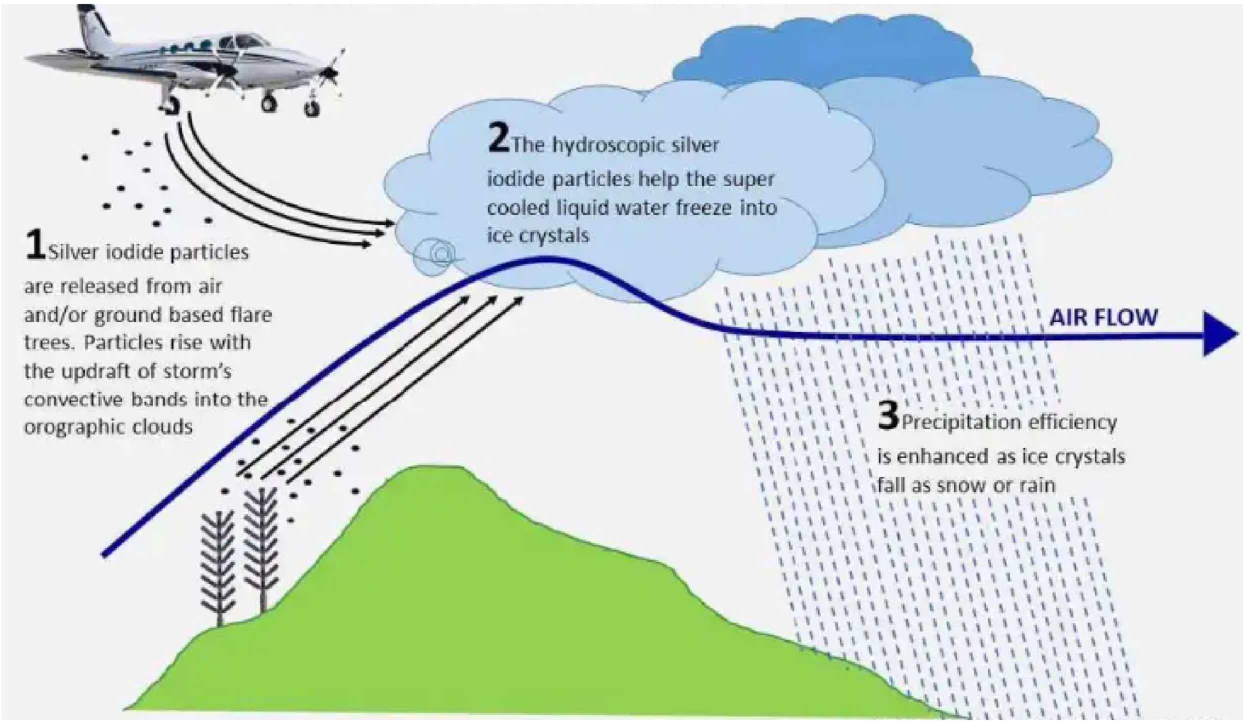

Figure 1A.1 Cloud seeding technology diagram. Source: Santa Barbara County, California, US

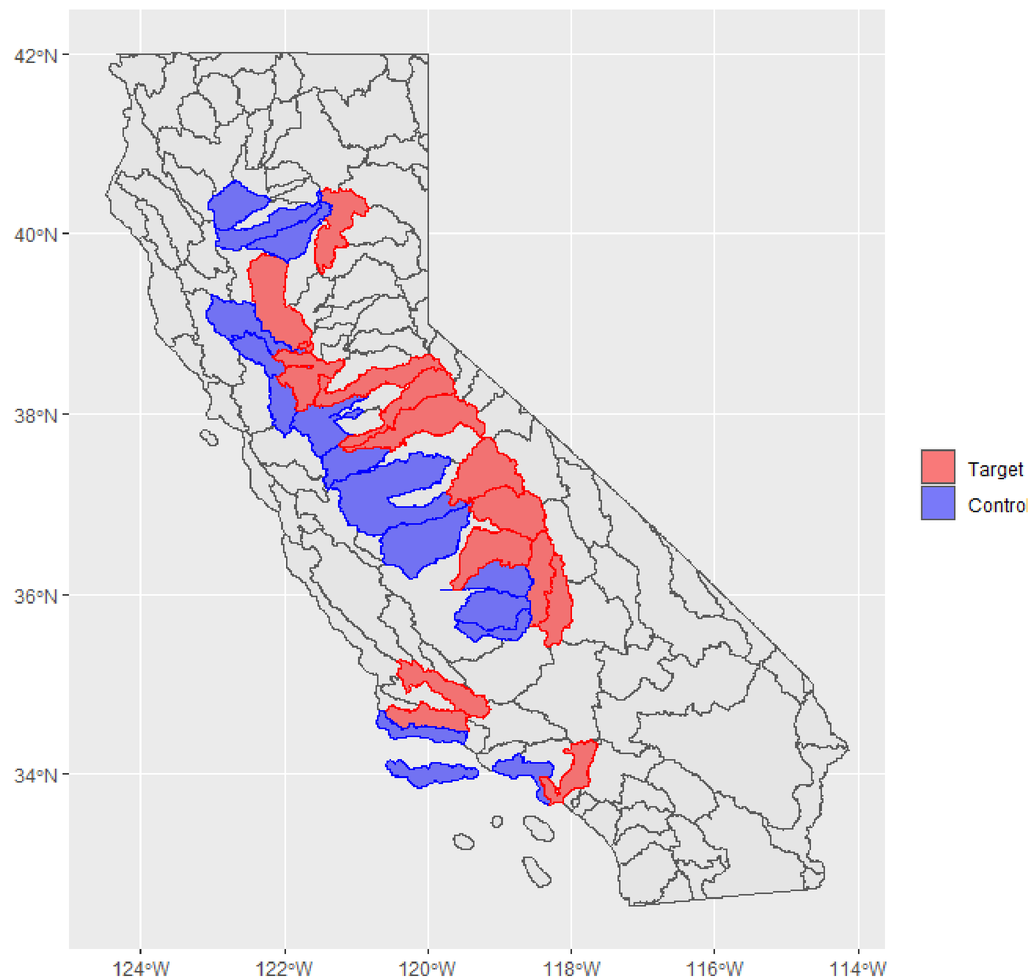

Figure 1A.2 Cloud seeding target and control 8-digit HUCs, California, US

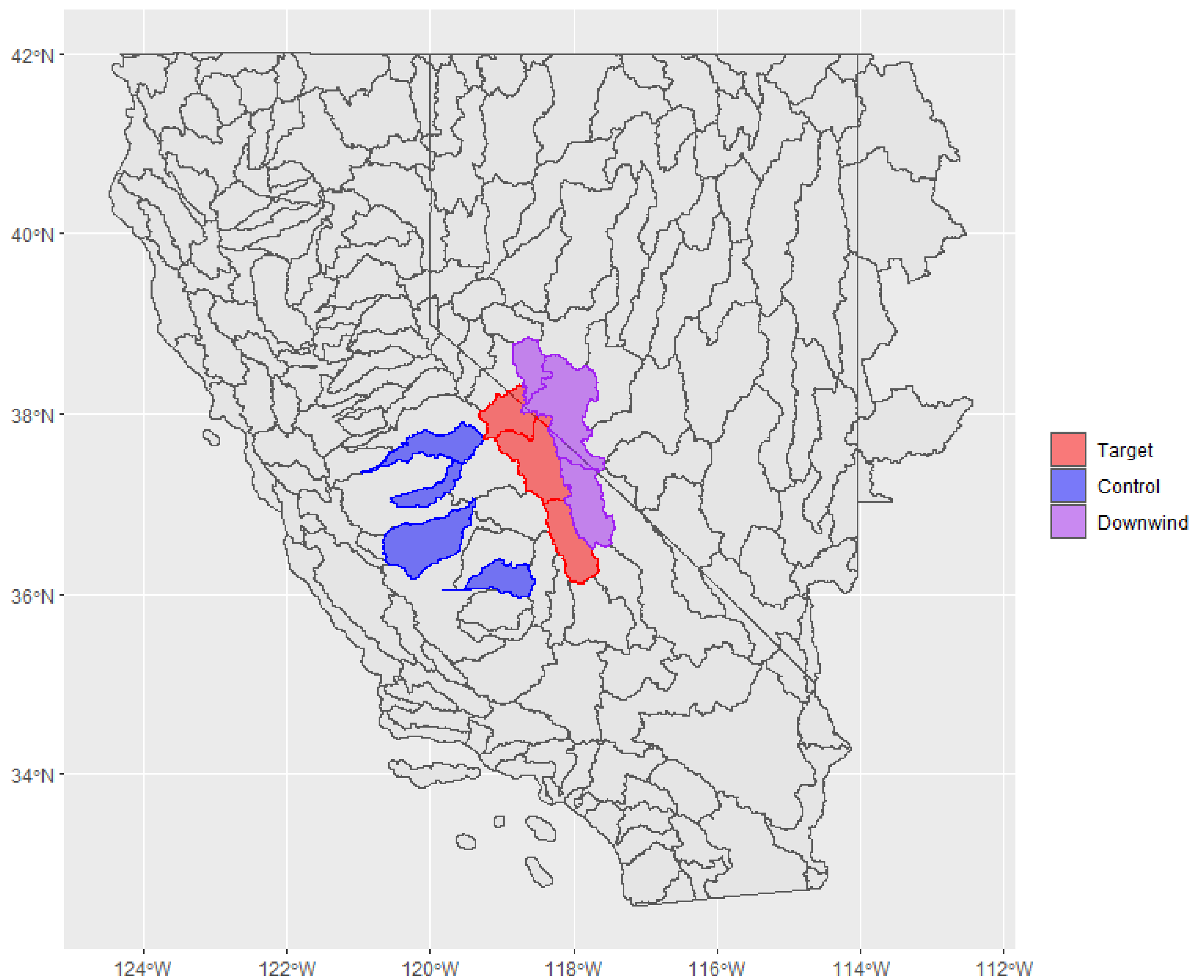

Figure 1A.3 Eastern Sierra Cloud seeding 8-digit HUCs, California and Nevada, US

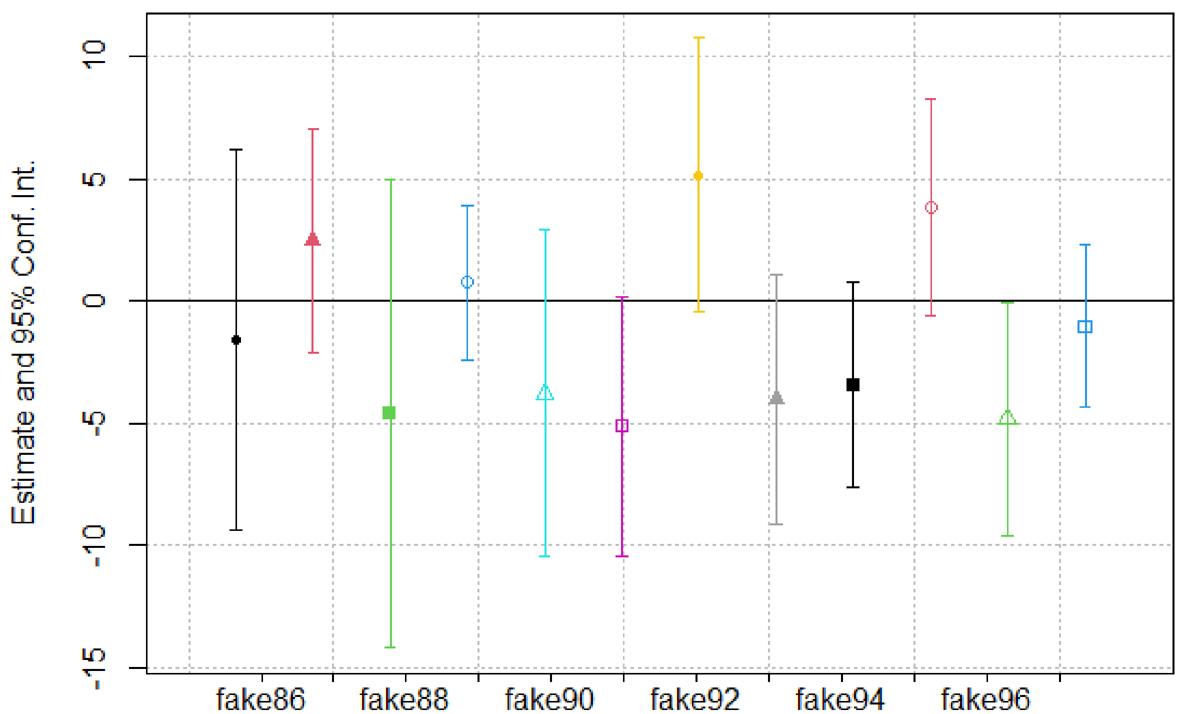

Figure 1A.4 Placebo test visualization representing the the average treatment effect of multiple "fake treatment" assignments on precipitation in mm per month

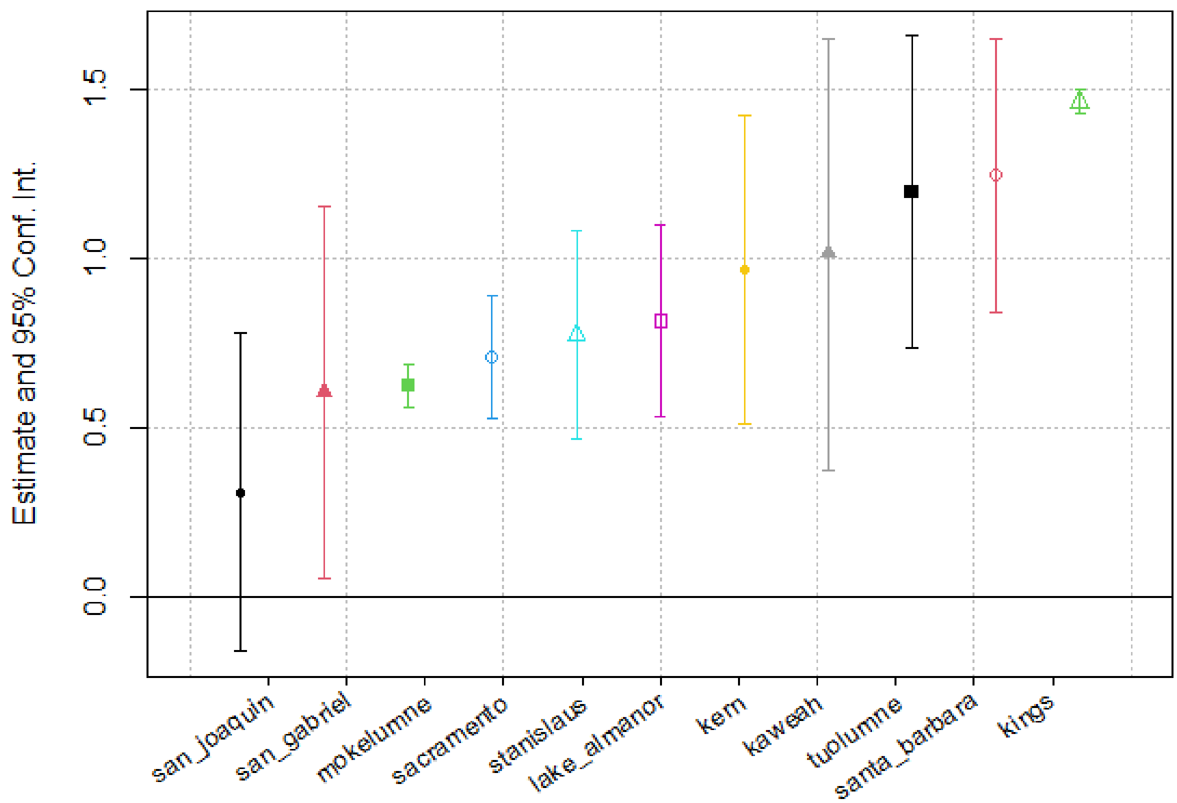

Figure 1A.5 Effects of Cloud Seeding on Precipitation (% per month), ordered from lowest to highest

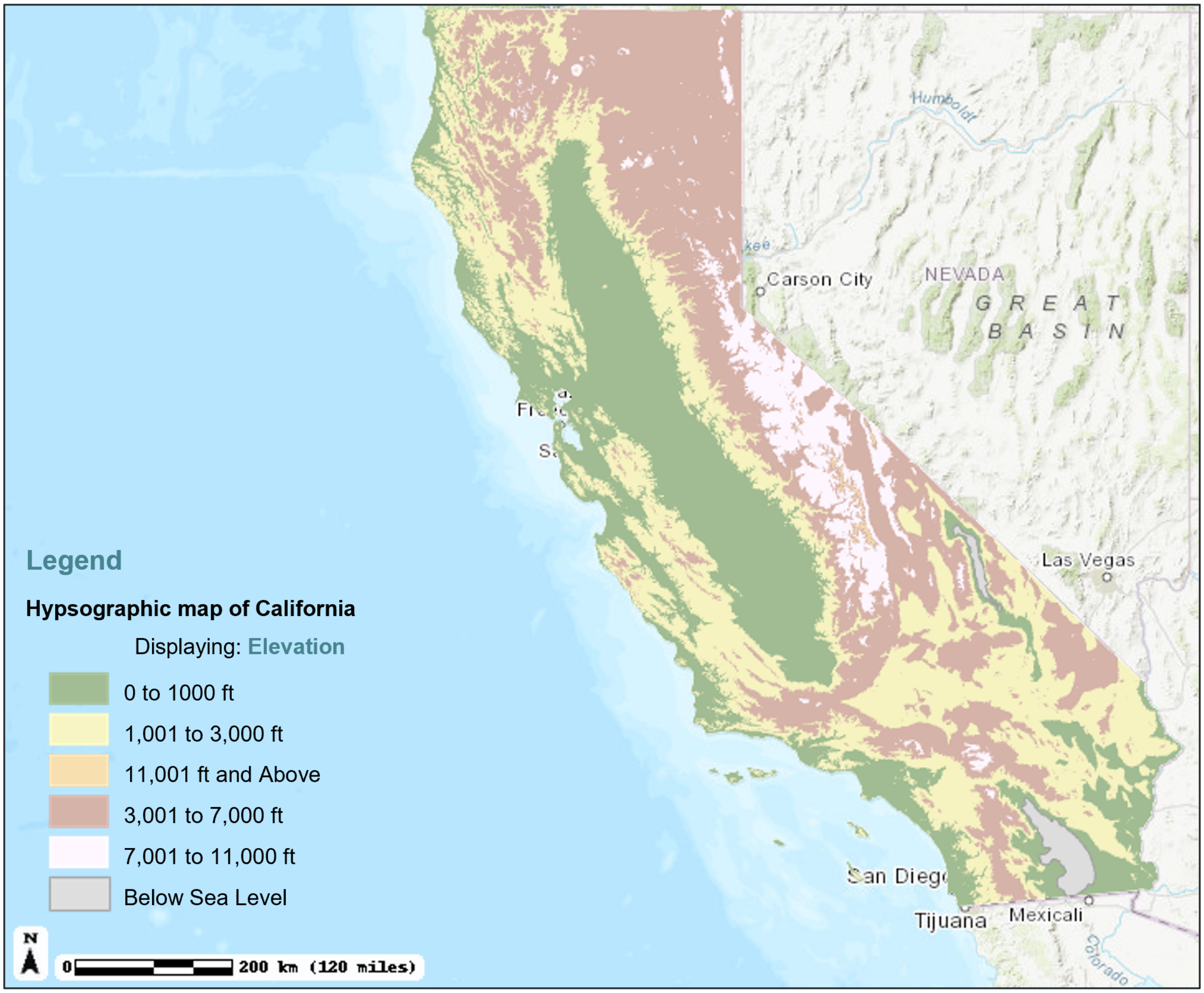

Figure 1A.6 Hypsographic map of California, US. Source: Data Basin.

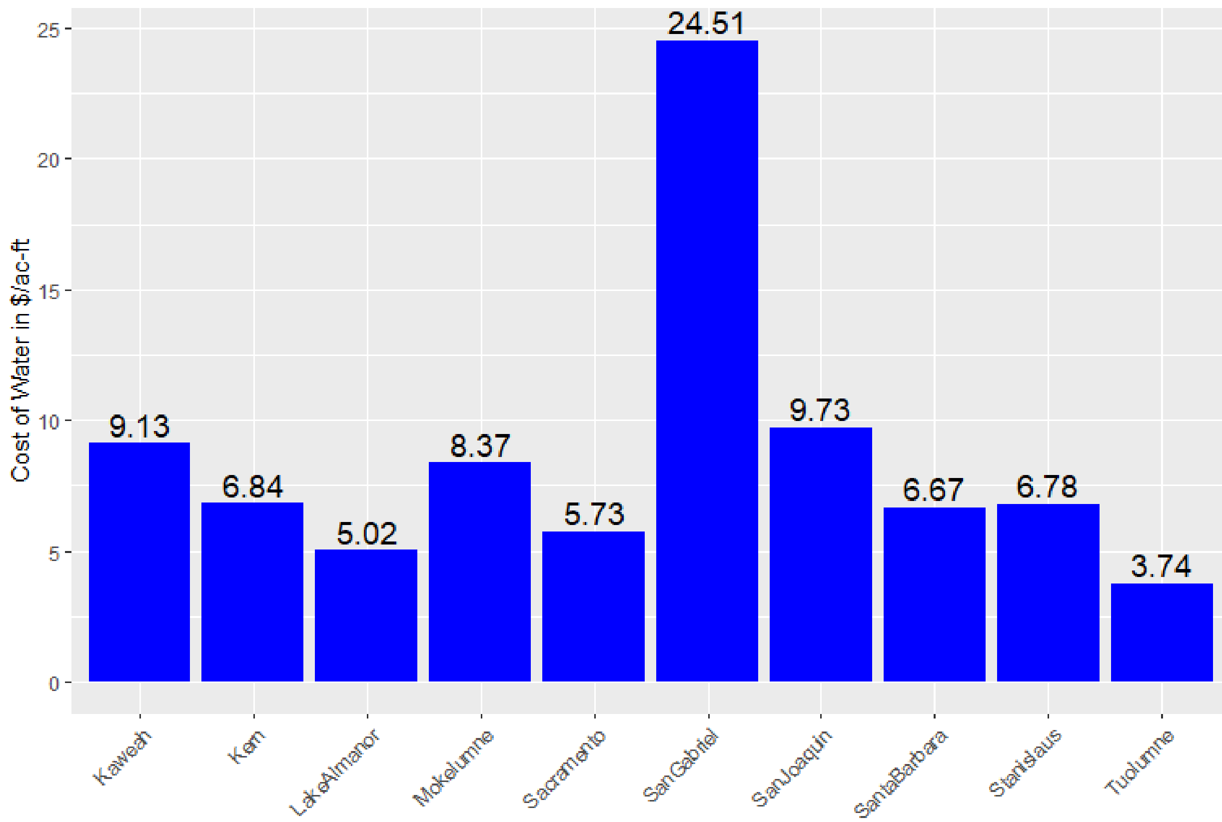

Figure 1A.7 Average unit cost ($/ac-ft) of additional surface water by project

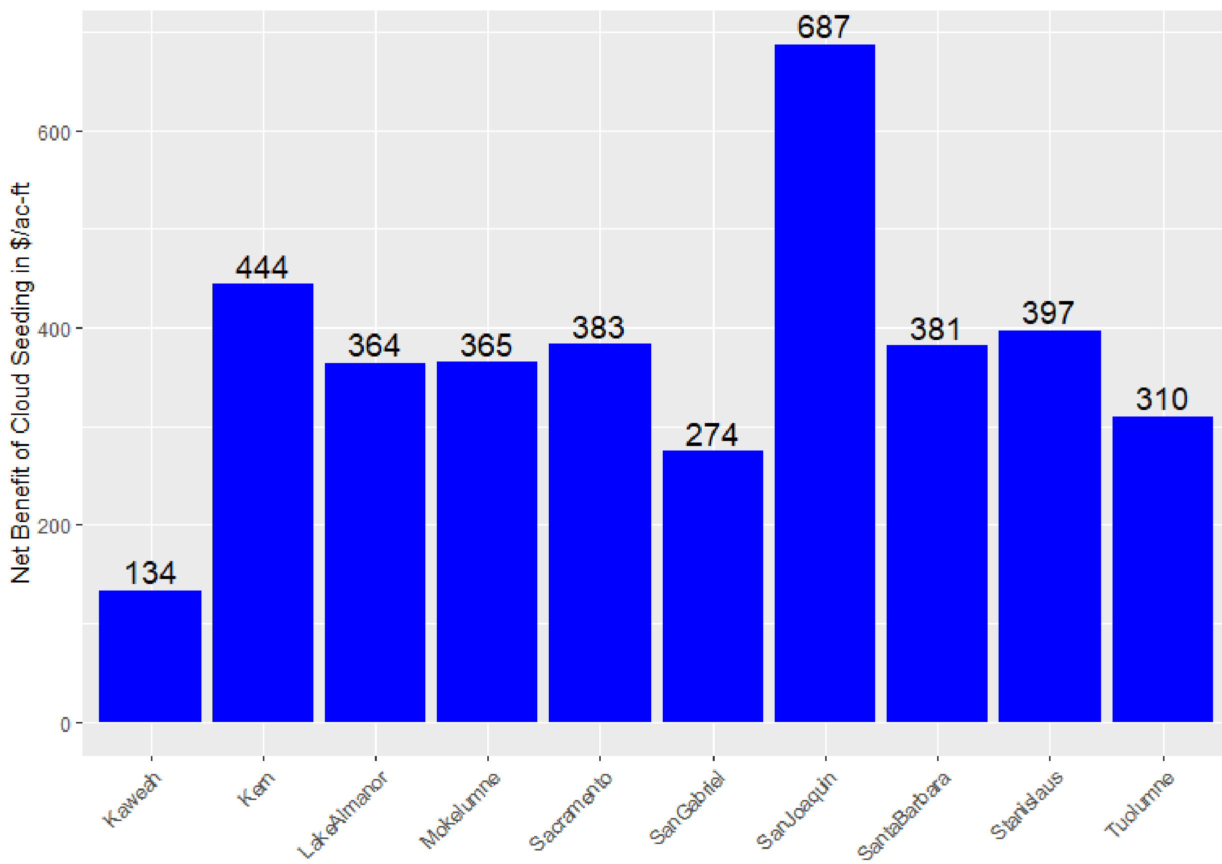

Figure 1A.8 Average Net Benefit ($/ac-ft) of additional water by project

APPENDIX 1B

TABLES

Table 1B.1 Sample cohorts

Cohort	Projects
1999	Sacramento
2000	Kern, Kaweah, Kings, Lake Almanor, Tuolumne, Mokelumne, San Joaquin
2001	San Gabriel, Santa Barbara
2007	Stanislaus

Table 1B.2 Average Monthly Precipitation Totals (mm) before (Pre: 1981-1999) and after (Post: 2000-2022) Cloud Seeding in Target and Control Hydrologic Units by Seeded and Unseeded Months, standard errors in parentheses

HUC	Pre	Post + unseeded	Post + seeded
Target	67.67 (96.16)	43.14 (74.14)	102.09 (98.15)
Control	49.48 (77.46)	41.42 (65.88)	41.42 (65.88)

Table 1B.3 Change in Average Monthly Precipitation (mm), Full Sample

Dependent Variable:	Precipitation
Cloud seeding	51.96***
	(8.313)
Fixed-effects	
Site	Yes
Month	Yes
Fit statistics	
Observations	15,120
Adjusted R^2	0.15512

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.4 Change in Average Monthly Precipitation (mm), Restricted Sample (precip>0)

Dependent Variable:	Precipitation
Cloud seeding	50.67***
	(8.184)
Fixed-effects	
Site	Yes
Month	Yes
Fit statistics	
Observations	14,904
Adjusted R^2	0.15460

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.5 Change in Average Monthly Precipitation (mm) by Technology of Cloud Seeding, Full Sample

	Precipitation		
Model:	(1)	(2)	(3)
method_airborne	40.74***		26.72***
	(7.470)		(9.563)
method_ground		40.89***	29.62***
		(4.783)	(7.520)
Fixed-effects			
Site	Yes	Yes	Yes
Month	Yes	Yes	Yes
Fit statistics			
Observations	11,934	11,937	12,225
Adjusted R^2	0.16263	0.16773	0.17409

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.6 Change in Average Monthly Precipitation (mm) by Technology of Cloud Seeding, Restricted Sample (precip>0)

Dependent Variable:	Precipitation
method_airborne	27.47***
	(9.382)
method_ground	29.22***
	(7.525)
Fixed-effects	
Site	Yes
Month	Yes
Fit statistics	
Observations	12,220
Adjusted R^2	0.17433

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.7 Changes in Average Monthly Precipitation (mm) as Fixed Effects are introduced

	Precipitation		
Model:	(1)	(2)	(3)
constant	49.25***		
	(4.004)		
Cloud seeding	52.15***	44.20***	51.96***
	(10.83)	(7.607)	(8.313)
Fixed-effects			
Site		Yes	Yes
Month			Yes
Fit statistics			
Observations	15,120	15,120	15,120
R^2	0.02292	0.10004	0.15909

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.8 Change in Average Monthly Precipitation (mm) with Covariates

	Precipitation			
Model:	(1)	(2)	(3)	(4)
Cloud seeding	51.96***	29.36***	17.82***	18.74***
	(8.313)	(6.934)	(6.370)	(6.334)
tdmean		-7.316***		0.8080
		(0.5443)		(1.084)
vpdmean			-6.656***	-6.963***
			(0.6836)	(0.8560)
Fixed-effects				
Site	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
Fit statistics				
Observations	15,120	15,120	15,120	15,120
Adjusted R^2	0.15512	0.25735	0.41501	0.41566

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.9 Change in Average Monthly Precipitation (mm) by Intensity of Cloud Seeding

	Precipitation			
Model:	(1)	(2)	(3)	(4)
NumberDays	13.56***			
	(1.675)			
HoursAirborne		6.279***		
		(1.026)		
HoursGround			0.176***	
			(0.0164)	
AmountSilverIodide_kg				10.16***
				(0.7697)
Fixed-effects				
Site	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
Fit statistics				
Observations	15,120	15,120	15,120	15,120
Adjusted R^2	0.17323	0.15566	0.15679	0.16976

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.10 Change in Average Monthly Precipitation (mm) using Heterogeneous Slopes

Dependent Variable:	Precipitation
Cloud seeding	50.65***
	(8.55)
Fixed-effects	
Month	Yes
Varying Slopes	
Site_num (Month)	Yes
Fit statistics	
Observations	15,120
Adjusted R^2	0.1041

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.11 Change in Average Monthly Precipitation (%), Full Sample

Dependent Variable:	Precipitation
Cloud seeding	0.574***
	(0.113)
Fixed-effects	
Site	Yes
Month	Yes
Fit statistics	
Observations	15,120
Squared Correlation	0.10129
Pseudo R^2	0.12235
BIC	1,255,328.3

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.12 Change in Average Monthly Precipitation (%), Restricted Sample (precip>0)

Dependent Variable:	Precipitation
Cloud seeding	0.555***
	(0.109)
Fixed-effects	
Site	Yes
Month	Yes
Fit statistics	
Observations	14,904
Squared Correlation	0.10115
Pseudo R^2	0.12228
BIC	1,235,469.5

Clustered (site & year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.13 Change in Average Monthly Precipitation (%) for the projects Kern, Kings, Lake Alamor, Sacramento, and Tuolumne.

	Precipitation				
Model:	Kern	Kings	L.Almanor	Sacramento	Tuolumne
Cloud seeding	0.969***	1.465***	0.817***	0.712***	1.197***
	(0.233)	(0.018)	(0.137)	(0.129)	(0.236)
Fixed-effects					
Site	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	1,872	504	1,728	2,340	1,008
Squared Correlation	0.20950	0.23998	0.13025	0.09630	0.27783
Pseudo R^2	0.20575	0.23918	0.13192	0.11362	0.27570
BIC	90,312.4	43,858.1	180,942.0	207,007.4	67,709.1

Clustered (year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.14 Change in Average Monthly Precipitation (%) for the projects Kaweah, San Gabriel, San Joaquin, Stanislaus, and Mokelumne.

	Precipitation				
Model:	Kaweah	S.Gabriel	S.Joaquin	Stanislaus	Mokelumne
Cloud seeding	1.014***	0.607***	0.311	0.776***	0.624***
	(0.309)	(1×10^{-5})	(0.227)	(0.131)	(0.032)
Fixed-effects					
Site	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	1,008	864	840	1,008	552
Squared Correlation	0.13727	0.08811	0.21529	0.21385	0.15912
Pseudo R^2	0.12855	0.11747	0.27283	0.24339	0.17311
BIC	58,654.4	58,863.7	57,062.4	75,555.2	40,448.4

Clustered (year) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.15 Change in Average Monthly Precipitation (% per month)

Poisson QMLE	Precipitation
Cloud seeding	-0.079***
	(0.029)
Fixed-effects	
Site	Yes
Month (site)	Yes
Fit statistics	
Observations	2,100
Adj. Pseudo R^2	0.201303

Clustered (site) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1B.16 Costs and benefits of the additional surface water produced by cloud seeding

Project	Volume (ac-ft)	Cost ($/ac-ft)	Benefit ($/ac-ft)	Net Benefit ($/ac-ft)
Kaweah	547431	9.13	143	134
Kern	731448	6.84	451	444
Lake Almanor	995448	5.02	369	364
Mokelumne	597463	8.37	373	365
Sacramento	872283	5.73	389	383
San Gabriel	204029	24.51	299	274
San Joaquin	513989	9.73	697	687
Santa Barbara	749769	6.67	388	381
Stanislaus	737119	6.78	404	397
Tuolumne	1335599	3.74	314	310

CHAPTER 2

EVENT-BASED ANALYSIS OF CLOUD SEEDING IN SOUTHERN CALIFORNIA

2.1 Introduction

Water availability is determined by the balance between supply and demand, with agriculture, thermoelectric power generation, and municipal requirements being the main drivers of freshwater demands. Future water balances are uncertain in many regions of the world due to various factors such as economic conditions, social behaviors, technological innovations, legal and policy drivers, increasing water demand, and climate change. Watersheds in the western United States (U.S.) are particularly vulnerable to low-flow events and projected long-term shifts in flow driven by climate change. Although agriculture is the main driver on the demand side of water stress in the western United States, decreased municipal water supply is predominant in southern California, indicating specific vulnerability in this region (Averyt et al., 2013).

Urban development and reliance on water-intensive crops have strained existing water infrastructure, prompting the need for sustainable and innovative water management strategies. Efforts to address this scarcity have focused on improving water use efficiency and conservation, as well as exploring the role of water markets in managing water allocation and demand (Schwabe et al., 2020). One potential solution to mitigate the impacts of water scarcity from the supply side is to implement precipitation enhancement programs which aim to increase the amount of rainfall using cloud seeding over a specific target area (Lachhab, 2024). Cloud seeding is especially important to alleviate the impact from drought (Ma et al., 2023).

Cloud seeding, which is the practice of introducing microscopic particles into clouds to promote snowflake formation and precipitation, has been pursued around the world for decades. While widely implemented, its effectiveness in boosting water supplies remains shrouded in uncertainty. Early studies, while paving the way for understanding the conditions conducive to cloud seeding success, fell short of quantifying the actual increase in precipitation it could induce (Malik et al., 2018). Despite the uncertainty, weather modification has continued in many arid regions in an effort to meet the increasing demand for water and because of the large potential cost benefit for production of additional water by cloud seeding (Rauber et al., 2019), especially compared to the cost of other water supply management techniques such as groundwater banking and desalination plants.

Experiments in Wyoming, North Dakota, and California reported an increase of 5-15% in rainfall (Acharya et al., 2011; Solak et al., 1987; Griffith et al., 2007), while experiments in Tasmania and Australia showed a 5-30% increase in rainfall (Morrison et al., 2009; Beare et al., 2010). Runoff[1] increase was linked to precipitation enhancement from cloud seeding by Yoo et al. (2022), who showed that most of the increased rainfall contributed to increased runoff.

However, the effectiveness of cloud seeding is contingent on many factors such as basin characteristics, catchment area size, reservoir volume, technology, and storm systems. The uncertainty surrounding the effectiveness of cloud seeding in boosting precipitation remains a contentious issue. This uncertainty has led to the closure of several programs, including a notable example in Israel. After 38 years of operation, the Israel Water Authority discontinued its cloud seeding program in 2021 due to underwhelming results. A key factor in this decision was a new randomized experiment conducted from 2013 to 2020. This study, designed to definitively assess the impact of cloud seeding on rainfall, yielded statistically insignificant results, suggesting an average increase of only 1.8%. Notably, this finding aligns with earlier concerns expressed by Levin et al. (2010), which previously raised questions about the program's effectiveness.

Cloud seeding projects often face scarce resources in terms of time, effort, and materials. Determining the optimal allocation of these resources is crucial for maximizing the project's net benefits. This chapter explores the trade-offs involved in allocating resources, such as deciding whether to focus on specific areas or employ a broader approach and choosing between ground-based or aerial seeding. Cloud seeding effectiveness can vary significantly between individual events, even within the same project. Factors like cloud characteristics, wind patterns, and seeding timing can influence the outcome of each operation. Analyzing individual events helps with gaining a more nuanced understanding of the factors influencing success and can identify potential areas for improvement in cloud seeding practices.

[1]Runoff refers to the proportion of rainfall that is not lost due to infiltration, percolation, evaporation, or evapotranspiration. It is thus the "useful" part of rainfall.

Building upon the foundation established in the first chapter's examination of 12 California cloud seeding projects, this chapter focuses on a unique, high-resolution analysis of a single project situated in Santa Barbara County, Southern California. This shift in focus offers several key distinctions, including (1) enhanced data granularity: while the first essay utilized monthly data on precipitation and cloud seeding operations offering a regional-level understanding, this essay leverages daily precipitation data alongside detailed cloud seeding records. This finer temporal resolution enables a more careful exploration of the immediate impacts of cloud seeding interventions on precipitation patterns within the project area, (2) scale and geographic focus: while the first essay explored a diverse range of projects spanning the breadth of California, this chapter narrows its lens to a specific coastal location within Santa Barbara County. This shift allows for a more in-depth analysis of the interactions between cloud seeding, local meteorological and hydrological characteristics, and precipitation patterns in this distinct coastal environment. Another distinction is the spatial resolution. Unlike the first chapter's analysis at the 8-digit hydrological unit level, this chapter benefits from precise knowledge of the target areas for seeding and the exact locations of ground generators used for burning flares. This level of spatial detail allows for a further examination of the spatial distribution of seeding effects and their potential impact on specific locations within the project area, as well as a deeper understanding of the targeting strategies and their potential influence on precipitation outcomes. Moreover, unlike the first essay, which relied on the National Oceanic and Atmospheric Administration (NOAA) cloud seeding data, this chapter benefits from highly detailed annual reports provided by the company directly implementing the cloud seeding operations. The access to precise locational information and daily data enables helps with understanding the operational aspects of cloud seeding in greater detail. The controlled environment of a single project allows for a more rigorous assessment of the causal relationship between cloud seeding and precipitation, minimizing the potential influence of external factors that might be present in a regional analysis.

In this paper, I focus on the program run by Santa Barbara County in the Upper Santa Ynez and the Huasna-Alamo watersheds. I combine data on cloud seeding operations spanning the period

1981-2023 from the weather modification project reports submitted to NOAA with operations reports from the Santa Barbara County Water Agency (SBCWA) and weather data from a network of precipitation gauges across the Santa Barbara County and the Oregon State PRISM project. The cloud seeding data consist of dates and duration of cloud seeding events, and the quantity of seeding agent used. The weather data consist of daily precipitation totals from gauges located within the targeted and control areas and span the period 1974-2023.

The resulting estimates suggest that the Santa Barbara weather modification program increases precipitation in the target areas by 0.26 inches (45% at the average) per event and that using more seeding material increases precipitation even further. Results also suggest that cloud seeding produces additional water resources at a cost of $5.6 per acre-foot. This cost is to be compared to the price of agricultural water, ranging from $45/act-ft to $600/ac-ft depending on the use and on water resource availability in a particular season. Hence, cloud seeding is dramatically more cost effective than other forms of agricultural water. Furthermore, our analysis shows that results obtained from a specific target area might not be extrapolated to other areas. This is true even for adjacent and seemingly similar catchment areas, like shown by the case of the Santa Barbara Precipitation Enhancement Project. The effect of cloud seeding from the same program is 6% lower and less statistically significant for Upper Santa Ynez than it is in Huasna Alamo. For policymakers, this means that results from other projects might not be relevant to inform the decision to invest (or not) in new weather modification projects.

Addressing a critical gap in both economic and scientific literature, this essay offers a framework to analyze the economic feasibility of cloud seeding programs as a potential alternative among water management strategies. This study breaks new ground in existing research by being the first to compare the effectiveness of ground-based flare seeding and aerial seeding methods at the project level. Its findings hold substantial relevance for local water management agencies, offering valuable information to guide informed decision-making about the optimal cloud seeding strategy for their specific needs and resource constraints. Furthermore, this analysis complements the broader regional perspective offered in the first chapter, ultimately contributing to a more comprehensive understanding of the potential and limitations of cloud seeding across diverse contexts.

The remainder of the manuscript is organized as follows: Section 2 briefly overviews cloud seeding in southern California and research efforts surrounding the effectiveness of the technique. Section 3 presents the data and empirical strategy. Results are summarized in Section 4, followed by a brief discussion on the implication of the results in Section 5. Section 6 draws conclusions from this study.

2.2 Background on Cloud Seeding in Southern California

Cloud seeding trials were initially carried out by Santa Barbara County in the 1950s, with subsequent experiments introduced in the 1970s to assess the potential benefits of cloud seeding for the coastal region of Southern California. Due to positive findings, cloud seeding has been conducted in Santa Barbara during most winter rainy seasons since 1981. Santa Barbara has targeted two primary sites, namely the Upper Santa Ynez region in eastern Santa Barbara County and the Twitchell Reservoir in northern Santa Barbara County and southern San Luis Obispo County.

While the initial trials showed some experimental evidence that cloud seeding may be effective, results thus far have largely not shown statistical significance. For example, the first Santa Barbara experiment from 1957-1960 found increases in precipitation of 45%, but results were not significant. In the second set of experiments from 1967-1973, some regions found statistically significant increases in precipitation: namely along the back of the Santa Ynez mountains, which experience higher natural precipitation. This study indicated that convection bands contributed approximately 50% of the total winter precipitation in this area. Two reports from North American Weather Consultants, Inc (NAWC) (Thompson and Griffith, 1988; Solak et al., 1996) provided a more precise quantification of the optimal seasonal seeding increases that might be expected at Juncal and Gibraltar Dams of 18-22%, respectively, from seeding convection bands. A more robust, causal analysis is needed to determine the impacts of cloud seeding on precipitation and streamflow.

Each seeding site is equipped with Automated High Output Ground Seeding (AHOGS) systems. These systems use flares with high concentrations of silver iodide, dispersed upon receiving a signal through cellular data connection to introduce seeding agents into storms systems (Yorty and

Cammans, 2022). Seeding decisions are made by a Weather Modification Association (WMA) certified project meteorologist based on radar images from the Vandenberg Air Force Base radar site. The project meteorologist has the option of firing flares individually in real time, or to order batch firing of any number of flares at selectable intervals at each site, e.g., three flares at 15-minute intervals, beginning at any selected time. Then, they are responsible for archiving relevant data from each event and report the number of flares used at the different seeding sites, the time of the seeding, and the type of flares which helps determine the composition and the weight of the seeding solution. If an aircraft is used, the report also includes the times of take-off and landing, as well as the amount of time spent actively seeding from the aircraft generators.

While the decision to seed or not to seed a certain cloud system is ultimately that of the meteorologist on call, meteorologist fixed effects [2] are not of concern since the decision to seed is based on a strict set of conditions that is known by the personnel working on the project and shared in official annual contracts. These conditions are based on a conceptual model of the dynamics of convection bands involving a frontal structure. The model explains that the primary low to mid-level inflow to these convection bands is expected to be along the leading edge of the bands, and that these inflow regions correspond to the accumulation zones of supercooled liquid cloud droplets. Practically speaking, this means that the desired region, within the cloud, for the introduction of the seeding material is known, and that flares burned from ground generators are timed to occur at the said leading edge of the bands. Other conditions considered in the decision include the possibility of targeting low-level winds from the surface up to the -5°C level, as well as the avoidance of seeding over areas that meet the established suspension criteria. Cloud seeding suspension criteria are invoked when the National Weather Service issues a severe storm, precipitation, flood warning, or flash flood warning that affects some part of the project area. There are general and area-specific suspension criteria to account for differences in hydrological conditions and previous burn areas from wildfires.

[2] Meteorologist fixed effects refer to unobserved, time-invariant characteristics associated with each meteorologist in the panel dataset, such as individual abilities, that could lead them to make a different decision to seed any given cloud.

2.3 Empirical Methods

I am interested in evaluating the effects of the implementation of a precipitation enhancement program using panel data in a setting where some locations, represented by rain gauges, are exposed to cloud seeding events starting in 1981 but other gauges located outside of the targeted areas are not treated. I exploit exogenous variation in the treatment to draw conclusions on the difference in potential outcomes. While precipitation enhancement programs are implemented in areas that suffer from decreased precipitation in an effort to extract as much water as possible from the clouds, the seasonal operations can be suspended for reasons unrelated to the weather, such as wildfires, contract timing, and budget constraints.

Compared to other studies that define seeding as a binary treatment only (Griffith et al., 2005), I consider both seeding as a binary variable and a continuous variable which allows me to include the treatment intensity in the analysis. To do this, I directly model the intensity of seeding measured as the total weight of seeding material multiplied by the number of flares used from each site or aircraft during a given seeding event. That is, intensity of treatment is defined in terms of grams of seeding material per seeding event. In the following subsections, I summarize the data used in the analysis and present the empirical methodology.

2.3.1 Data

I obtain data about cloud seeding from the yearly operations and evaluation reports prepared by North American Weather Consultants, Inc. for Santa Barbara County Water Agency. I supplement the data using the final weather modification activity reports submitted to the U.S. Secretary of Commerce, via the National Oceanic and Atmospheric Administration (NOAA) Weather Program Office. I compile daily precipitation data from the public records of a network of weather stations via the National Centers for Environmental Information of NOAA. I also use precipitation data publicly available from the Water Resources Division of the Santa Barbara County Public Works.

NAWC has conducted winter cloud seeding programs in Santa Barbara County beginning in 1981. Prior to this, Santa Barbara County had conducted two research programs. The earliest program, Santa Barbara I, ran for the period 1957-1960 and was sponsored by various organizations including the State of California, The University of California, Santa Barbara and Ventura counties, the National Science Foundation, the U.S. Weather Bureau, and the U.S. Forest Service (NAWC, 2022). The second research program, known as Santa Barbara II, was conducted during the winter seasons of 1967 to 1973. To avoid confounding the effects of the ongoing program with the research phase effect, I restrict historical data on precipitation to the period 1974-2021.

For each season, November through April, a number (1-6) of seeding sites are used to target two watersheds, Upper Santa Ynez and Huasna-Alamo, e.g. Twitchell, both located within the overall project area. These sites, depicted in Figure 2A.1, are chosen by NAWC to fit optimal seeding conditions criteria, particularly elevation, proximity to the target area, and location with regard to the wind direction. I follow Griffith et al. (2005) to choose the first set of target and control sites such that the control sites are not only located upwind but also bracket the target area, which provides better correlations between control and target sites. Similar to their approach, I use double-mass plots[3] to eliminate some outlier sites from consideration, due to long-term changes which did not correspond with the timing of any cloud seeding programs. I expand the sample to include some weather stations that were not available in Griffith et al. (2015), and test different combinations of data to keep sites that are more correlated in terms of historical precipitation data.

I compile data from the NAWC annual reports and construct a panel dataset with variables indicating for each day of the wintertime whether a seeding event had occurred, whether it was ground based or an aircraft was used, the number of flares used, the amount of seeding agent used, and the amount of time spent seeding conditional on the use of an aircraft, as well as the record of precipitation registered on each site.

2.3.1.1 Descriptive Statistics

In Table 2B.3, I summarize the treatment and intensity of treatment variables for each month and for each mode of conducting cloud seeding in our sample. The monthly frequency of seeding

[3]Double-mass analysis is used to test the consistency of a rainfall record by comparing the cumulative annual values of the station of interest with those of a reference station (Ponce, 1989). The reference station is usually the mean of several neighboring stations. The cumulative pairs are plotted in an arithmetic coordinate system, and the plot is examined for trend changes. The rainfall record at the station of interest is consistent if the plot is linear, and inconsistent otherwise.

events in the sample varies, with February being the month where most events occur at a mean of 3.2 ground-based events per month and 2 aircraft-based events per month. November and April are on the boundaries of the defined wintertime season. With fewer winter storms and fewer convective bands passing through the Santa Ynez and Twitchell watersheds, there are fewer cloud seeding opportunities, hence the low average of 0.3 and 0.5 ground-based operations per month in November and April respectively and 0.1 aircraft-based events per month for the same months. It is also notable that there are more ground than aerial seeding events for all months. Although both methods have been used to conduct the cloud seeding operations in Santa Barbara County, NAWC and the Santa Barbara County Water Agency (SBCWA) agreed that ground-based seeding is probably more fiscally efficient. The quantity of seeding agent is directly proportional to the number of flares used in each event (Yorty and Cammans, 2022). The number of flares used in an event is decided by the meteorologist on site, based on the physical characteristics of the storm. However, the reports indicate that the number of flares can vary randomly for the same type of storm because the contracted company was trying to test a possible correlation between the quantity used and the outcome of the operation.

The average number of events and quantity of seeding material can be driven down by the variance in the length of the operational period in each season. I include November to April to account for all possible events that occurred since 1981. However, Table 2B.4, which summarizes historical information about the operational season for each water-year, shows that the length of the season varies from 3.5 to 6 months, and can be even shorter for the airborne program, when operational. NAWC adjusts the seasonal period and targeted watershed within the overall program area for the needs expressed by SBCWA (Yorty and Cammans, 2022) and can be adjusted to fit various hydrological and budgetary circumstances. Hydrological constraints are typically binding after a fire that produces a large burn area, which creates concerns about the potential for excessive erosion and mud slides.

2.3.2 Estimation Strategy

The hypothesis is that cloud seeding operations increase precipitation in target areas. To investigate this hypothesis, I adopt a difference-in-differences methodology and estimate the average treatment effects on the treated (ATTs). The target and control areas, denoted $i \in N$, are uniquely identified panel units. Let precip_{it} denote the total daily precipitation at time t in area i. Let $CloudSeeding_{it}$ be the cloud seeding event indicator, such that $CloudSeeding_{it} = 1$ if unit i is treated in day t. The usual Two-Way Fixed-Effects estimator is obtained from estimating the following equation, such that α_i and γ_t are site and time fixed effects, respectively:

$$precip_{it} = \beta_{it} \cdot CloudSeeding_{it} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{2.1}$$

This initial equation is set up to estimate the average change in precipitation associated with cloud seeding events, controlling for both site- and time- fixed effects. The second equation investigates the heterogeneity in the effect of cloud seeding by differentiating the treatment variable based on the specific seeding method used:

$$precip_{it} = \beta^{G}_{it} \cdot CloudSeeding^{ground}_{it} + \beta^{A}_{it} \cdot CloudSeeding^{aerial}_{it} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{2.2}$$

Equation 2.2 is set to compare the effectiveness of the two seeding methods by estimating their differential impacts on precipitation levels. $CloudSeeding^{ground}_{it}$ and $CloudSeeding^{aerial}_{it}$ are binary variables indicating whether ground-based seeding or aerial seeding was conducted at location i at time t. This third equation, Equation 2.3, is used to understand the relationship between the intensity of the cloud seeding intervention, measured by the amount of seeding agent in kilograms, and the resulting change in precipitation.

$$precip_{it} = \beta^{q}_{it} \cdot quantity_{it} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{2.3}$$

According to the hypothesis that cloud seeding increases precipitation, the ATT, β_{it}, is expected to be positive.

The potential outcomes framework implicitly assumes the stable unit treatment value assumption that unit i's outcome is independent of the treatment status of unit $j \neq i$, which rules out spillover effect. To ensure this, I choose control units that are located upwind, based on Griffith et al. (2015), and on maps intersected with data on wind direction. Moreover, the no-anticipation assumption arguably holds because cloud seeding has no causal effect on precipitation prior to the implementation of the precipitation enhancement program, or even prior to a single cloud seeding event.

The identification assumption is that the difference in daily precipitation between areas that were affected by cloud seeding and those that were not would be constant over time in the absence of a precipitation enhancement program. This amounts to an assumption of parallel trends. The assumption of parallel trends is a key assumption in the difference-in-differences (DiD) framework, and it is crucial for obtaining unbiased estimates of treatment effects. The parallel trends assumption posits that, in the absence of treatment, the average treatment and control groups would have followed similar trends over time. In other words, the groups should have exhibited parallel or similar trajectories in the outcome variable even if they were not exposed to the treatment. That is, in the absence of cloud seeding, the average difference in precipitation between the target areas and control areas would remain constant over time. This assumption is crucial because it is the basis on which changes in the average treatment group outcomes relative to the control group can be attributed to the treatment itself. If the parallel trends assumption holds, then the DiD estimator provides an unbiased estimate of the causal effect of the treatment. However, if the assumption is violated, it introduces the risk of biased estimates, as divergent trends could be mistaken for treatment effects.

While testing this assumption is not possible given the unobservability of the counterfactual, plotting historical data for the treated and untreated units gives an indicative visualization of the pattern in which the data changes. In general, DiD methods are applied in cases where there is a substantial number of treated units, and the assumption of "parallel trends" can be made with a high degree of certainty. I visually test the parallel-trend assumption. In Figure 2A.2, I plot the outcome variable, precipitation, for the treated versus untreated group. This test suggests that the "parallel trends" assumption might not hold, even when control units are carefully chosen to be meteorologically and hydrologically similar to the target units. I find irregularities in the data and note that the two watersheds targeted by the same program and their respective upwind neighboring areas might be too different to confidently assume that the parallel trends assumption holds without adjusting the sample. However, it is possible to argue that weather data from different stations cannot show pre-trends comparable to the ones expected from economic variables. Unlike economic variables, which are often influenced by human decisions and may exhibit gradual changes over time, precipitation data inherently exhibits natural fluctuations and non-linear trends due to factors like long-term climate cycles, seasonal variations, and random weather events. This inherent variability makes it difficult to definitively establish clear pre-treatment trends for precipitation in both treated and control groups. Given these challenges, instead of focusing solely on pre-treatment trends, it might be sufficient to adopt a more relaxed approach for precipitation data. This approach emphasizes the existence and direction of pre-existing correlations between the treated and control groups. In simpler terms, I aim to demonstrate that historical data exhibits correlation as this indicates that both groups experience similar patterns of fluctuation, even if the exact magnitude of precipitation might differ. Additionally, both the treated and control groups should exhibit trends that move in the same direction (e.g., both increasing or decreasing) during the entire study period, including the pre-treatment and post-treatment phases. This suggests that external factors affecting both groups are likely similar, and any observed changes post-treatment can be more confidently attributed to cloud seeding. By demonstrating these two points, as shown in Figure 2A.2, I establish a weaker form of the parallel trends assumption, which is sufficient for drawing valid causal inferences within the DiD framework when applied to precipitation data.

It is important to acknowledge that this approach does not completely eliminate concerns about potential violations of the parallel trends assumption. Although it sets ground to proceed with the DiD analysis while acknowledging the inherent limitations associated with natural phenomena like precipitation. I also consider conducting additional robustness checks, such as exploring different control groups to strengthen the validity of the findings. To do this, I adjust the sample to obtain better matching between the location of the generator and the location of target areas. The initial setup, as explained in the data section, leads to a set of target units that all receive treatment at the same time. I leverage data on detailed descriptions of individual treatment events to identify which generators are activated conditional on weather data, specifically wind direction and speed. This information is descriptive and cannot be formally codified. This is because relevant weather data is obtained from the Vandenberg Air Force Base radar, with a range that includes all project areas, without the possibility of splitting the observations by location. This means that the radar data cannot be used directly in the analysis, since there is no heterogeneity in wind data between different areas. However, it is possible to identify a general trend that links specific ground generators to specific areas of the project, from the description of daily operations provided by the project meteorologist. Following this adjustment, I split the sample into two treated groups: Santa Ynez and Huasna-Alamo are subsets of the Santa Barbara project. The resulting sub-samples are shown in Table 2B.6.

While the synthetic difference-in-differences method (SDiD) developed by Arkhangelsky et al. (2021) is used in the literature to weaken the reliance on the parallel trend assumption, it relies on balanced panels with uniform pre-treatment periods, which is not the case of cloud seeding events data. Therefore, it is also unsuitable for the setting of this study. As a robustness check, and to alleviate concerns about potential violations of the assumption, I explicitly model time-varying unobserved heterogeneity by adding sub-basin-specific linear time trends to the traditional TWFE models (Wooldridge, 2021).

Another concern for the identification of the treatment effect is selection bias. Selection bias refers to a situation where the sample used in the analysis is not randomly selected from the population, leading to a distortion of the estimated relationships between variables. This bias can arise when there are systematic differences between the included and excluded observations, and these differences are related to the variables of interest. In this setup, a possible source of selection bias is the non-random assignment of the treatment (cloud seeding) to the treatment group. The fact that cloud seeding is only done on cloudy days introduces a form of "self-selection" bias because the decision to apply cloud seeding is related to the weather conditions, specifically cloudiness. When estimating the average partial effect of cloud seeding using a difference-in-differences approach with a two-way fixed effects estimator, the estimate could be biased upwards. Cloud seeding is applied selectively on cloudy days, and this decision is likely related to the precipitation levels. Cloudiness and precipitation are correlated, and this self-selection creates a non-random assignment of the treatment. The treatment group may have inherently different characteristics from the control group on cloudy days versus non cloudy days, and this could affect the estimated treatment effect.

To address this issue, one might consider implementing strategies to control for the self-selection bias, such as (1) Propensity Score Matching to estimate the propensity score, which is the probability of receiving cloud seeding given observed characteristics, and use matching techniques to create a balanced comparison group, (2) Instrumental Variables such that the binary variable of cloud seeding is replaced by another variable (instrument) that is related to cloud seeding but is not directly related to precipitation in order to address endogeneity issues arising from self-selection, or (3) include relevant control variables that may be correlated with both the assignment of cloud seeding and the precipitation to reduce bias. Unfortunately, data on relevant variables to be used in any of these strategies is not available from the weather stations. Another approach that could potentially help mitigate some of the selection bias present in the data is "thresholding"[4]. It consists of using a restricted sample where only observations that meet a certain threshold are kept. I restrict the sample to keep observations on days with precipitation (i.e., precipitation value > 0). By focusing only on days with precipitation, the analysis is restricted to situations where cloud seeding is more likely to occur. This can help reduce the self-selection bias, as the comparison is now made between cloudy days with cloud seeding to other cloudy days without cloud seeding. By focusing on days with precipitation, the sample is more homogeneous in terms of weather conditions. This improves the comparability between the treatment and control groups, potentially reducing confounding factors related to varying weather patterns.

Following the discussion about the identification of the true effect of cloud seeding operations on the quantity of precipitation, I show the results of the estimation of Equation 2.1 using three samples: (1) the full unadjusted and unrestricted sample, (2) the full adjusted but unrestricted sample, and (3)

[4]This approach assumes that the treatment effect of cloud seeding is constant across different levels of precipitation. If the treatment effect varies with precipitation intensity, the estimated average partial effect may still be biased. This dimension is beyond the scope of the analysis

the full adjusted and restricted sample. The different datasets are explained in Table 2B.5. While this exercise is interesting in showing the importance of understanding the structure of the data and the implications on the estimation results, inference is more robust and plausible using the estimates obtained using the adjusted and restricted sample since in corrects for issues with identification. Specifically, this strategy addresses the self-selection bias and potential violations of the parallel trends assumption, thus providing more valid estimated treatment effects and consequently allowing for more reliable conclusions about the impact of cloud seeding on precipitation.

2.4 Results

2.4.1 Changes in Precipitation

Table 2B.7 in the appendix summarizes the average partial effects of the selected covariates for the TWFE model with the full sample prior to adjustment. First, without additional differentiation between the aerial program and the ground program, I find a statistically significant correlation between precipitation and the cloud seeding-based precipitation enhancement program that has been operational for the period 1981-2022 in Santa Barbara County (Column 1). On average, cloud seeding increases precipitation by 0.26 inches per operation, which is a 45% increase on average. Using the adjusted sample, which identifies a treatment as it relates to a specific target, the estimates, presented in Table 2B.8, are higher in magnitude, and suggest that cloud seeding increases precipitation by 0.45 inches per operation on average.

Column 2 of Table 2B.7 shows the change in precipitation, in inches per operation, attributable to cloud seeding, using the full sample without adjustment, while column 2 of Table 2B.8 shows the results of the same estimation using the adjusted sample. Ground-based seeding seems to be more effective than aircraft-based seeding in both estimations. Using the full sample without adjustment and the full adjusted sample respectively, the estimates in column 2 of Table 2B.7 (Table 2B.8) suggest that ground-based seeding has a larger and more significant effect of 0.31 (0.37) additional inches of precipitation compared to aircraft-based seeding, evaluated at an average increase of 0.18 (0.27) inches of precipitation in the target areas compared to the control sites. The joint significance using the Wald test suggests that there is strong evidence to reject the null hypothesis that the coefficients on *seed_aerial* and *seed_ground* are statistically equal, in both estimations. This means that ground-based seeding is more effective in increasing precipitation. This might be due to the fact that aerial seeding is often used as a complementary measure. Most of the moisture in the cloud would have already been captured using ground flares. The coefficient on aerial seeding is still statistically different than zero, when using the adjusted sample, meaning that there is still a significant benefit from employing an aircraft in some operations to increase precipitation.

The third column of Table 2B.7 and column 3 of Table 2B.8 show the estimated effects of the new technology of flares on precipitation, using the full sample without adjustment and after adjustment, respectively. The *LW83* flares used from 1981 to 2000 weighed 400 grams and between 1 and 12 flares were fired during each seeding event. After the technological change in 2001, the weight of the new *ICE* flares changed to 150 grams, which corresponds to approximately 16 grams of the seeding agent, silver iodide. I find no statistical differences attributable to the change in technology from *LW83* to *ICE* flares (Column 3). The technological innovation made the pyrotechnic flares used at the AHOGS sites more effective in terms of nuclei production. The higher effectiveness is attributed to a change in the operative channel through which the flare works, evolving from slow contact nucleation to a fast condensation-freezing mechanism. This also means that less seeding material is needed to attain the same results. The result on the technology coefficient is explained by the fact that the higher effectiveness was used to decrease the weight of the flares, which nullifies potential increases in the effect on precipitation that could have happened if the same amount of seeding agent continued to be used.

The level of effort is modeled using the quantity of seeding agent used each day as the explanatory variable in Equation 2.1. The estimation is done using the full adjusted sample that corrects for the parallel trends assumption. As shown in Table 2B.9, one additional kilogram of seeding agent, silver iodide in this case, increases precipitation by 0.31 inches (57%) on average. This effect is an aggregate of the total material used in ground-based operations and aircraft-based operations. The disaggregated effects, as presented in column 2 of Table 2B.9, show an increase if 0.38 inches versus an increase of 0.25 inches, for ground and aircraft operations, respectively. I test the null hypothesis $H_0 : \beta_{kg\_AgI\_aerial} - \beta_{kg\_AgI\_ground} = 0$ and find enough statistical evidence to reject the null at the 1% significance level. This means that the effect of an additional unit of seeding agent during a ground-based cloud seeding operation is significantly higher than that of an additional unit during an aircraft-based operation. Again, this is explained by the operational logistics of cloud seeding in the Santa Barbara project, where ground flares are used as the base treatment whereas the aerial program is used as a complementary measure.

Furthermore, the spatially refined analysis relating generators to specific targets for each operation (see Table 2B.6 for sub-samples) reveals a difference in effectiveness of cloud seeding by sub-basin. As shown in Table 2B.10, cloud seeding seems to be slightly more effective in Huasna Alamo, with an increase of 0.34 inches per event, equivalent to 80% increase in precipitation at the average, compared to Upper Santa Ynez where the increase is estimated at 0.45 inches, or 76% increase in precipitation per event at the average. The Wald test of the null hypothesis $H_0 : \beta_{\text{treat\_alamo}} - \beta_{\text{treat\_ynez}} = 0$ gives a p-value of 0.067, which suggests marginal significance, but it does not reach conventional levels of significance. Thus, there is not enough evidence to reject the null hypothesis, although there may be a suggestion of a trend toward significance. Practically, this result does not show a significant difference in the effectiveness of cloud seeding between the sub-basin of Upper Santa Ynez and the sub-basin of Huasna Alamo.

Results of the estimation using the restricted sample[5] are presented in Table 2B.11. As anticipated, the coefficients are smaller in magnitude because the sample includes more homogeneous treatment and control groups conditional on weather. This correction improves the validity of the estimated treatment effects and offers more reliable conclusions about the impact of cloud seeding on precipitation, which is estimated at an increase of 0.30 inches per operation. Column 2 of the same table shows that an additional kilogram of silver iodide is predicted to increase precipitation by 0.24 inches on average, with a difference between the marginal effect of the quantity by method of seeding such that ground based flares are estimated to increase precipitation by 0.28 inches while an incremental increase in silver iodide dispersed by aircraft generators increases precipitation by 0.21

[5] The restricted sample only includes observations on days with precipitation > 0.

inches on average. Based on the potential limitations of the full sample DiD, I strongly favor the restricted sample DiD as the preferred specification. This approach actively addresses the challenges of selection bias and non-parallel trends, leading to a more reliable and unbiased estimate of the causal effect of cloud seeding on precipitation.

Finally, results of the separate regressions of cloud seeding on precipitation by sub-basin, using the restricted sample, are presented in Table 2B.12. Estimates are closer in magnitude such that an additional cloud seeding operation on a cloudy day with optimal seeding conditions is estimated to increase precipitation by 0.32 inches in the Huasna-Alamo sub-basin and by 0.27 inches in the Upper Santa Ynez catchment area. These additional inches of precipitation correspond to a 49% and a 26% increase for each sub-basin at the average, respectively. The estimated percentage increase for Upper Santa Ynez is not only smaller in magnitude when using the restricted sample, but its statistical significance also diminishes. This could be the result of a combination of factors; the estimation with no correction for selection bias produces estimates that are biased upwards, so the estimation conditional on more homogeneous target and control groups, terms of weather conditions, ought to produce smaller estimates. Moreover, the sample splitting strategy leads to a reduced sample size by deleting observations on days without precipitation. A smaller sample size can lead to less precise estimates and reduced statistical power, potentially limiting the ability to detect true treatment effects. Table 2B.4 shows that historically, the Santa Barbara county project has targeted the Huasna-Alamo catchment area in the Twitchell target area more often than it has targeted Upper Santa Ynez, which explains why this caveat is more relevant in the case of the latter sub-sample. The partial effect of cloud seeding on precipitation in Upper Santa Ynez is still statistically significant at the 5% level.

2.4.2 Costs and Benefits

In considering the optimal allocation of resources for cloud seeding operations in different areas of the Santa Barbara precipitation enhancement program, economic theory suggests considering the marginal cost and marginal benefit of the intervention in each target area. The efficiency equimarginal principle states that net benefits are maximized when the marginal benefits from an allocation equal the marginal costs. Thus, resources should be allocated where the marginal benefit equals the marginal cost. In this context, the marginal benefit is reflected in the percentage increase in precipitation attributed to cloud seeding, while the marginal cost encompasses the expenses associated with cloud seeding operations. The marginal benefit should be converted to a dollar amount to be able to draw such conclusions.

The first step in this approach is to quantify the effect of cloud seeding on precipitation for each area, which is what the results from the analysis so far have provided. The next step is to estimate the marginal cost to produce an additional unit of usable water. The choice to evaluate the cost of an additional unit of surface water rather than the cost of an additional unit of precipitation is grounded in the practical consideration of the ultimate goal of cloud seeding operations and the relevance to water resource management. Water reservoirs, such as Twitchell, Cachuma, Jameson, and the Gibraltar dam serve as storage facilities for water resources. Cloud seeding is intentionally performed over the catchment areas of these reservoirs to facilitate surface water runoff into them. Water managers are more concerned with the volume of water that can be stored in these reservoirs, as this directly impacts water supply, agricultural needs, and other water-dependent activities. The relationship between precipitation and surface water is not one-to-one. Precipitation is just the initial phase, and factors like runoff, infiltration, and evaporation influence how much of the precipitation contributes to surface water in reservoirs. Water resource managers need information on the cost-effectiveness of cloud seeding in terms of increasing the volume of water in reservoirs. The cost of an additional unit of surface water directly informs decisions related to the allocation of resources for cloud seeding operations. Measuring the cost of an additional unit of surface water is more practical and directly aligned with the operational goals of water resource management. It provides a tangible and actionable metric for decision-makers. In regions where water scarcity is a significant concern, the focus is on ensuring an adequate and cost-effective water supply. Assessing the cost of increasing surface water directly addresses this concern and helps prioritize interventions like cloud seeding.

To calculate the marginal cost of cloud seeding in each target area, I use the estimated effects of cloud seeding on precipitation obtained from the difference-in-differences (DiD) estimation and incorporate the rainfall-to-runoff conversion coefficient. The rainfall-to-runoff conversion coefficient represents the proportion of precipitation that eventually becomes runoff and contributes to surface water. The ratio used in Griffith et al. (2005) is $\rho = 0.67$, whereas a more conservative ratio reported in the annual operations reports is $\rho = 0.4$. For each target area (Huasna-Alamo and Santa Ynez), I calculate the additional runoff resulting from cloud seeding using the estimated effects obtained from the DiD estimation: Additional Runoff Surface Water$_i = \beta_i \cdot \rho$. Then, I convert the additional runoff to acre-feet. The conversion is based on the total target area, in acres, for each region, such that: Additional Surface Water$_i$(ac-ft) = (Additional Runoff)$_i$ $\cdot$ (Total Target Area)$_i$. To find the marginal cost, I divide the proportion of the total cost of cloud seeding for each area by the additional acre-feet of water generated in that area. This provides the cost per acre-foot of additional water: Marginal Cost$_i$(per ac-ft) = Total Cost of Cloud Seeding$_i$/Additional Surface Water$_i$(ac-ft). Since the cost to run the project in each target area individually is not provided, I use an approximation based on the following information: (1) The annual program cost is $300,000 U.S., which is shared on a 50/50 basis between the county and the local water purveyors; (2) The two targeted watersheds in the county cover approximately 447,230 acres, and the Upper Santa Ynez target area is slightly larger than the Huasna Alamo area. The working assumption is that Santa Ynez represents approximately 60% of the total area of the project. I vary the share of the annual costs dedicated to operating the cloud seeding program in each target area to include different possible allocations of resources.

To estimate the marginal benefit of cloud seeding, I refer to the water rates in the county of Santa Barbara (Water Wise, 2023a). The water use in Santa Barbara County report (Water Wise, 2022) shows that, for the year 2022, only 5% of the total metered water deliveries served agriculture[6]. The other 95% of water is directed to municipal and industrial use. The additional surface water in the catchment areas of Jameson Lake, Lake Cachuma, and the Gibraltar reservoir mainly adds to the supplies used by the water purveyors of the city of Santa Barbara and the Montecito Water District (Water Wise, 2023b). Water rates in both locations follow a block rate structure, which means that the water is priced at a different rate, in dollars per hundred cubic feet ($/HCF), for each

[6]Agencies in the county of Santa Barbara that report water use in agriculture are Carpinteria Valley Water District, Goleta Water District, Montecito Water District, City of Santa Barbara, and Santa Ynez Water District.

volume or "block" of water used, with rates increasing with each higher volume. If a consumer uses multiple blocks of water, then the water bill includes multiple rates. An excerpt from the report for the locations of interest is presented in Table 2B.13. Surface water from Lake Cachuma is also a significant source of water in the public water systems of the Carpinteria Valley water district and the Goleta water district, where water use represents 54% and 24% of total metered water deliveries, respectively, at the rates of 2.02 and 2.33 $/HCF.

The estimated costs and benefits for different cost allocation values between areas, alternative values for the rainfall-to-runoff ratio, and different block rates for water are presented in Table 2B.14 for Huasna Alamp and Table 2B.15 for Upper Santa Ynez. Many combinations of marginal cost and marginal benefit are possible, each producing a different estimated value of net benefits. In general, net benefits are higher in Santa Ynez compared to Huasna Alamo due to both higher marginal benefits and lower marginal costs. Most of the water from the Upper Santa Ynez reservoirs is used by municipalities at relatively higher rates than those for agricultural water use. Although the estimated percentage increase in precipitation is higher in Huasna Alamo, the volume of surface water produced via cloud seeding is smaller conditional on the area considered as the targeted catchment area. Changing the working assumption about areas to one that posits that both catchment areas are similar (in acres) produces marginal costs and benefits depicted in Table 2B.16.

In light of the results obtained from the analysis, the allocation of effort in cloud seeding operations can be strategically guided by the principle of efficiency. The assessment of net benefits reveals a nuanced interplay between allocating efforts to Santa Ynez and Huasna Alamo. Despite higher marginal costs in Santa Ynez, the overall net benefits are elevated in this region. This outcome is attributed to the confluence of various working assumptions made at different levels of the analysis. Despite statistical evidence that indicates a higher percentage increase in precipitation due to cloud seeding in Huasna Alamo, the economic implications paint a different picture. In Santa Ynez, where the marginal benefits of cloud seeding are discerned to be higher, the destination of water use is primarily municipal and industrial, attracting a higher price per acre-foot. On the contrary, the Huasna Alamo, with its elevated precipitation impact, predominantly channels water into agriculture, where the price per acre-foot is comparatively lower. It is important to note that, due to limitations in available data, it is not feasible to differentiate further based on the specific costs associated with different types of flares or the choice between aerial and ground seeding methods. This limitation prevents a more detailed cost-benefit analysis at the level of specific implementation techniques.

2.5 Discussion and Conclusion

This paper estimates the impact of cloud seeding on water resources and the cost of providing additional water supply through a weather modification program. In this study, I undertake a rigorous statistical analysis to assess the impact of cloud seeding on precipitation, employing sophisticated methods borrowed from the realm of econometrics, and using daily data on cloud seeding operations and precipitation. I use site fixed effects and time fixed effects to control for potential omitted variables that could influence the observed outcomes. This approach aims to enhance the internal validity of the study by capturing site-specific and temporal nuances, providing a robust foundation for inference. Addressing concerns related to selection bias and potential violations of the parallel trends assumption, I implement methodological strategies to fortify the credibility of the estimated treatment effects. My consideration of these statistical intricacies sought to ensure that the findings could be confidently attributed to the cloud seeding intervention rather than confounding factors.

The results suggest that Santa Barbara County's cloud seeding program has been successful in increasing precipitation in the target areas by 26-49%. The success rate of cloud seeding is heterogeneous across locations and across technologies. Ground-based seeding seems to have higher effects on precipitation than aircraft-based operations. This analysis serves as the first step to answer economic questions about the viability of cloud seeding for a long-running project in Southern California, as well as about the optimal allocation of effort levels between different target areas within the same project. This was done by calculating both marginal costs and benefits under varying assumptions. The additional water supply generated by cloud seeding comes at a cost of $6.56 per acre-foot, which is relatively low compared to other water supply management techniques such as groundwater banking and desalination plants.

The nuanced exploration of these economic metrics involves a thoughtful consideration of factors such as water destination (municipal and industrial versus agricultural use) and the associated pricing structures. This comprehensive economic evaluation aims to provide a holistic understanding of the efficiency and cost-effectiveness of cloud seeding initiatives in the water-scarce region of Southern California. In an effort to interpret the findings within the scope of a specific project, I conduct a comparative analysis between two distinct areas: Huasna Alamo and Upper Santa Ynez. Despite the higher percentage increase in precipitation due to cloud seeding in Huasna Alamo, the nuanced economic evaluation uncovers a more intricate picture. In Upper Santa Ynez, where the marginal benefits of cloud seeding are greater, the destination of water use emerges as a critical factor. The predominance of municipal and industrial water use in this region translates into a higher market price per acre-foot, amplifying the economic benefits associated with increased precipitation.

Beyond the internal discussion of allocation based on efficiency, the findings from this study can be contextualized within the literature in three ways. First, there is a comparison to be made between the estimated cost of additional water supply to the value of water in different sectors in California. Agricultural water in California is valued from $40 to $50 per ac-ft and up to $175 per ac-ft during drought (Hunter et al., 2007). The value of water for hydroelectric use by PG&E is $100 per ac-ft, while Municipal and industrial values range from $300 to $600 per ac-ft, according to a 2007 California Energy Commission (CEC) report, which was made publicly available in 2018. Second, I compare these costs with other alternatives to increase the availability of water supply. The variable cost of groundwater banking projects is between $150-$250 per ac-ft (Hunter et al., 2007). The cost of desalinated water in California is between $700/ac-ft (Hunter et al., 2007) and $980/ac-ft (Griffith et al., 2005). The cost of cloud seeding water is also less expensive than pumping ground water, evaluated at approximately $200-$300/ac-ft (Griffith et al., 2005).

Finally, I compare the findings of this essay to the costs estimated by other studies. The Wyoming pilot project estimates a weather modification cost between $3.96 and $7.91 per ac-ft, while the Utah Division of Water Resources has stated that the estimated direct cost of water from an 8% to 12% increase in snowpack from cloud seeding in key mountain watersheds is about $1 per ac-ft

(Hasenyager et al., 2012). The estimated costs from the cloud seeding projects in Nevada and Colorado are $6 to $12 per ac-ft and "less than" $20/ac-ft respectively, with new projects that only bear variable costs by building on existing infrastructure estimated to supply additional water at one-third of the pre-stated costs.

This research offers crucial insights with far-reaching implications for policymakers tasked with formulating effective climate change mitigation strategies. As the specter of water scarcity looms over various regions, especially those characterized by arid and semi-arid climates, understanding the potential benefits and cost-effectiveness of cloud seeding programs becomes paramount. The findings from this study underscore the relevance of cloud seeding as a viable tool in the arsenal of water resource management strategies. Policymakers can leverage these insights to tailor interventions that not only address current water scarcity concerns but also proactively mitigate the anticipated impacts of climate change on water resources. Furthermore, the research prompts a reevaluation of resource allocation, emphasizing the need for strategic investments in cloud seeding programs in regions where the cost-benefit analysis proves favorable. The integration of such initiatives into broader climate adaptation plans could enhance resilience in the face of changing precipitation patterns and evolving water resource dynamics.

BIBLIOGRAPHY

Acharya, A., Piechota, T. C., Stephen, H., and Tootle, G. (2011). Modeled streamflow response under cloud seeding in the North Platte River Watershed. *Journal of Hydrology*, 409(1-2):305–314.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Averyt, K., Meldrum, J., Caldwell, P., Sun, G., McNulty, S., Huber-Lee, A., and Madden, N. (2013). Sectoral Contributions to Surface Water Stress in the coterminous United States. *Environmental Research Letters*, 8(3):035046.

Beare, S., Chambers, R., and Peak, S. (2010). Statistical modeling of rainfall enhancement. *The Journal of Weather Modification*, 42(1):13–32.

Griffith, D., Yorty, D., and Beall, S. (2015). Target/control analyses for Santa Barbara County's operational winter cloud seeding program. *The Journal of Weather Modification*, 47(1):10–10.

Griffith, D. A., Solak, M. E., Almy, R. B., and Gibbs, D. (2005). The Santa Barbara cloud seeding project in coastal southern California, summary of results and their implications. *The Journal of Weather Modification*, 37(1):21–27.

Griffith, D. A., Solak, M. E., Yorty, D. P., and Brinkman, B. (2007). A level ii weather modification feasibility study for winter snowpack augmentation in the Salt River and Wyoming Ranges in Wyoming. *The Journal of Weather Modification*, 39(1):76–83.

Hasenyager, C., McGettigan, S., and Cole, D. (2012). Utah Cloud Seeding Program, Increased Runoff/Cost Analyses. *Utah Department of Natural Resources, Division of Water Resources*.

Hunter, S. M. et al. (2007). Optimizing cloud seeding for water and energy in California. *A pier-final Project Report for California Energy Commission, March*.

Lachhab, R. (2024). Silver bullets: Cloud seeding and water resources in California.

Levin, Z., Halfon, N., and Alpert, P. (2010). Reassessment of rain enhancement experiments and operations in Israel including synoptic considerations. *Atmospheric Research*, 97(4):513–525.

Ma, J.-H., Yoo, C., Song, S.-U., Na, W., Cho, E., Song, S.-K., and Chang, K.-H. (2023). Different effect of cloud seeding on Three Dam Basins, Korea. *Water*, 15(14):2555.

Malik, S., Bano, H., Rather, R. A., and Ahmad, S. (2018). Cloud seeding; its prospects and concerns in the modern world-a review. *Int. J. Pure App. Biosci*, 6(5):791–796.

Morrison, A. E., Siems, S. T., Manton, M. J., and Nazarov, A. (2009). On the analysis of a cloud seeding dataset over Tasmania. *Journal of Applied Meteorology and Climatology*, 48(6):1267–1280.

Ponce, V. M. (1989). *Engineering Hydrology, Principles and Practices*. Prentice Hall, Englewood Cliffs, NJ.

Rauber, R. M., Geerts, B., Xue, L., French, J., Friedrich, K., Rasmussen, R. M., Tessendorf, S. A., Blestrud, D. R., Kunkel, M. L., and Parkinson, S. (2019). Wintertime orographic cloud seeding—a review. *Journal of Applied Meteorology and Climatology*, 58(10):2117–2140.

Schwabe, K., Nemati, M., Landry, C., and Zimmerman, G. (2020). Water markets in the Western United States: Trends and opportunities. *Water*, 12(1):233.

Solak, M., Girdzus, J., and Griffith, D. (1996). Precipitation augmentation potential from convection band cloud seeding in Santa Barbara County. *Prepared for Santa Barbara County Flood Control and Water Conservation Dist. and Water SBCWA NAWC Report WM*, pages 3–96.

Solak, M., Henderson, T., Allan, R., and Duckering, D. (1987). Winter orographic cloud seeding over the Kern River Basin in California. *The Journal of Weather Modification*, 19(1):36–40.

Thompson, J. and Griffith, D. (1988). Precipitation augmentation potential from convective band cloud seeding in Santa Barbara County. *NAWC report WM*, pages 7–87.

Water Wise, S. B. C. (2022). Water use in Santa Barbara County report. Technical report.

Water Wise, S. B. C. (2023a). Water rates in Santa Barbara County report. Technical report.

Water Wise, S. B. C. (2023b). Water supply portfolio of pulic water systems in Santa Barbara County. Technical report.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. Available at SSRN 3906345.

Yoo, C., Na, W., Cho, E., Chang, K.-H., Yum, S. S., and Jung, W. (2022). Evaluation of cloud seeding on the securement of additional water resources in the Boryeong Dam Basin, Korea. *Journal of Hydrology*, 613:128480.

Yorty, D. and Cammans, G. (2022). Annual cloud seeding report. Prepared for Santa Barbara County Department Of Public Works, Water Agency Division.

APPENDIX 2A

FIGURES

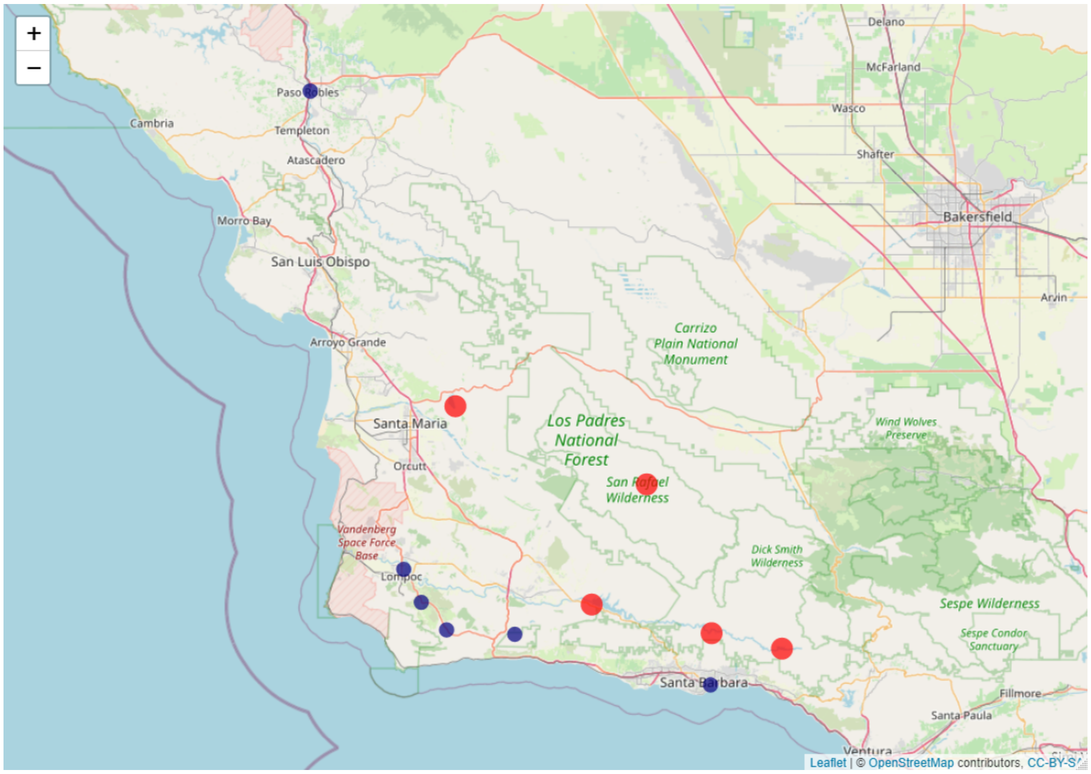

Figure 2A.1 Control (blue) and Target (red) sites in Twitchell and Upper Santa Ynez watersheds, California, US

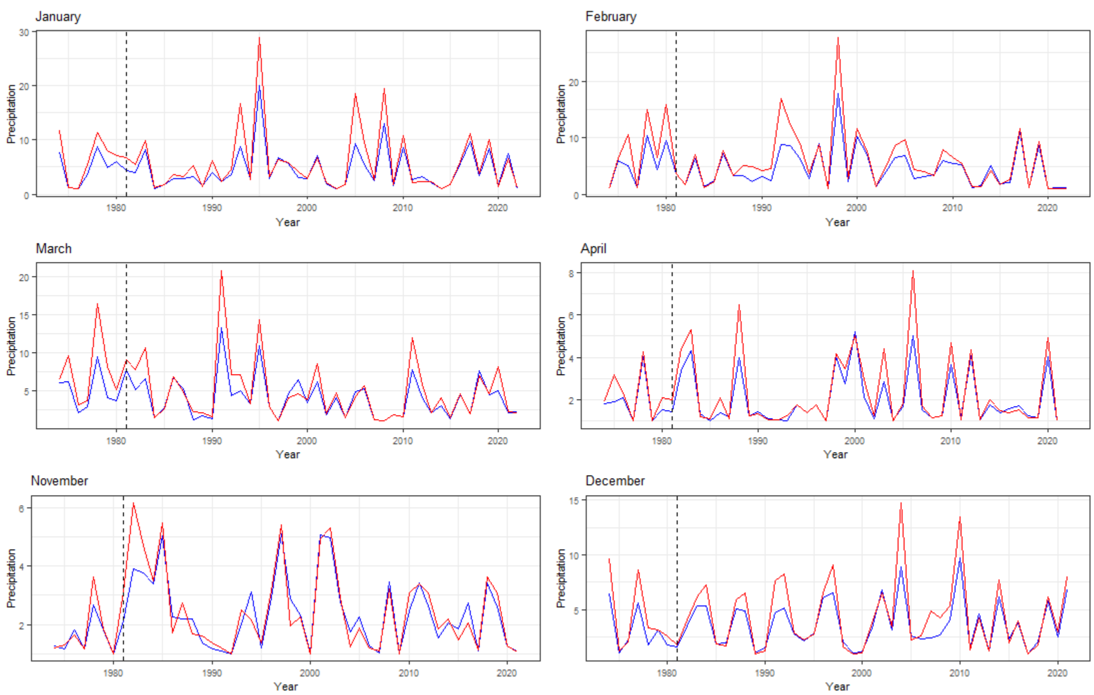

Figure 2A.2 Visual test of the Parallel Trends Assumption. Average monthly precipitation totals in target (blue) and control (red) locations during the winter season (1974-2022)

APPENDIX 2B

TABLES

Table 2B.1 1974-1981 Monthly Average Precipitation (inches) In Target and Control Sites

Site	Jan	Feb	Mar	Apr	Nov	Dec
Control Sites						
Lompoc	3.27	3.68	4.25	0.89	0.41	1.98
PasoRobles	3.35	3.41	2.64	0.80	0.66	1.90
SanJulian	4.65	4.63	5.53	1.08	0.37	2.90
Sisquoc	3.07	4.18	3.92	0.92	0.49	1.67
Salispuedes	4.05	4.57	5.00	0.95	0.71	2.40
SantaBarbara	3.55	4.33	4.31	0.49	0.72	2.17
Target Sites						
Cachuma	4.50	5.16	5.83	1.10	0.65	2.73
Gibraltar	6.59	7.04	7.39	1.31	0.69	3.73
Jameson	6.92	8.12	8.06	1.22	0.78	4.13
Juncal	6.92	8.20	7.92	1.15	0.78	4.13
Twitchell	2.95	4.11	4.38	1.38	0.54	2.25

Table 2B.2 1981-2021 Monthly Average Precipitation (inches) In Target and Control Sites

Site	Jan	Feb	Mar	Apr	Nov	Dec
Control Sites						
Lompoc	3.07	2.97	2.67	0.89	1.32	2.31
Paso Robles	3.11	2.89	2.65	0.76	1.21	2.43
San Julian	5.20	5.06	3.98	1.41	1.88	3.40
Sisquoc	2.88	2.94	2.86	1.07	1.39	2.53
Salispuedes	3.83	4.04	2.97	1.19	1.48	2.95
Santa Barbara	3.54	3.65	2.70	0.87	1.35	2.47
Target Sites						
Cachuma	4.46	4.32	3.59	1.09	1.49	3.20
Gibraltar	5.20	5.05	3.59	1.19	1.23	3.59
Jameson	6.36	6.36	4.93	1.79	1.99	3.96
Juncal	5.62	5.85	4.56	1.74	1.82	3.96
Twitchell	2.59	2.22	2.44	1.01	1.18	2.55

Table 2B.3 Descriptive statistics of the seeding events and flares

Month	Unit	Mean	Max	Min	S.D.
Ground Seeding Events					
January	Count	3.1	17	-	3.6
February	Count	3.2	14	-	3.1
March	Count	2.8	12	-	2.8
April	Count	0.5	4	-	1.0
November	Count	0.3	3	-	0.7
December	Count	2.4	8	-	2.5
Ground-based Flares					
January	Grams	1856	7500	-	2105
February	Grams	2404	11100	-	2733
March	Grams	2319	8800	-	2342
April	Grams	501	3150	-	906
November	Grams	215	3150	-	633
December	Grams	1901	9000	-	2240
Aerial Seeding Events					
January	Count	1.9	12	-	2.7
February	Count	2.0	9	-	2.7
March	Count	1.4	14	-	2.7
April	Count	0.1	1	-	0.2
November	Count	0.1	3	-	0.5
December	Count	1.3	9	-	2.3
Aircraft-generated Flares					
January	Grams	906	4158	-	1287
February	Grams	1486	7380	-	2131
March	Grams	823	7452	-	1541
April	Grams	15	330	-	67
November	Grams	24	930	-	150
December	Grams	571	5580	-	1197

Note: Only seasons with active precipitation enhancement program are included.

Table 2B.4 Santa Barbara County Historical Program Information

Season	Months	Ground Program	Airborne Program
2010-2011	4.5	Nov.15 - Mar.31	Nov. 15 - Mar 31
2011-2012	4.5	Dec. 1 - Apr.22[1]	N/A
2012-2013	3.5	Dec. 1 - Mar.15[2]	N/A
2013-2014	5	Nov.15 - Apr.15	Dec.15 - Mar.15
2014-2015	5	Nov.15 - Apr.15	Jan. 1 - Mar.31
2015-2016	6	Nov. 1 - Apr.30	Dec. 1 - Mar.31
2016-2017	6	Nov. 1 - Apr.30	Jan.01 - Mar.31
2017-2018	5	Nov.15 - Apr.15[1]	N/A
2018-2019	5	Nov.15 - Apr.15[1]	N/A
2019-2020	4.5	Dec. 1 - Apr.15[1]	N/A
2020-2021	4	Dec. 1 - Mar.31[1]	N/A
2021-2022	4	Dec. 1 - Mar.31	N/A

[1] Program only conducted for the Twitchell Target area.
[2] Season shortened due to the likelihood of no significant runoff occurring. No aircraft were included due to large burn areas present in Santa Barbara County.
Source: NAWC 2022 Annual Cloud Seeding Report

Table 2B.5 Explanation of samples used in the estimation

Dataset	Obs	Treatment	Treated group	Control group
Full unadjusted and unrestricted	All	= 1 for all units in the treated group if there is record of a cloud seeding activity from any generator; = 0 otherwise	All units inside target areas after 1981	All units outside target areas + all units inside target areas before 1981 + all units on days with no precipitation
Adjustment: Matching a subset of units to specific generators using wind direction and detailed description of cloud seeding operations that identify which generators were used to target which areas				
Full adjusted but unrestricted	All	= 1 for some units in the treated group that are intentionally targeted by cloud seeding activity on that day; = 0 otherwise	Some units inside target areas after 1981 that intentionally received treatment that day	Some units outside target areas, located south-west of units in the corresponding treated sub-group +some units inside target areas before 1981 +all units on days with no precipitation
Restriction: Keep only observations on days with precipitation (i.e., precipitation value > 0)				
Full adjusted and restricted	All such that precip> 0	= 1 for some units in the treated group that are intentionally targeted by cloud seeding activity on that day, conditional on nonnegative value of precipitation; = 0 otherwise	Some units inside target areas after 1981 that intentionally received treatment that day, conditional on nonnegative value of precipitation that day	Some units outside target areas, located south-west of units in the corresponding treated sub-group +some units inside target areas before 1981

Table 2B.6 Construction of the adjusted full sample

	Adjusted Sample	
Treatment group subsample	Huasna-Alamo	Upper Santa Ynez
Weather Stations	Twitchell Paso Robles Salispuedes San Julian Sisquoc Lompoc	Cachuma Juncal Jameson Gibraltar Santa Barbara
Ground Generators	Berros Peak Mount Lospe Harris Grade Sudden Peak Suey Bixby Reagan Twitchell	Dos Vistas Gaviota West Camino Gibraltar El Deseo Corona Del Mar Lacumbre Graham
Common Generators	Mt. Tranquillion, Solomon, Figueroa	

Table 2B.7 Estimates of the average partial effect of cloud seeding on precipitation in inches/event

	Precipitation		
	(1)	(2)	(3)
Seeding	0.260***		0.287***
	(0.030)		(0.028)
Seeding_Aerial		0.184	
		(0.092)	
Seeding_Ground		0.308***	
		(0.039)	
Tech			-0.027
			(0.014)
Seeding × Tech			0.063
			(0.046)
Fixed-effects			
Month FE	Yes	Yes	Yes
Site FE	Yes	Yes	Yes
Site × Month FE	Yes	Yes	Yes
Fit statistics			
Observations	94,710	94,710	94,710
Adjusted R^2	0.069	0.066	0.067

Clustered (site & month) standard-errors in parentheses
Signif.: ***: 0.01, **: 0.05, *: 0.1

Table 2B.8 Average partial effect of cloud seeding on precipitation in inches/event, using the adjusted sample

	Precipitation		
	(1)	(2)	(3)
Seeding	0.425***		0.451***
	(0.030)		(0.083)
Seeding_Aerial		0.267**	
		(0.081)	
Seeding_Ground		0.370***	
		(0.057)	
Tech			-0.018
			(0.013)
Seeding × Tech			-0.057
			(0.127)
Fixed-effects			
Month FE	Yes	Yes	Yes
Site FE	Yes	Yes	Yes
Fit statistics			
Observations	94,699	94,699	94,699
Adjusted R^2	0.032	0.034	0.033

Clustered standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 2B.9 Average partial effect of 1kg of Silver Iodide (AgI) on precipitation in inches/event, using the adjusted sample

	Precipitation	
	(1)	(2)
kg_AgI	0.315***	
	(0.041)	
kg_AgI_Aerial		0.251***
		(0.049)
kg_AgI_Ground		0.389***
		(0.082)
Fixed-effects		
Month FE	Yes	Yes
Site FE	Yes	Yes
Fit statistics		
Observations	94,699	94,699
Adjusted R^2	0.028	0.029

Clustered standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 2B.10 Average partial effect of cloud seeding on precipitation in inches/event, by sub-basin

	Precipitation			
	(1)	(2)	(3)	(4)
Seeding	0.425***			
	(0.030)			
Seed_Huasna_Alamo		0.344***		0.340***
		(0.005)		(0.004)
Seed_Santa_Ynez			0.450***	0.455***
			(0.041)	(0.042)
Fixed-effects				
Month FE	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
Fit statistics				
Observations	94,699	51,654	43,045	94,699
Adjusted R^2	0.032	0.023	0.034	0.032

Clustered standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 2B.11 Average partial effect of cloud seeding on precipitation in inches/event, using the restricted sample

	Precipitation			
	(1)	(2)	(3)	(4)
Seeding	0.303***			
	(0.045)			
kg_AgI		0.240***		
		(0.035)		
Seeding_Aerial			0.243**	
			(0.094)	
Seeding_Ground			0.232*	
			(0.092)	
kg_AgI_Aerial				0.201***
				(0.010)
kg_AgI_Ground				0.278**
				(0.082)
Fixed-effects				
Month FE	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
Fit statistics				
Observations	17,377	17,377	17,377	17,377
Adjusted R^2	0.068	0.069	0.070	0.069

Clustered standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 2B.12 Average partial effect of cloud seeding on precipitation inches/event by sub-basin, using the restricted sample

	Precipitation (1)	(2)
Seed_Huasna_Alamo	0.328***	
	(0.006)	
Seed_Snta_Ynez		0.276**
		(0.062)
Fixed-effects		
Month FE	Yes	Yes
Site FE	Yes	Yes
Fit statistics		
Observations	10,143	7,234
Adjusted R^2	0.050	0.046

Clustered standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 2B.13 2022-2023 Water Rates in $/Hundred Cubic Feet in two water agencies in the county of Santa Barbara, California, US

Agency	Residential		Commercial		Agriculture	
	Block	Rate	Block	Rate	Block	Rate
Montecito WD	0 – 9	6.94	Uniform	10.18	0 – 9	6.94
	10 – 35	11.79			> 9	5.82
	> 36	13.02				
City of Santa Barbara	1 – 4	4.85	< mo budget[1]	7.4	< mo budget	3.63
	5 – 16	14.46	> mo budget	27.1	> mo budget	26.93
	> 17	27.19				

[1] Monthly budget.

Table 2B.14 Estimated marginal cost (MC) and marginal benefit (MB) of cloud seeding in Huasna Alamo in $/ac-ft

Runoff Ratio	Cost share	$MC(\$/ac-ft)$	$MB(\$/ac-ft)$
0.40	0.3	4.77	3.50
	0.5	7.95	
	0.7	11.13	
0.67	0.3	2.85	5.85
	0.5	4.75	
	0.7	6.65	

Table 2B.15 Estimated marginal cost (MC) and marginal benefit (MB) of cloud seeding in Santa Ynez in $/ac-ft

Runoff Ratio	Cost share	MC ($/ac-ft)	Block rate	MB ($/ac-ft)
0.40	0.3	3.82	1	6.31
	0.5	5.64	2	13.69
	0.7	7.89	3	20.83
0.67	0.3	2.02	1	10.56
	0.5	3.36	2	22.93
	0.7	4.71	3	34.89

Table 2B.16 Estimated marginal costs (MC) and marginal benefit (MB) in Huasna Alamo and Santa Ynez in $/ac-ft, assuming 50-50 split of total project area

Runoff Ratio	Cost share	MC ($/ac-ft)		MB ($/ac-ft)	
		Huasna Alamo	Upper Santa Ynez	Huasna Alamo	Upper Santa Ynez
0.40	0.3	3.82	4.10	4.37	5.25
	0.5	6.36	6.76		
	0.7	8.91	9.47		
0.67	0.3	2.28	2.42	7.31	8.79
	0.5	3.80	4.03		
	0.7	5.31	5.65		

CHAPTER 3

JOINT ESTIMATION OF IRRIGATION ADOPTION AND CROP CHOICE IN ETHIOPIA

3.1 Introduction

Agriculture in Ethiopia is crucial for food security and economic growth. The country's crop production is diverse, with five major cereals: teff, wheat, maize, sorghum, and barley, forming the core of the agricultural and food economy (Winer, 1989). However, the sector faces challenges such as low yields, poor resilience to adverse weather, and limited access to technology and inputs (Bekabil, 2014). Improving production levels and reducing variability are essential for enhancing food security and increasing household incomes. Ethiopia's agricultural sector is heavily reliant on smallholder farmers, who account for a significant portion of agricultural value added (Taffesse et al., 2012). Adequate policy support to farmers facing such challenges requires the implementation of agricultural policies that help them decrease variability and increase production. In particular, water supply risk is an important determinant of production, especially in arid and semi-arid areas, significantly affecting both cropping and investment decisions (Marques et al., 2005; Koundouri et al., 2006; Feinerman and Tsur, 2014).

Policymakers attempt to mitigate the water-related production risk by implementing various policies targeting the optimization of water use, increasing water supply, preventing floods, increasing drainage and irrigation capacities, or implementing water pricing policies to (dis)incentivize farmers' use of certain sources of water. Investment in irrigation infrastructure has been proposed as an adaptive response to sustain agricultural yields under higher temperatures and irregular rainfall (Duflo and Pande, 2007). Irrigation infrastructure is used to regulate water flows and smooth rainfall irregularities. Ideally, this offers more security to farmers who are facing high risks of weather variability and attenuates the adverse effects of climate change on crop choices. Arellano-Gonzalez and Moore (2020) provide empirical evidence that reducing drought risk through access to large scale water storage enables higher value but riskier production in California agriculture. In the absence of insurance, farmers have been shown to adopt risk-averse strategies, favoring lower but less variable returns to higher value but risky options. Karlan et al. (2014) find that rainfall risk is a binding constraint on the investment decisions of farmers in northern Ghana. (Wang et al., 2008) show that farmers in China use both irrigation choice and crop choice to adapt to climate variability.

However, access to irrigation infrastructure does not ensure full adoption of irrigation technology by small-scale farmers. Irrigation adoption in developing countries is influenced by various factors such as socio-economic characteristics, psychological factors, availability of resources, and institutional support. Factors like strong norms, risk perceptions, and attitudes play a role in the adoption of irrigation technology and practices (Hornum et al., 2023). Additionally, the adoption of small-scale irrigation technologies varies across countries, with different factors influencing adoption patterns. For example, the intensity of agricultural labor and the use of inorganic fertilizers are positively associated with the adoption of small-scale irrigation in Ethiopia (Hatch et al., 2022).

The adoption of new agricultural technologies is a critical driver of improving farm productivity, enhancing resource efficiency, and ensuring long-term sustainability in the agricultural sector (Tilman et al., 2002). The factors that influence a farmer's decision to adopt a new agricultural technology, as identified by Feder et al. (1985), include farm size, labor availability, credit constraints, human capital, risk preferences, access to input and output markets, and land tenure systems (Headey and Jayne, 2014; Jayne et al., 2014; Holden and Lunduka, 2014; Melesse and Bulte, 2015; Wakeyo and Gardebroek, 2013). Farmers are more likely to adopt technologies with a clear economic benefit, considering both potential gains in yield or efficiency and associated investments in acquiring and implementing the technology. Land tenure security, infrastructure development, and access to markets create the institutional environment within which adoption decisions are made. In this sense, secure land rights and reliable market access can incentivize technology adoption by mitigating risks and enhancing potential returns. Moreover, the complexity of the technology, its compatibility with existing practices, as well as access to technical support significantly influence adoption rates (Sumberg, 2005). User-friendly technologies that integrate well with existing farming systems and offer readily available technical support are more likely to be adopted by farmers due to reduced learning curves and enhanced problem-solving capacities.

Among these determinants, farmer knowledge plays a particularly crucial role in facilitating technology adoption. To obtain this knowledge, farmers require access to reliable information about new technologies, encompassing their potential benefits and drawbacks, as well as the necessary skills for successful implementation. Hence, one way to encourage the adoption of irrigation is through extension. Agricultural extension programs have a significant impact on farmers' production choices in general, and on technology adoption in particular. These programs play a crucial role in improving knowledge penetration in the agricultural sector and empowering farmers (Shangshon et al., 2023). Extension programs enhance agricultural production efficiency and water resource management through the adoption of technologies and management strategies (Rudnick et al., 2020). Access to information provided by extension services is important for small-scale farmers in particular (Amghani et al., 2023) and has been found to influence the adoption of irrigation technologies and practices by farmers. The adoption of improved production technologies and participation in extension delivery programs have been shown to improve farmers' productivity, efficiency, and welfare (Mohammed and Abdulai, 2022).

Although agricultural producers face considerable production risk due to weather and unreliable water supplies, they also have adaptation strategies available, such as the ability to choose alternative crops. Many field crops require different quantities of water for production and can be planted across different growing seasons (Fisher and Huber-Lee, 2012). Knippenberg et al. (2020) explore this concept further by showing that crop diversification, as a result of land fragmentation in Ethiopia, helps mitigate weather risks and increase the diversity of foods available. They claim that higher land fragmentation makes crop diversification feasible for households living in a heterogeneous agro-ecological environment with diverse parcel characteristics in terms of slope, elevation, and wetness. Such diverse parcel characteristics are prevalent in Ethiopia given its topography-induced spatial variability. Crop choice has been repeatedly observed as an adaptation strategy in agriculture. Farmers adjust their crop choice and inputs in response to observed or forecast changes in seasonal rainfall (Rosenzweig and Udry, 2013; Taraz, 2017). Seo and Mendelsohn (2007) find that South American farmers adapt to climate by changing crops, switching away from maize, wheat, and potatoes towards squash, fruits, and vegetables.

While the effect of agricultural extension on technology adoption, and the impact of irrigation water availability on crop choice have been extensively studied (Mesfin et al., 2022; Pan et al., 2018; Genius et al., 2014; Krishnan and Patnam, 2014; Koundouri et al., 2006; Lichtenberg, 1989; Moore et al., 1994), a research gap exists in understanding the specific effects of extension on simultaneous decisions like crop choice and irrigation adoption. Existing studies often focus on one decision at a time, neglecting the potential jointness in these choices. This paper aims to address this gap by investigating whether access to agricultural extension services influences the adoption of irrigation technologies and practices by Ethiopian farmers and how agricultural extension and irrigation availability affect the choices farmers make regarding the crops they cultivate. I propose a theoretical framework to analyze the determinants of the joint decision of irrigation adoption and crop choice for smallholder farmers facing a subsistence constraint in Ethiopia. Understanding the relationship between extension services, irrigation technology, and crop choice can inform policy and interventions aimed at improving food security and farmer livelihoods in Ethiopia. While previous studies have explored the independent effects of extension services on irrigation adoption and the influence of irrigation on crop choice, this paper examines the combined impact of these factors. Additionally, the paper contributes to the literature by focusing on Ethiopia, a context with limited resources and where extension services might play an even more critical role compared to previously studied settings. By examining the combined effects of extension services and irrigation on crop choices in the specific context of Ethiopia, this paper offers valuable insights that can contribute to the broader body of knowledge on agricultural adaptation strategies in water-scarce environments.

I use Ethiopia's Living Standards and Measurement Study-Integrated Survey on Agriculture (LSMS-ISA). This data source provides information on cropping choices at the field level, characteristics of the field, weather, household demographics, input costs and output prices. I estimate a nested logit model, which aligns with the theoretical framework of the farmer's optimization problem by including the effect of various determinants, such as access to agricultural extension, on the nested decision choosing a crop type and adopting irrigation.

The remainder of the paper is organized as follows. I briefly present a background on Ethiopia, focusing on agriculture, irrigation, and extension. Then, I present a theoretical model of crop choice and irrigation adoption. In section 3, I explain the empirical methods and data. I present the estimation results in Section 4. Section 5 concludes the paper with a summary of the results and a discussion of policy relevance.

3.2 Background on Agriculture in Ethiopia

With a population of approximately 123 million people (2022), growing at a rate of 2.6% annually, Ethiopia faces a unique food security challenge, despite the significant progress it has made in agricultural production since the 1980's. An estimated 85% of its population relies on land-intensive subsistence agriculture, contributing 45% to the national Gross Domestic Product (GDP). This dependence is further amplified by the country's low urbanization rate, with only 15% of the population residing in urban areas. This scenario leads to a shrinking average land holding per household over time. Currently, the national average land size per household is 0.96 ha (Headey et al., 2014) with variations among regions. This land is predominantly used for staple crop production (Nuru and Seebens, 2008). Ethiopia's agricultural sector is dominated by smallholder farming, with these households contributing over 90% of both total agricultural output and cultivated land. Smallholder farmers rely on their production for income generation and subsistence (Bekabil, 2014). The country's diverse landscape, encompassing varied agro-ecological zones, topography, and natural vegetation, has led Ethiopian small farmers to develop complex farming methods and diverse cropping patterns. Ethiopia boasts a diverse range of fruits, vegetables, root crops, and tubers, adapted to various agro-ecological zones and altitudes. Small-scale farmers are the primary producers of horticultural crops, primarily relying on rainfed agriculture with limited irrigation.

Recurring droughts and growing population pressure drove the Ethiopian government to prioritize irrigated agriculture in the country's development agenda to reduce the food deficit (FAO, 2016). The area under irrigation is developed primarily with small irrigation schemes, requiring lower capital and technical investments, and reaching small communities. Despite the increases in irrigated land area achieved through multiple quinquennial plans, the harvested irrigated crops were in 2014/2015 not yet using this new potential fully. The expansion of irrigation has faced challenges related to study and design, construction or implementation, irrigation management, and lack of support services such as extension (Shitu et al., 2022). The challenges of irrigating crops in Ethiopia also include inadequate awareness of irrigation water management, inadequate knowledge on improved and diversified irrigation agronomic practices, and shortage of basic technical knowledge on irrigation methods (Shitu and Almaw, 2020).

Studies conducted in Ethiopia by the Food and Agriculture Organization (FAO) and Water Resources Development Authority identified the need to establish improved, sustainable farming systems in the Central Highlands of Ethiopia, in view of the high population pressure, high erosion hazard, loss of forest reserves and declining per capita food production. In the same report (Ojukwu et al., 2001), the African Development Fund (ADF) declares that the government emphasis is to develop the irrigation sub-sector by assisting and supporting farmers to improve irrigation management practices and the promotion of modern irrigation systems on small (less than 200ha), medium (200 to 3000 ha) and large-scale (over 3000 ha) schemes.

For several decades, Ethiopian agricultural development has been heavily reliant on government-driven research and technology generation (Gebreselassie et al., 2017). This commitment dates back to the institutionalization of agricultural research activities over five decades ago. Currently, a diverse network of research institutions at various levels engages in generating, disseminating, and applying agricultural technologies. Ethiopia's agricultural extension system, while extensive, faces limitations that hinder its effectiveness in supporting irrigated agriculture. A key challenge lies in the knowledge gap among Development Agents (DAs) regarding irrigation practices (Leta et al., 2017). Several factors contribute to this shortcoming. First, the agricultural extension system lacks a long-term strategic vision. Frequent policy changes and inconsistencies create an unstable environment for DAs and farmers alike. This is particularly evident in irrigation practices, where the system's emphasis has swung dramatically between promoting national-scale water harvesting initiatives and encouraging rainfed agriculture (Leta et al., 2017). Second, the sheer number of extension agents deployed does not necessarily translate to effective service. In fact, there is a mismatch between the large number of extension experts and their low capacities to provide efficient services (Leta et al., 2017). Gebremariam et al. (2021) examines the determinants of farmers' interaction with agricultural extension agencies in Ethiopia.

Ethiopia has been experiencing rapid population growth and, in many regions, there is no additional land to be brought into cultivation (Ojukwu et al., 2001). Thus, food for the growing population must be obtained by intensifying production. The ADF recognizes that farmers have an influence on cropping patterns and that their choices might not always match what the policymakers have envisioned while deciding the scope of an irrigation scheme (Ojukwu et al., 2001). Discrepancies between the optimal projections and the actual choices of farmers could have significant impacts on the success of irrigation projects, since these are built based on a scenario of certain crops with given water needs and evapotranspiration. When the water needs are too different than the ones projected, the project might not offer enough resources to satisfy water demand. This could result in crop damage, which in turn may encourage producers to not choose high-value, water-intensive crops. For this reason, the irrigation projects offices analyze which crops to grow each season based upon market prices and water usage, providing a list of possible crops to be grown to the farmers. The farmers then make the final decision of what they will grow by choosing from this list. This also makes extension a simultaneous policy to guide farmers adjusting to the introduction of a new irrigation scheme.

3.3 Theoretical Framework: Understanding Farmer Decision-Making in Irrigation Adoption and Crop Choice

This section develops the conceptual framework to consider how Ethiopian farmers make two crucial decisions: adopting irrigation and choosing crops. While these choices might appear independent at first glance, they are actually closely linked. One possible framework could treat these choices as separate entities, analyzing irrigation adoption and crop choice independently. However, this approach can lead to a problem of endogeneity. This arises because the choice to adopt irrigation can be influenced by the crop choice and vice versa. For instance, water-intensive crops like vegetables are more likely to benefit from irrigation, incentivizing adoption. Conversely, adopting irrigation technology might open up the possibility of cultivating crops with higher water requirements that were previously unsuitable due to limited water availability. Ignoring this interdependence can lead to biased estimates and inaccurate conclusions about the true drivers of each decision.

The rationale for this approach is supported by findings from various studies. Green et al. (1996) highlight the importance of considering crop types when analyzing irrigation technology adoption, emphasizing the varying water requirements and potential benefits of irrigation across different crops. Similarly, Shrestha and Gopalakrishnan (1993) and Green and Sunding (1997) estimate technology choices conditioned on the type of crop produced, further solidifying the need to consider the endogeneity of these decisions.Lichtenberg (1989) provides a compelling example of jointly modeling irrigation and crop choices. He employed a multinomial logit model to understand land allocation and irrigation decisions, modeled as six different crop-technology combinations, across major crops in Nebraska. Similarly, Moreno and Sunding (2005) looked at the joint estimation of irrigation technology adoption and land allocation.

3.3.1 A Theoretical Model of Crop Choice and Irrigation Adoption

Under the assumption that crop choice and technology adoption are driven by a farmer's objective of maximizing expected utility of profit, one can focus on the choice problem that relates the simultaneous decision-making process to farmer, field, crop, and technology characteristics. Farmers who rely on their on-farm production to meet their household food consumption needs might face a different utility function that takes into account their optimal consumption level, rather than simply maximizing expected profit. There are many ways to think about this, such as assuming that they can purchase the food they need with the income they generate through their farm. This does not practically change the farmer's problem from that of a profit-maximizing one. However, if using a certain level of production for direct household consumption is important, then conceptually one can express that by imposing a constraint to the farmer's maximization problem.

Let π_{it}^c represent the profit for farmer i in year t from producing a single output q_{it}^c, where the superscript c denotes the crop of choice, through a technology described by a production function $f^c(\cdot)$, such that $f^c(\cdot)$ is continuous, twice differentiable, and exhibits diminishing marginal productivity to variable inputs. The subscripts i and t are cumbersome to carry throughout the argument, hence they are dropped as the model is presented for a representative farmer i in year t. The farmer uses a crop-specific vector of inputs to production $\mathbf{x}^c$. Let p^c denote output price and $\mathbf{r}^c$ the vector of input prices, such that p^c and $\mathbf{r}^c$ are non-random under the assumption that farmers are price-takers in both the input and output markets. The farmer can choose to use irrigation as an additional input, incurring a cost r^w where the superscript w denotes irrigation. Irrigation adoption is taken to be a choice over two alternatives: to irrigate or not irrigate. The choice to irrigate is thus a variable that follows a Bernoulli distribution, and the cost incurred by the farmer is expressed using an indicator function such that:

$$\mathbb{1}[w]r^w = \begin{cases} r^w & \text{if } \mathbb{1}[w] = 1 \\ 0 & \text{if } \mathbb{1}[w] = 0 \end{cases} \tag{3.1}$$

The productivity of crop c not only depends on the adoption of irrigation, but also on a vector of other farm management practices and heterogeneous farm resources and farmer characteristics, denoted $\mathbf{s}$. This is captured by including a productivity function $y^{c,w}(\mathbf{s})$ in the production function, which is written as $q^{cw} = f^c[y^{c,w}(\mathbf{s}), \mathbf{x}^c]$.

Under the assumption that the farmer makes the irrigation adoption and crop choice decisions simultaneously, the farmer's problem[1] in year t is to maximize the expected utility function of profit:

$$\begin{aligned} &\max_{\pi^{cw}} U(\pi^{cw}) = \max_{\mathbf{x}^c} U(p^c f^c[y^{c,w}(\mathbf{s}), \mathbf{x}^c] - \mathbf{r}'\mathbf{x}^c - \mathbb{1}[w]r^w) \\ &\text{s.t. } f^c[y^{c,w}(\mathbf{s}), \mathbf{x}^c] \geq H^c \end{aligned} \tag{3.2}$$

where H^c is the consumption level of the product from crop c needed by the household. The farmer's decision whether or not to adopt irrigation is a binary choice, where the farmer can choose to adopt ($w = 1$) or not ($w = 0$) irrigation. Irrigation is assumed to increase crop productivity given other inputs and farm/farmer characteristics, i.e., $y^{c,1}(\mathbf{s}) > y^{c,0}(\mathbf{s})$, but also increases the total cost

[1]The farmer's problem can be expanded to include production risk. Following Koundouri et al. (2006), risk can be represented by a random variable ε, with distribution $G(\cdot)$, exogenous to the farmer's actions. The farmer thus maximizes the *von Neumann-Morgenstern* utility function of profit, i.e $\max_{\pi^{cw}} E[U(\pi^{cw})] = \max_{\mathbf{x}^c} \int U(p^c f^c[y^{c,w}(\mathbf{s}), \mathbf{x}^c] - \mathbf{r}'\mathbf{x}^c - \mathbb{1}[w]r^w) dG(\varepsilon)$

of production by r^w[2]. Furthermore, the farmer's decision on crops is a choice over a set of M crops. For $c = 1, ..., M$ crop types and $w = 0, 1$, under the simultaneous decision assumption, the farmer makes a choice over a set of M^2 mutually exclusive combinations of c and w. The farmer will choose the combination $j = (c, w)$ if the utility with this combination is greater than the utility with other combinations, i.e.,

$$U(\pi_{it}^{j}) > U(\pi_{it}^{m}) \text{ for all } m \neq j \tag{3.3}$$

3.4 Empirical Methods

3.4.1 Nested Logit

Building upon the theoretical model of the farmer's joint decision regarding crop choice and irrigation adoption, I apply a Nested Logit Model (NL) for the empirical analysis. Discrete choice models provide a valuable tool for analyzing and predicting discrete decision-making and are particularly relevant when individuals face a set of finite alternatives, such as the selection of crop-irrigation combinations. While the Multinomial Logit Model (MNL) is a commonly employed choice model, its underlying assumption of independence of irrelevant alternatives (IIA) presents a potential limitation in this context. The IIA assumption postulates that the relative probability of choosing one alternative compared to another is independent of the availability of other alternatives. However, in the context of the farmer's decision, this assumption might not hold true. The decision to adopt irrigation for a specific crop might be influenced by the irrigation requirements of other crops under consideration. More generally, crop choice and irrigation adoption are considered jointly and assumed to be influenced by a shared set of factors, such as crop prices, input costs, farm characteristics, and climatic conditions. The Nested Logit Model offers a solution to this limitation by introducing a hierarchical structure that acknowledges potential dependencies between choices. This structure recognizes that certain decision-making processes might involve a nested sequence, in no particular time order. The nested logit model is more appropriate when a decision maker faces a choice set than can be partitioned into subsets such that two properties hold: (1) For any

[2]Since irrigation adoption is a binary decision, the decision on the quantity of water used conditional on irrigation adoption is not considered. Thus, the cost corresponding to adoption is fixed such that $r^1 = b$ and $r^0 = 0$, where $b > 0$ is a constant.

two alternatives in the same subset, the ratio of probabilities is independent of the attributes or existence of other alternatives in the subset; (2) for any two alternatives in different subsets, the ratio of probabilities can depend on the attributes of other alternatives in the two subsets. In other words, the IIA assumption holds within each subset but does not hold for alternatives in different subsets.

The choice set for crops can be constructed based on various clustering assumptions. For instance, crops can be grouped into major crop categories such as cereals, pulses, fruits, vegetables, etc. This would result in a choice set of M crop groups such that $M > 2$. An alternative clustering, based on the value of crops, would result in a binary choice between "staple crops" and "cash crops". For flexibility, the possible crop choices are assumed to be multiple ($M > 2$). The farmer chooses the combination of crop and irrigation that maximizes their utility. This structure is modeled as a two-level nested choice where the farmer chooses irrigation adoption and crop jointly. Figure 3A.1 is a visual representation[3] of the nested choice model when irrigation is available for all crop types. The tree structure, although it shows a partitioning of the irrigation-crop choice, is not to be confused with a sequential choice model. Rather, it is used to represent flexible patterns of substitution among correlated crop choices based on the grouping assumption. While the choices at each level of the tree structure are conceptually interchangeable since the decision is simultaneous, it still is important to impose an *ex ante* structure on substitution patterns because of the IIA. In the structure depicted in Figure 3A.1, factors influencing crop selection will primarily influence irrigation decisions within the chosen crop category. For example, with a discouraging factor such as high irrigation costs, a farmer who chooses vegetables with high irrigation needs is more likely to substitute towards another vegetable variety with lower water requirements. Substitution across categories (e.g., vegetables to cereals) might be weaker due to factors like market demands or crop rotation practices. If the order is reversed such that crop choice is nested within the irrigation decision, the substitution pattern becomes less intuitive and potentially unrealistic in this context. This structure implies that a farmer would not consider irrigation needs until after deciding to irrigate, which might not reflect the reality of joint decision-making.

[3]Figure 3A.2 represents the nested tree structure when some nests are degenerate because only the rainfed option is available.

Let the set of alternative crop-irrigation combinations $j = (c, w)$ be partitioned into $K = M^2$ mutually exclusive nests (i.e., subsets) denoted $B_1, B_2, \ldots, B_K$. The utility that farmer i obtains from alternative j in nest B_k in year t is denoted $U_{ijt} = U(\pi_{it}^j)$. The farmer is assumed to choose the alternative from which she derives the highest utility. The utility of each alternative is determined by a vector of characteristics of the farmer and the alternatives. The researcher has information on some, but not all, of the determinants of choice. This is reflected by splitting the utility into a deterministic component V_{ijt} and a stochastic component ϵ_{ijt}:

$$U_{ijt} = V_{ijt} + \epsilon_{ijt} \tag{3.4}$$

The probability P_{ijt} that farmer i chooses some crop-irrigation alternative j in year t is equal to the probability of U_{ijt} being the largest of all $U_{i1t}, \ldots U_{iKt}$. This probability is written as:

$$\begin{aligned} P_{ijt} &= \mathrm{P}\left(U_{ijt} > U_{imt} \quad \forall m = 1, \ldots, K : m \neq j\right) \\ &= \mathrm{P}\left(\epsilon_{imt} - \epsilon_{ijt} \leq V_{ijt} - V_{imt} \quad \forall m = 1, \ldots, K : m \neq j\right) \end{aligned} \tag{3.5}$$

Given the deterministic component of the utility functions $V_{i1t}, \ldots, V_{iKt}$, this probability depends on the assumptions on the distribution of the stochastic error terms $\epsilon_{i1t}, \ldots, \epsilon_{iKt}$. For the nested logit model, the vector of unobserved utilities, $\epsilon_{it} = \langle \epsilon_{i1t}, \epsilon_{i2t}, \ldots, \epsilon_{iKt} \rangle$ has a Generalized Extreme Value (GEV) cumulative distribution, such that:

$$G^{GEV}(\epsilon_{ijt}) = \exp\left(-\sum_{k=1}^{K}\left(\sum_{j \in B_k} e^{-\epsilon_{ijt}/\lambda_k}\right)^{\lambda_k}\right) \tag{3.6}$$

For a logit model, each ϵ_{ijt} is independent with a univariate extreme value distribution, while the ϵ_{ijt} are correlated within the nests. For any two alternatives j and m in nest B_k, ϵ_{ijt} is correlated with ϵ_{imt}. For any two alternatives in different nests, the unobserved portion of utility is $\mathrm{cov}\left(\epsilon_{ijt}, \epsilon_{imt}\right) = 0$ for any $j \in B_k$ and $m \in B_\iota$ with $\iota \neq k$. The parameter λ_k denotes the "dissimilarity parameter", which is an inverse measure of the correlation, and captures the degree of unobserved utility independence among alternatives within nest B_k. A higher value of λ_k indicates greater independence and less correlation between the alternatives in the nest. This implies that the alternatives are less similar in terms of unobserved factors that influence utility. Conversely, a lower value of λ_k suggests stronger correlation and greater similarity between the alternatives within the nest, implying the unobserved factors influencing utility are more alike. The value $\lambda_k = 1$ represents complete independence within the nest, meaning the unobserved utility components for each alternative are entirely independent. Interestingly, if $\lambda_k = 1$ for all nests, the nested logit model collapses to the standard logit model. This is because the nested logit model reduces to the product of independent extreme value terms under this condition, essentially removing the nesting structure and its associated parameters.

Under the held assumption that the vector of unobserved utility has a GEV distribution, the probability that farmer i chooses crop-irrigation alternative j in year t from the complete choice set can be written as:

$$P_{ijt} = \frac{e^{V_{ijt}/\lambda_k}\left(\sum_{j \in B_k} e^{V_{imt}/\lambda_k}\right)^{\lambda_k - 1}}{\sum_{\iota=1}^{K} \left(\sum_{m \in B_\iota} e^{V_{imt}/\lambda_\iota}\right)^{\lambda_\iota}} \tag{3.7}$$

While acknowledging the joint nature of these decisions, the primary interest of this paper lies in understanding the impact of agricultural extension on irrigation adoption behavior and crop choice. To achieve this, I decompose the joint probability of adopting a specific crop-irrigation pair into conditional and marginal probabilities. This approach, as outlined by Maddala (1983), simplifies the estimation process. This step corresponds to writing a more intuitive expression for the choice probabilities by decomposing the observed component of utility into two parts: (1) a part that is constant for all alternatives within a nest, denoted W_{ikt}, and (2) a part that varies over alternatives within a nest, denoted L_{ijt}. Then, Equation 3.4 becomes:

$$U_{ijt} = W_{ikt} + L_{ijt} + \epsilon_{ijt} \text{ for } j \in B_k \tag{3.8}$$

Thus, the probability of farmer i choosing crop-irrigation alternative j in year t is equal to the product of the probability to choose *an* alternative in nest B_k and the conditional probability to choose *the* alternative j given some alternative in the same nest B_k is chosen. In other words, the nested logit probability is the product of two standard logit probabilities, such that the probability of choosing alternative $j \in B_k$, P_{ijt}, equals the marginal probability of choosing the nest B_k, P_{iB_kt}, multiplied by the conditional probability of choosing alternative j given that alternative in B_k is chosen, $P_{ijt|B_k}$:

$$P_{ijt} = P_{iB_kt} \times P_{ijt|B_k} \tag{3.9}$$

such that

$$P_{ijt} = \frac{e^{W_{ikt}+\lambda_k I_{ikt}}}{\sum_{\iota=1}^{K} e^{W_{i\iota t}+\lambda_\iota I_{i\iota t}}} \times \frac{e^{L_{ijt}/\lambda_k}}{\sum_{m\in B_k} e^{L_{imt}/\lambda_k}} \tag{3.10}$$

where

$$I_{ikt} = \ln \sum_{m\in B_k} e^{L_{imt}/\lambda_k} \tag{3.11}$$

I_{ikt} is the inclusive value (IV), which corresponds to the expected value of the utility that farmer i obtains from the alternatives in nest B_k. It is represented in Equation 3.11 as the log of the denominator of the second level (the rescaled measure of the attractiveness of the nest B_k). It is important that the inclusive value enters as an independent variable in the first level because it works as a link between the two levels of the nested logit model by bringing information from the bottom level into the upper level.

Estimation of the nested logit model is done by full information maximum likelihood to allow the most efficient use of available information (Forinash and Koppelman, 1993).

The estimation of the nested logit model produces one parameter estimate for each alternative-specific variable and one parameter estimate for each farmer-specific variable for each alternative, except for the baseline alternative for which the set of parameters is set to 0. It also estimates the coefficient λ_k on the inclusive value I_{ikt}. As mentioned in the previous section, this coefficient represents the dissimilarity parameter, reflecting the degree of independence among the unobserved components of the utility function for the alternatives in each nest. For the model to be consistent with a random utility model, it must be the case that $0 < \lambda_k < 1$ for all k. If the dissimilarity parameter is negative, this implies that the model is inconsistent with utility-maximizing behavior (Hensher et al. (2005), section 13.6). If the dissimilarity parameter is greater than one, then this indicates that the model is consistent with utility-maximizing behavior for only some range of the independent variables.

Computationally, fitting a nested logit model becomes more challenging with an increasing number of parameters. To account for this, it is important to choose a concise set of variables that makes the model close to the true data-generating process without being too complicated. Furthermore, convergence might not be achieved with a small number of observations for some alternatives. To remedy this, I choose a slightly different nested structure that reflects the low irrigation adoption rate for certain crop groups. This results in dropping irrigated pulses, irrigated roots, and irrigated spices as possible alternatives in the farmer's choice set. A nest that contains only one alternative is called a degenerate nest and the dissimilarity parameter of such nests is not defined in the random utility nested logit model. Numerically, the issue of nonidentifiable or undefined parameters can be solved by setting constraints on them. I impose three constraints on the model to set the dissimilarity parameters of the degenerate nests (pulses, roots, and spices) to be equal to 1.

3.4.2 Conditional Logit Model

As mentioned in the previous section, the nested logit model is appropriate under the assumption that the IIA holds within each nest but not between alternatives in different nests. Furthermore, the dissimilarity parameter λ_k captures the degree of independence between alternatives in each nest, such that $\lambda_k = 1$ means complete independence. The hypothesis that $\lambda_k = 1$ for all k can be tested after a Nested Logit Model estimation. If there is not enough statistical evidence at the conventional levels to reject the hypothesis, then a more appropriate discrete choice model would be some variation of the standard logit model, since the nested structure would be unnecessary. The multinomial logit and pure conditional logit models are the most widely used tools in this case (Heiss, 2002). These two models can be combined into a version of the conditional logit model (simply called the conditional logit, as opposed to the pure conditional logit) that, similarly to the nested logit model, uses the characteristics of an individual, e.g. farmer i, as well as the characteristics of an alternative, e.g. crop-irrigation choice j, to explain the probability that the decision-maker i chooses the alternative j in time t.

Starting with the same utility decomposition in Equation 3.4, the conditional logit model assumes that the vector of error terms, $\epsilon_{it} = \langle \epsilon_{i1t}, \epsilon_{i2t}, \ldots, \epsilon_{iKt} \rangle$ has a Gumbel distribution, i.e., a Type-I extreme value distribution, such that:

$$G^{Gumbel}(\epsilon_{ijt}) = \exp\left(-\exp^{-\epsilon_{ijt}}\right) \tag{3.12}$$

Under the assumption that the error terms are independent and identically distributed (iid) extreme value, the choice probabilities are derived as the product of the individual cumulative probability distributions from Equation 3.5, such that:

$$P_{ijt} = \int \left(\prod_{j \neq m} e^{-e^{-\left(\epsilon_{ijt} + V_{ijt} - V_{imt}\right)}} \right) e^{-\epsilon_{ijt}} e^{-e^{-\epsilon_{ijt}}} d\epsilon_{ijt} \tag{3.13}$$

McFadden (1974) shows that this probability has a straightforward analytical solution. Thus, the probability that farmer i chooses a crop-irrigation alternative j in year t is:

$$P_{ijt} = \frac{e^{V_{ijt}}}{\sum_j e^{V_{imt}}} \tag{3.14}$$

While the nested logit model aligns better with the assumption that the irrigation adoption decision is nested within the crop choice, the conditional logit model assumes a higher independence between irrigated and rainfed choices within the crop nests. After running the nested logit, I examine the behavior of the dissimilarity parameters. If for some specifications, the estimated dissimilarity parameters are very close to 1, this suggests a high degree of independence between the alternatives within the nests. In simpler terms, it implies the choices within each nest behave much like completely separate options. Thus, a dissimilarity parameter close to 1 indicates the nested structure might be unnecessary, and the conditional logit model, which assumes independence between all alternatives, becomes a good robustness check.

3.4.3 Determinants of Choice

This section details the types of variables and coefficients used to model the utility function in the context of farmers' decisions regarding crop selection and irrigation adoption, based on the specified theoretical framework. The deterministic component of the utility function captures the systematic factors influencing the farmer's choice. It is typically modeled as a linear function (Heiss, 2002) of (1) farmer and farm attributes, which represent characteristics specific to the farmer or their farm operation that might influence their decision-making process, such as farmer knowledge, farm assets, and physical quality of the land, and (2) alternative attributes, which represent characteristics of the available crop-irrigation combinations that influence their utility for the farmer, such as crop prices and input costs. Coefficients associated with these variables represent the marginal change in the farmer's expected utility for a unit change in the corresponding variable. For example, a positive coefficient on crop price indicates that a higher expected price increases the utility of choosing that crop-irrigation combination. Conversely, a negative coefficient on labor suggests that higher labor needs make the combination less attractive to the farmer. On the other hand, the random component of the utility function captures the influence of unobserved factors specific to each crop-irrigation combination, such as individual farmer preferences beyond those captured by the deterministic component. Farmer-specific variables are collected in a vector $\mathbf{z}_{it}$ for each farmer i in year t, while alternative-specific variables are collected in a vector $\mathbf{d}_{ijt}$ for each farmer i and for each alternative j in year t.

3.4.3.1 Human Capital Assets

Education: Studies have consistently employed farmer education level as a factor influencing the adoption of agricultural innovations (e.g., Nkamleu and Adesina (2000)). It functions as a proxy for several factors that can influence decision-making such as (1) access to information, as higher education levels are often associated with greater access to information sources, such as extension services, media, or other farmers. This information can play a crucial role in raising awareness about new technologies, improved practices, and market opportunities, potentially influencing both irrigation adoption and crop selection; (2) understanding technical aspects since education can equip farmers with the capacity to comprehend the technical complexities associated with different crops and irrigation technologies. This understanding can empower them to make informed decisions about adopting new practices and evaluating their potential benefits and drawbacks; and (3) profitability assessment in the sense that education can also enhance a farmer's ability to analyze the economic viability of different crop choices and irrigation options. By considering factors like yields, input costs, and market prices, farmers with higher education levels may be better equipped to make choices that maximize their net benefit. I hypothesize that, similar to studies like Zerfu and Larson (2010), the effect of education on technology adoption in Ethiopia is positive.

Extension: Agricultural extension programs serve as a crucial bridge between research and practice, disseminating valuable information directly relevant to agricultural production, particularly in contexts like Ethiopia where access to information is limited for farmers. These programs can equip farmers with the knowledge and skills necessary to understand new technologies. Extension agents can offer guidance on agricultural practices, water management techniques, and crop management strategies, empowering farmers to improve their yields and resource utilization. They can also offer access to information about market trends, prices, and demand which can can help farmers make informed decisions about crop selection and improve their bargaining power in the market. Studies conducted in various contexts (e.g., Adesina and Zinnah (1993)) have demonstrated a positive correlation between participation in extension programs and the adoption of land-improving technologies. This evidence suggests that extension services play a significant role in influencing farmer behavior and promoting agricultural development. I hypothesize that extension will have a positive effect on irrigation adoption for high value crops such as fruits, vegetables, and cash crops.

3.4.3.2 Farm Assets

Land tenure: Land ownership serves as a proxy for wealth and influences crop selection in various ways. The cost of adopting irrigation technology may be more easily justified and recouped with secure land property rights, potentially increasing the likelihood of adoption on farms with greater landholdings. Farmers with secure land tenure may be more willing to invest in long-term improvements like terracing (Gebremedhin and Swinton, 2003) or irrigation, as they have the assurance of reaping the benefits over time. Conversely, those with limited land ownership or relying heavily on rented land might prioritize short-term returns and choose less water-intensive crops due to uncertainties about future access to the land.

Land area: Arable land area is also assumed to be a determinant of crop choice and irrigation adoption decisions for smallholder farmers. With smaller plots, farmers could prioritize crops that offer high yields per unit area to maximize their output. This might favor cash crops or high-value vegetables and fruits over lower-yielding staples. Additionally, irrigation becomes a more attractive option for smaller landholdings. The limited space necessitates maximizing productivity, and irrigation offers greater control over water availability, potentially leading to higher yields. However, smaller plots also mean that the farmer has limited financial resources that could hinder technology adoption and put higher pressure on subsistence farming, leading to the choice of staple foods to meet the household consumption needs.

Livestock: Raising livestock is not only a separate agricultural activity for Ethiopian farmers but also an indirect driver of crop choice. Farmers might cultivate certain crops specifically to meet the feed needs of their livestock. This can include fodder crops, legumes rich in protein, or other crops that complement the existing feed resources available. Certain crops, like maize, might be chosen not only for human consumption but also for their byproducts like stover, which can be used as livestock feed. This allows farmers to extract additional value from their crops and potentially benefit from integrated value chains. On the other hand, farmers might dedicate a portion of their land for grazing or cultivating fodder crops, potentially limiting the land available for specific crop production. This can influence the choice of crops that are more efficient in terms of land use or offer higher yields on smaller land areas. As a financial asset, raising livestock acts as a form of risk management for farmers as it can provide income even during years of poor crop yields, offering financial security and stability. In general, livestock sales can provide additional income, allowing farmers to invest in improved inputs or invest in diversification strategies, including exploring new crop varieties or cultivating niche market crops. However, raising livestock requires additional labor for feeding, cleaning, and managing animals. This can influence farmers to choose crops with lower labor demands, especially if they have limited family labor available or rely heavily on hired labor.

Labor: Access to labor is crucial for various agricultural activities, including land preparation, planting, irrigation management, and harvesting. The availability of sufficient labor can influence crop choices, as some crops crops, like vegetables or fruits, often require more intensive labor compared to others, like cereals. Farmers with limited access to labor may favor less labor-intensive crops, even if they might be less profitable, due to practical constraints in managing labor demands.

Larger households typically have more readily available family labor, potentially offering greater flexibility in adopting labor-intensive crops or managing larger landholdings. However, other factors, like education levels and participation in off-farm activities, can also influence the availability of household labor within a family.

Oxen: In the context of Ethiopian agriculture, oxen represent a valuable form of capital. Owning oxen can be a proxy for both wealth and labor availability as it provides farmers with greater autonomy and control over their agricultural practices. Unlike renting oxen, which can be expensive and subject to availability constraints, owning them allows farmers to manage their workload and plan their cropping activities more effectively. The ability to harness the power of oxen opens up a wider range of crop choices for farmers, since oxen are essential for tillage and land preparation, particularly for larger landholdings or crops requiring deep plowing. Owning oxen allows farmers to cultivate these lands more efficiently, potentially leading them to consider crops that might not be feasible with manual labor alone. Oxen can be used for transporting agricultural inputs and outputs, such as seeds, fertilizers, and harvested crops. This can be particularly crucial for farmers located farther from markets or input suppliers, allowing them to access resources and transport their products more easily.

Access to credit: Financial resources play a vital role in adopting new technologies and purchasing necessary inputs. Limited access to credit can restrict the ability of farmers to invest in irrigation equipment, fertilizers, or improved seeds, potentially hindering their adoption of certain irrigation technologies or crop varieties. Conversely, access to credit can empower farmers to explore new options and invest in practices that might offer higher potential returns. In particular, the initial cost of irrigation systems can be a significant barrier for farmers, especially those with limited financial resources. Additionally, adopting certain irrigation practices might require the use of specific fertilizers to optimize crop growth and nutrient uptake. Limited access to credit can restrict farmers' ability to purchase these essential inputs, potentially hindering the effectiveness of irrigation even if the technology is adopted. I use the existence of a microfinance institution in the community as proxy to farmers' access to credit.

Off-Farm Business: Owning an off-farm business or engaging in off-farm work presents Ethiopian farmers with additional income opportunities, but it also introduces new considerations when making crop choices. The additional income can allow farmers to invest in agricultural technologies that can potentially reduce labor requirements or improve efficiency, such as small-scale irrigation systems. This can free up time for off-farm work while maintaining or even increasing crop production. Off-farm work can sometimes bring farmers into closer contact with markets and expose them to new information about diverse crops and marketing opportunities. On the other hand, off-farm work can reduce the labor available on the farm, especially during critical periods like planting and harvesting. This can influence crop choice by shifting towards less labor-intensive crops as farmers might choose crops requiring less labor throughout the growing season, such as those with longer germination periods or less demanding weeding schedules. This allows them to dedicate time to off-farm activities while managing their farm workload. In some cases, farmers might prioritize crops with higher market value even if they are more labor-intensive. The potential for increased income from these crops can incentivize them to hire additional labor or dedicate more family labor to farm work, potentially offsetting the time constraints imposed by off-farm activities.

3.4.3.3 Physical Land Quality

Physical land characteristics significantly shape irrigation and crop choice decisions, presenting both opportunities and constraints for Ethiopian farmers.

Soil: Soil type directly affects a crop's suitability for a given location. Soils with good fertility and water-holding capacity support a wider range of crops and irrigation options. In contrast, less fertile or sandy soils with a lower water-holding capacity may necessitate crops that are more drought-tolerant or require less intensive irrigation techniques. Soils with poor drainage can hinder root development and promote waterlogging, limiting the feasibility of irrigation and influencing crop selection in favor of those that can tolerate these conditions.

Potential Wetness Index[4]: The Potential Wetness Index (PWI) provides a measure of areas

[4]The "Basic Information Document" provided as an appendix to the Ethiopia Rural Socioeconomic Survey 2011-2012 describes this measurement as follows: "Local upslope contributing area and slope are combined to determine the potential wetness index: $WI = ln(As/tan(b))$ where As is flow accumulation or effective drainage area and b is slope

likely to accumulate water. This information is valuable for identifying land with higher irrigation potential and for making informed decisions about the types of irrigation systems best suited to specific areas. The PWI can inform decisions about crop selection by highlighting areas where specific crops might thrive due to existing moisture levels, reducing the need for intensive irrigation or favoring water-loving crops and varieties. This index is highly correlated with slopes. Steeper slopes may necessitate substantial land leveling to implement effective irrigation infrastructure, which can significantly increase costs and influence the types of irrigation technologies that are suitable. Irrigation on sloping terrain, if not managed carefully, can exacerbate soil erosion risks, potentially influencing farmers to either forgo irrigation altogether or prioritize crops and practices that promote soil conservation.

Climate: The amount and distribution of rainfall are crucial for both crop selection and irrigation decisions. Areas with limited rainfall necessitate choices of drought-tolerant crops and the adoption of irrigation technologies designed to maximize water-use efficiency. Conversely, areas with ample rainfall might open up a broader range of crop options and reduce the urgency for extensive irrigation infrastructure. Elevation directly impacts the local climate that crops can tolerate. Higher elevations experience cooler temperatures, potentially limiting the suitability of certain heat-loving crops. Farmers at higher elevations would be more likely to choose crops adapted to these cooler conditions, such as potatoes, peas, or certain temperate fruits. Conversely, lower elevations offer warmer environments ideal for heat-tolerant crops like maize, sorghum, or cotton. Elevation also influences rainfall patterns. Higher elevations might receive more consistent and abundant rainfall, potentially reducing the need for irrigation for some crops. Farmers in these areas might prioritize rain-fed crops that perform well with the natural precipitation. Conversely, lower elevation areas could experience less rainfall or erratic patterns, making irrigation a more critical factor for ensuring crop survival and maximizing yields. Farmers at lower elevations might be more inclined to choose drought-resistant crops or adopt irrigation systems to mitigate the risk of water scarcity. Elevation is highly correlated with mean temperature and annual precipitation, therefore I drop the latter

gradient."

variables to avoid multicollinearity while also reducing the parameters to be estimated.

3.4.3.4 Economic Incentives

Beyond the physical constraints and practical considerations, farmers are also driven by economic incentives. Farmers are economically rational actors who aim to maximize utility. The prices of different crops, coupled with the expected yields under varying irrigation scenarios, play an important role in their decision-making. Farmers who are market-oriented will be inclined to adopt irrigation technologies and choose crops that offer the highest potential return on investment based on prevailing market prices. Prices also influence how farmers manage risk. Crops with higher market prices might incentivize adoption of irrigation even if it involves greater initial costs or water usage, as the potential for higher profits justifies the risk. Conversely, crops with lower market prices may lead farmers to prioritize water-saving techniques or drought-resistant varieties, even if it means sacrificing some potential yield, to minimize risks and ensure a decent return even in the face of fluctuating market conditions. Furthermore, low output prices discourage farmers from adopting new technologies like efficient irrigation practices, as they often cannot afford the initial investment. This creates a disincentive for them to seek out or actively utilize irrigation advice from extension agents (Leta et al., 2017). Crop prices in the dataset are generated by dividing the total revenue generated from selling all of part of the harvest by the quantity sold.

However, prices don't operate in isolation. They interact with other factors in the model. The distance to markets and transportation costs can significantly impact the final price received by farmers. Even if a crop has a high market price overall, the associated costs of getting it to the market can reduce its profitability for individual farmers, influencing their choices. Moreover, perishable goods, such as fruits and vegetables, are particularly sensitive to the distance to markets and the quality of transportation infrastructure. Perishable goods have a limited shelf life, making timely transportation to markets crucial. The longer the distance to markets, the greater the risk of spoilage, quality deterioration, and loss of value. Farmers located further from markets are at a disadvantage and may be forced to sell at lower prices to compensate for potential quality losses during transit. They might even be driven out of the market and choose a consumption based crop choice rather than a market-oriented optimization problem. Farmers close to markets have more flexibility in choosing high-value perishable crops, while those in remote areas may opt for less perishable commodities, potentially resulting in lower income potential.

3.4.4 Data

I use data from the Living Standards Measurement Study-Integrated Survey on Agriculture (LSMS-ISA) for Ethiopia. The LSMS-ISA is a socio-economic panel dataset collected at the household level, with a specific focus on capturing detailed information related to agricultural practices. This dataset offers a useful resource for analyzing the joint decision-making processes of Ethiopian farmers regarding crop selection and irrigation adoption. The Ethiopian LSMS-ISA data encompasses three survey rounds[5] conducted across a three-year period: 2011/2012, 2013/2014, and 2015/2016. The initial round included data from 3,776 rural households. The second round expanded the sample size to 5,262 by incorporating households from urban areas. However, since I focus on rural agricultural practices, I exclude the additional 1,486 urban households recruited in the second round due to their low participation in farming activities. This exclusion, along with the application of population weights, ensures the remaining sample is representative of rural households at both the national and regional levels in Ethiopia. Data collection for each round was conducted in three waves annually. This multi-wave approach captures critical information at different stages of the agricultural cycle, including post-planting and post-harvest periods. The dataset is organized with different observation levels and includes variables categorized within distinct subsets. The specific subsets utilized in this analysis include (1) Community-level data which provides information on the proximity of households to credit services and public irrigation schemes, (2) Household-level data which encompasses basic demographic characteristics such as gender, age, and education level of the household head, farm type classification, and ownership of non-farm assets, (3) Field-level data, such that detailed information is collected on each individual field, including reported area, soil type, potential wetness index, slope, cultivated crops, applied inputs, harvest yields, access to extension services, and irrigation practices, and (4) Climate data such as mean annual temperature

[5] There is a fourth survey wave in 2017/2018 with a completely new sample, which would now allow for a panel data structure. Also note that the data is not affected by the Tigray war which happened in 2020-2022.

and cumulative precipitation. The complete dataset incorporates a vast array of 125 different crops. However, to accommodate the limitations imposed by household-specific climate and soil variables within the chosen analytical framework, the analysis focuses on a strategically selected subset of crops. This subset is derived by clustering crops according to their agronomic group. This categorization process results in eight distinct crop categories: cereals, pulses, roots, vegetables, fruits, spices, oilseeds, and cash crops. Tables 3B.1 and 3B.2 show the frequency of crops in the dataset. Another useful way of thinking of the crop-irrigation alternative choice is to look at Tables 3B.3 and 3B.4, which summarize the frequency of selection of crop types and crop with irrigation adoption, respectively.

The selection of explanatory variables for the estimation of the structural model is guided by the underlying theoretical framework. These variables aim to comprehensively explain the choice probabilities associated with crop selection and irrigation adoption. The initial set of variables underwent a screening process to identify and eliminate collinear variables. This ensures the model avoids potential issues arising from multicollinearity, ultimately leading to more robust and reliable estimates.

3.4.4.1 Descriptive statistics

To ensure a clear link between the conceptual framework and the empirical model, Table 3B.5 explicitly shows how each variable included in the empirical model (e.g., distance to nearest main road) serves as a proxy for a broader conceptual variable outlined in the theoretical framework (e.g., market access). Descriptive statistics are reported in Table 3B.6. These descriptive statistics provide a foundational understanding of the Ethiopian smallholder farming context and the characteristics of the farmers included in this study. The data reveals a resource-constrained environment with limited access to education, credit, and irrigation. Educational attainment appears limited, with only 36% of farmers reporting ever attending school. Extension program participation is moderately high, with over half (56.5%) of farmers participating. Land tenure security varies, with 58% of households possessing a certificate for their parcels. Land area ranges from 0 to over 4 million square meters, with an average of approximately 0.145 hectares. Livestock ownership is common, with 42.1% of farmers owning livestock. Oxen ownership is less prevalent, with an average of 1.2 oxen per owner. Off-farm business activity is present but not widespread, with 31.7% of households reporting someone owning an off-farm business. Access to credit services is limited, with only 24% of farmers utilizing them. The average size of household is relatively small, with an average of 4 members providing labor on the farm. Distance to the nearest major road presents a challenge, averaging 14.7 kilometers. While the existence of irrigation schemes availability is notable, with 62.1% of communities having access to one, Irrigated land is a small portion of the total cultivated area, with only 6% of fields using irrigation. Plot elevation varies considerably, ranging from 192 to 3,419 meters above sea level. The Potential Wetness Index (PWI) indicates a moderately dry environment, with an average score of 12.6%. The table also includes average prices for all crop categories, showing that cereals have the highest price and to fruits having the lowest. Crop prices are calculated by dividing the total amount of money received for selling all or part of the harvest by the quantity sold. The crop prices are then averaged across all crop within a given crop category and across years. The resulting average crop type prices are reported in the table. While the code book of the survey shows that prices must be reported in Birr/kg, there seems to be some discrepancies in these price data, particularly with the inconsistency of the order of magnitude of prices. Table 3B.7 serves as a reference to compare the reported prices to those found in other sources such as the Ethiopian Central Statistical Agency (CSA, 2011) and the Food and Agriculture Organization Statistical Database (FAOSTAT, 2024). Descriptive statistics of soil types, presented in Table 3B.8, show that Vertisol and Luvisol dominate soil types in the sample. Finally, Table 3B.9 reports descriptive statistics of variables relating to the subsistence nature of Ethiopian farmers in the sample. Only 17% of respondents sell any of their harvest, and the average distance of households to the nearest market stands at 68 kilometers. The variable *HH consumption* represents the proportion of crop product that is used directly for household consumption, and not for sale or payments in kind. This is calculated as an average across households in the sample to allow it to enter as a generic alternative-specific variable in the choice model.

The analysis that follows will explore how these factors and others influence the joint decision-making of crop choice and irrigation adoption.

3.5 Results

This study employs a multi-pronged approach to model selection and robustness checks. I estimate and compare the results of several models, including nested logit and conditional logit, with and without crop price (including both the original survey data and the adjusted price version). The selection of the preferred model is guided by theoretical consistency with the decision-making framework (following Lau (1986)) and goodness-of-fit statistics. Among the estimated models, the nested logit with 13 alternatives and the original sample crop price offers the highest log-likelihood function value, as shown in Table 3B.10. I interpret the results based on this preferred model. Furthermore, to assess the sensitivity of the main findings, I conduct robustness checks by excluding crop price, using adjusted crop prices, and expanding the choice set to include 16 alternatives. These checks are used to evaluate whether the core conclusions regarding the impact of extension services, irrigation availability, and household consumption on crop choices remain consistent. Tables 3B.11 and 3B.12 summarize the results from the nested logit choice model estimation, using all eight crop groups with three degenerate nests and the baseline alternative defined as cereals. I report farm-specific estimates of the effects of the chosen independent variables in rows of Table 3B.11. The coefficients on generic variables and the crop-specific estimates of the IV parameters are reported in Table 3B.12. While the focus is on the effect of agricultural extension on crop choice and irrigation adoption, there are two important considerations according to which I structure this section: the estimated effect of crop prices on decision making, and the conclusions on the preferred model specification.

Looking at the determinants of choice in Table 3B.11, results show that, compared to cereals, being part of an extension program increases the likelihood of choosing cash crops, but decreases the probability of choosing fruits, and has no statistically significant effect on the likelihood of other choices. This result confirms that extension in Ethiopia, specifically through farmers' cooperative unions, typically focuses on strategic commodities like coffee (Jena et al., 2012; Leta et al., 2017). Education has a positive and statistically significant effect on choosing pulses, vegetables, fruits, and cash crops, and no effect on choosing any of the other crop types. Surprisingly, land tenure is shown to negatively impact the likelihood of choosing cash crops and fruits while positively impacting the likelihood of choosing pulses and oilseeds. The interpretation of the coefficient on 'land tenure' merits further discussion due to the way crops are categorized. The current grouping (cash crops and fruit crops) combines annual and perennial crops. This presents a challenge, as secure land tenure is often more critical for decisions regarding perennial crops, which require long-term investments. The "cash crop" category, for example, includes annual crops like sesame and tobacco alongside perennial trees. Similarly, the "fruits" category mixes tree fruits (perennial) with watermelons (annual). This clustering might mask the true relationship between land tenure security and crop choices, particularly for perennial crops that benefit more from long-term land use rights. Land area, on the other hand, has no meaningful effect on crop choice as the estimates are too small. Owning livestock and having access to credit are variables used to proxy farm assets and the level of wealth of the farmer. Results show that, compared to cereals, owning livestock increases the likelihood of choosing cash crops and roots. Additionally, access to credit increases the likelihood of choosing oilseeds while decreasing the probability of choosing pulses and fruits. Oxen is usually needed extensively in preparing soil for cereals and is less needed for labor related to fruits or cash crops. This is supported by the negative sign on the coefficient on oxen for the latter crop types. Seeing that participating in off-farm activities increases the likelihood of choosing cash crops, fruits, and also roots and spices, it follows from the hypothesis formulated in the previous section that off-farm activities offer more venues to market crops as well as more revenue to support farming of these crops, without constraining available labor. As expected, increased distance from a household to a main road decreases the likelihood of choosing fruits, which are the most perishable products in the choice set, hence the hardest to transport to a market while preserving quality. Results also show that the existence of an irrigation scheme in the community does not increase the likelihood of choosing crops that typically have higher water needs such as fruits and vegetables. The effects of other control variables such as soil type, elevation, and PWI are also reported in the same table.

To explain the counter-intuitive result, in Table 3B.12, suggesting that the likelihood of choosing a crop decreases with its price, I report the crop prices in Table 3B.6. The reported prices in the sample show that fruits have by far the lowest prices. This leads to the conclusion that, under the held assumption that a profit-maximizing farmer would choose crops which generate higher revenues, either (1) prices are misreported in the survey, (2) the crop clustering is misspecified, or (3) farmers are optimizing a different utility function. To stay consistent with the theoretical framework and the data source, I keep the price variable and add the variable representing the average household consumption share for each crop. As posited in the theoretical model of the farmer's problem, the constraint facing subsistence farmers is meeting the household's level of food consumption needs. As expected from the hypothesis, the likelihood of choosing a crop increases with the share of harvest that is used for direct household consumption of that crop.

An intuitive robustness check is to re-estimate the nested logit choice model after excluding the crop prices from the vector of alternative-specific variables. The results of this specification are reported in Tables 3B.13 and 3B.14. The significance level of the coefficient on household consumption remains the same and the magnitude of the effect increases only slightly. However, the unconstrained dissimilarity parameters in this specification are greater than one. This indicates that the model is not fully consistent with utility-maximizing behavior for some ranges of the independent variables. Therefore, the specification that includes crop prices is preferred, keeping in mind the reported crop prices. Results also show that the effect of extension on crop choices aligns closely with the results from the first nested logit model specification, exhibiting consistency in both sign, magnitude, and statistical significance.

While ideally the analysis would rely on the crop prices reported directly by respondents in the survey, there is reason to believe these prices, particularly for fruits and possibly spices, might be misreported. This suspicion arises from observing fruit prices being an order of magnitude lower compared to other crops. Consulting external sources like the Ethiopian CSA (CSA, 2011) and the FAO Statistical Database (FAOSTAT, 2024) reveals that fruit prices, while still lower than cereals, are closer to the prices of other crop types (see Table 3B.7). To address this potential misreporting and assess its impact on the results, a robustness check is employed. Specifically, the price of fruits is adjusted by multiplying it by a factor of 10. This represents a conservative approach that maintains some consistency with the original survey data while correcting for the suspected underestimation. The nested logit model is then re-estimated using the price-adjusted data. The results from this robustness check are particularly noteworthy. The sign of the estimated coefficient on price flips, becoming positive, as shown in Table 3B.16. This aligns more closely with economic theory and the formulated theoretical model, where higher prices for a crop-irrigation alternative are expected to incentivize its choice. The effect of extension remains consistent with prior results.

Looking back at the dissimilarity parameters estimates in Table 3B.12, note that $\lambda_k = 1$ for the three nests has been deliberately imposed on the model since the occurrence of irrigated oilseeds, roots, and spices is so low that the sensible choice of dropping them from the sample had to be made. While the log-sum coefficients on cereals, fruits, and vegetables are close to 1, I find from testing the null hypothesis $H_0 : \lambda_k = 1 \forall k$ that I can reject the null at the 1% significance level, meaning that the dissimilarity parameters are sufficiently different from zero to justify a nested logit structure instead of a standard logit. Nonetheless, only the coefficient on cash crops indicates a relatively high dependence between irrigated and rainfed choices. This suggests that a conditinal logit can also be fitted, allowing more independence between the crop-irrigation alternatives.

Results of the conditional logit are outlined in Tables 3B.17, 3B.18, and 3B.19. Although all estimates come from the same estimated model, farm-specific estimates of the effects of the chosen independent variables are reported in rows of Table 3B.17 for the crop types cereals, pulses, oilseeds, and vegetables, and Table 3B.18 for the remaining crop types for readability. I report the marginal effect of crop price and household consumption share on the probability of crop-irrigation outcomes from the conditional logit choice model in Table 3B.19. The main difference between the nested logit and the conditional logit is the assumption that the alternatives are completely independent in the latter, hence the full set of alternative-specific estimates instead of the crop type-specific estimates. The first remark is that these coefficients are similar to the ones obtained from the nested logit model. The statistical significance of the coefficient on household consumption is higher when estimated using the conditional logit, and its magnitude is closer to the one obtained with the nested logit that excludes price as an explanatory generic variable. Moving to the estimate of interest, results show that, compared to irrigated cereals, agricultural extension increases the likelihood of choosing all other rainfed crops, but has no effect on the irrigated crops. If this effect is thought of as a separate effect on crop choice and on irrigation adoption, it would suggest that given all possible choices, extension has no statistically significant effect on irrigation adoption for any crop type. Results also show that while the existence of an irrigation scheme in the community does not increase the likelihood of choosing crops that typically have higher water needs, it does negatively impact the likelihood of choosing rainfed crops.

While the independence of irrelevant alternatives assumption is more restrictive in the conditional logit model, it offers an opportunity to include the crop-irrigation alternatives that were dropped in the previous specification without losing the general concavity of the log-likelihood function. As a robustness check, I estimate the effect of the chosen determinants of choice on the probability of choosing a crop-irrigation alternative when all 16 possible alternatives are included in the farmer's choice set, using a conditional logit model. The results of this specification are reported in Tables 3B.20, 3B.21, and 3B.22. Starting with the estimated effects of the alternative-specific variables in Table 3B.20, it is clear that, similar to the previous specifications, these results show that the likelihood of choosing a crop-irrigation alternative increases with the share of the crop harvest used for direct household consumption, further confirming the hypothesis formulated in the theoretical framework about the constraint faced by subsistence farmers. This hypothesis posits that smallholder farmers in Ethiopia are subsistence farmers, relying on their farming operations to meet their dietary needs, which effectively translates to a constraint on the share of crop that is used directly in household consumption.

The results from the 16 alternatives specification of the conditional logit for the farm-specific variables are presented in Tables 3B.21, and 3B.22. Results suggest that extension significantly increases the likelihood of choosing all rainfed crops, relative to cereals. Extension has less of an impact on irrigated crops, with no significant effect on the likelihood of the added alternatives of irrigated oilseeds and irrigated spices, and a negative effect on the probability of choosing irrigated roots relative to cereals. The marginal effects of the independent variables on the likelihood of alternative outcomes remain the same between the specification that includes 16 alternatives compared to 13 alternatives, suggesting that the IIA assumption holds. Specifically, the marginal effects of agricultural extension on the likelihood of rainfed oilseeds, roots, and spices remain the same after introducing the possibility of irrigating these specific crops. Results also suggest that having an irrigation scheme in the village has a negative and significant effect on the likelihood of choosing rainfed crops. While coefficients are positive, we do not see a significant effect of irrigation schemes on irrigated crops. This may be due to the low take-up of irrigation in the overall sample, as shown in Table 3B.23. Although not causal, Table 3B.23 also provides some descriptive evidence that irrigation take-up is higher when irrigation schemes exist.

In total, with similar coefficients across models, including those in Table 3B.11, results suggest that irrigation schemes, combined with extension services, are not making a substantial impact on increasing irrigation adoption. This is not surprising given the context of the current literature. The lack of expertise in irrigation development among extension agents has been documented in Leta et al. (2017), suggesting that farmers are typically encouraged to keep practicing crops that they are familiar with, rather than adopting new technology. Gebremariam et al. (2021) find further evidence that extension through water users' associations often does not meet the farmers' needs in terms of access to irrigation. Another layer pertaining to governance issues in Ethiopia adds to the complexity of the joint crop choice and irrigation adoption decision making. Gebremariam et al. (2021) suggest that Ethiopian farmers are afraid that the government would take away their irrigated land with no fair compensation, hence their reluctance to either adopt irrigation or disclose such technology uptake.

Instead, results suggest that extension strongly increases the probability of choosing cash crops, and generally encourages rainfed alternatives, although it does not necessarily discourage irrigation. The negative coefficient of extension on fruits can be interpreted by looking at the effect of the distance to the nearest road on the likelihood of choosing fruits. Market avenues for cash crops are well established in Ethiopia, especially for coffee. However, fruits bear higher risk for farmers due to the limited access to local markets. Household consumption significantly impacts farmers' choices. Finally, results from the log-likelihood ratio test for the IIA assumption, which is a joint test that $\lambda_k = 1 \forall k$, reveals that the null that all of the log-sum coefficients are 1 can be rejected with 99% confidence, although the dissimilarity parameters are very close to 1. This indicates that while the conditional logit is informative, namely by looking at predictive margins in Table 3B.24, there is still reason to believe that the nested logit is slightly more appropriate, upholding the assumption that crop choice and irrigation adoption decisions are made jointly.

3.6 Conclusion

Agriculture plays a critical role in Ethiopia's food security and economic growth. However, the sector faces challenges like low yields, weather variability, and limited access to technology and inputs. Improving productivity and reducing variability are essential for enhancing food security and farmer livelihoods. Ethiopian agriculture is dominated by smallholders, whose decision-making regarding crop choice and irrigation adoption significantly impacts production outcomes.

This study explores the determinants of joint decision-making regarding crop choice and irrigation adoption among smallholder farmers in Ethiopia, leveraging data from the Ethiopian Living Standards Measurement Study-Integrated Survey on Agriculture (LSMS-ISA). I start by formulating a theoretical model that explicitly considers the farmer's optimization problem as a simultaneous decision on crop choice and irrigation uptake, while incorporating a crucial constraint on household consumption, reflecting the subsistence needs of these farmers. I then use a nested logit choice model for the empirical application, aligning with the theoretical framework's nested structure of choices, such that irrigation adoption is nested within the crop category. The analysis focuses on eight distinct crop categories and includes a comprehensive set of explanatory variables such as crop prices, household consumption patterns, access to agricultural extension services, land tenure arrangements, and various household characteristics. I also consider the possibility of some independence between irrigated and rainfed choices within certain crop categories. To address this question, I complement the analysis by using a conditional logit model. This model offers a different assumption of independence between alternatives, allowing for a more nuanced understanding of the decision-making process.

A key insight from this study is the confirmation that Ethiopian smallholder farmers prioritize subsistence needs over pure profit maximization when making crop choices. This finding aligns with the understanding that these farmers are primarily resource-constrained and operate within a subsistence agricultural context. The small effect of crop prices on crop choice likelihood, despite its counter-intuitive nature, can be interpreted in this light. It suggests factors beyond simply maximizing revenue influence decision-making. The strong positive influence of household consumption share on crop choice further reinforces this notion. Farmers are more likely to choose crops that directly contribute to meeting their household's food needs, highlighting the critical role of subsistence production in their decision-making process. This focus on subsistence needs is further supported by the mixed impact of agricultural extension services on crop choices. While extension positively affects cash crop selection, potentially reflecting a desire for increased income, it discourages farmers from choosing fruits. Fruits, being perishable and often requiring specific market access for good returns, might be perceived as riskier by subsistence farmers compared to staple crops that directly contribute to household food security. This is supported by the negative effect of the distance to the nearest main road on the likelihood of choosing fruits.

Overall, I find that irrigation schemes, combined with agricultural extension services, have minimal to no effect on increasing irrigation adoption in Ethiopia. Agriculture dominates Ethiopia's water use, consuming a staggering 92% of withdrawals. Despite this, only 7% of the sampled cultivated land benefits from irrigation. The prevalent form, spate irrigation, relies on unpredictable floodwaters, limiting its effectiveness in mitigating rainfall variability. While long-term water efficiency goal from making irrigation available might involve switching to water-loving crops, the initial focus is likely to be on increasing existing crop production, directly impacting food security through staple crops familiar to local dietary needs.

This study sheds light on the complex decision-making processes of Ethiopian smallholder farmers regarding crop selection and irrigation adoption. By examining the combined effects of extension services and irrigation on crop choices, this paper offers valuable insights that contribute to the understanding of agricultural adaptation strategies in water-scarce environments. These findings can inform the design of agricultural extension programs and policies that better target the needs of Ethiopian farmers, ultimately contributing to improved agricultural productivity and food security. In particular, this study highlights the crucial role of subsistence needs in driving crop choice, with farmers prioritizing crops directly contributing to household food security. This underscores the importance of extension programs that consider this context. Tailoring extension services to address both cash crop opportunities and subsistence needs, such as promoting high-yielding food staples alongside market-oriented cash crops, could enhance food security while generating income. Additionally, the limited impact of extension on fruit selection, potentially due to remoteness and perishability concerns, suggests a need for complementary interventions. Investments in infrastructure like improved rural roads alongside extension services focused on post-harvest handling and marketing strategies for fruits could incentivize their production and improve market access for Ethiopian smallholder farmers. Furthermore, the findings pertaining to the very limited effect of irrigation schemes and agricultural extension services on irrigation adoption highlight a need for policymakers to explore additional strategies alongside extension services, such as addressing financing constraints, technical training on water management practices, or investigating alternative, more affordable irrigation technologies that might be better suited to the resource limitations of smallholder farmers.

BIBLIOGRAPHY

Adesina, A. A. and Zinnah, M. M. (1993). Technology characteristics, farmers' perceptions and adoption decisions: A tobit model application in Sierra Leone. *Agricultural Economics*, 9(4):297–311.

Amghani, M. S., Mojtahedi, M., and Savari, M. (2023). An economic assessment of the effect of agricultural extension activities on increasing production and farmers' revenue: The case of irrigated wheat production-extension model sites in Iran. *Scientific Reports*.

Arellano-Gonzalez, J. and Moore, F. C. (2020). Intertemporal arbitrage of water and long-term agricultural investments: Drought, groundwater banking, and perennial cropping decisions in California. *American Journal of Agricultural Economics*, 102(5):1368–1382.

Bekabil, U. T. (2014). Review of challenges and prospects of agricultural production and productivity in Ethiopia. *Journal of Natural Sciences Research*, 4(18):70–77.

CSA, E. (2011). Area and Production of Crops and Crop Utilization. Data retrieved from CSA, https://www.statsethiopia.gov.et/our-survey-reports/.

Duflo, E. and Pande, R. (2007). Dams. *The Quarterly Journal of Economics*, 122(2):601–646.

FAOSTAT (2024). Producer prices. Data retrieved from FAOSTAT, http://https://www.fao.org/faostat/en/data/PP.

Feder, G., Just, R. E., and Zilberman, D. (1985). Adoption of agricultural innovations in developing countries: A survey. *Economic development and cultural change*, 33(2):255–298.

Feinerman, E. and Tsur, Y. (2014). Perennial crops under stochastic water supply. *Agricultural economics*, 45(6):757–766.

Fisher, F. and Huber-Lee, A. (2012). *Liquid Assets: An Economic Approach for Water Management and Conflict Resolution in the Middle East and Beyond*. Routledge.

Forinash, C. V. and Koppelman, F. S. (1993). Application and interpretation of nested logit models of intercity mode choice. *Transportation Research Record*, (1413).

Gebremariam, Y. A., Dessein, J., Wondimagegnhu, B. A., Breusers, M., Lenaerts, L., Adgo, E., Ayalew, Z., Minale, A. S., and Nyssen, J. (2021). Determinants of farmers' level of interaction with agricultural extension agencies in northwest Ethiopia. *Sustainability*, 13(6):3447.

Gebremedhin, B. and Swinton, S. M. (2003). Investment in soil conservation in northern Ethiopia: the role of land tenure security and public programs. *Agricultural Economics*, 29(1):69–84.

Gebreselassie, S., Haile, M., and Kalkuhl, M. (2017). The wheat sector in ethiopia: Current status and key challenges for future value chain development. *ZEF working paper*.

Genius, M., Koundouri, P., Nauges, C., and Tzouvelekas, V. (2014). Information transmission in irrigation technology adoption and diffusion: Social learning, extension services, and spatial effects. *American Journal of Agricultural Economics*, 96(1):328–344.

Green, G., Sunding, D., Zilberman, D., and Parker, D. (1996). Explaining irrigation technology choices: a microparameter approach. *American Journal of Agricultural Economics*, 78(4):1064–1072.

Green, G. P. and Sunding, D. L. (1997). Land allocation, soil quality, and the demand for irrigation technology. *Journal of Agricultural and Resource Economics*, pages 367–375.

Hatch, N. R., Daniel, D., and Pande, S. (2022). Behavioral and socio-economic factors controlling irrigation adoption in Maharashtra, India. *Hydrological Sciences Journal*, 67(6):847–857.

Headey, D., Dereje, M., and Taffesse, A. S. (2014). Land constraints and agricultural intensification in Ethiopia: A village-level analysis of high-potential areas. *Food Policy*, 48:129–141.

Headey, D. D. and Jayne, T. S. (2014). Adaptation to land constraints: Is Africa different? *Food Policy*, 48:18–33.

Heiss, F. (2002). Structural choice analysis with nested logit models. *The Stata Journal*, 2(3):227–252.

Hensher, D. A., Rose, J. M., and Greene, W. H. (2005). *Applied Choice Analysis: a Primer*. Cambridge University Press.

Holden, S. T. and Lunduka, R. W. (2014). Input subsidies, cash constraints, and timing of input supply. *American Journal of Agricultural Economics*, 96(1):290–307.

Hornum, S. T., Bolwig, S., and Trærup, S. L. M. (2023). Enabling drip irrigation technology diffusion and adoption among African smallholders: a document analysis of climate action plans and donor projects. *Climate and Development*, pages 1–17.

Jayne, T. S., Chamberlin, J., and Headey, D. D. (2014). Land pressures, the evolution of farming systems, and development strategies in Africa: A synthesis. *Food Policy*, 48:1–17.

Jena, P. R., Chichaibelu, B. B., Stellmacher, T., and Grote, U. (2012). The impact of coffee certification on small-scale producers' livelihoods: a case study from the Jimma Zone, Ethiopia. *Agricultural Economics*, 43(4):429–440.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129(2):597–652.

Knippenberg, E., Jolliffe, D., and Hoddinott, J. (2020). Land fragmentation and food insecurity in Ethiopia. *American Journal of Agricultural Economics*, 102(5):1557–1577.

Koundouri, P., Nauges, C., and Tzouvelekas, V. (2006). Technology adoption under production uncertainty: Theory and application to irrigation technology. *American Journal of Agricultural Economics*, 88(3):657–670.

Krishnan, P. and Patnam, M. (2014). Neighbors and extension agents in Ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, 96(1):308–327.

Lau, L. J. (1986). Functional forms in econometric model building. *Handbook of Econometrics*, 3:1515–1566.

Leta, G., Kelboro, G., Stellmacher, T., and Hornidge, A.-K. (2017). The agricultural extension system in Ethiopia: operational setup, challenges and opportunities. *ZEF working paper series*.

Lichtenberg, E. (1989). Land quality, irrigation development, and cropping patterns in the northern high plains. *American Journal of Agricultural Economics*, 71(1):187–194.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Number 3. Cambridge University Press.

Marques, G. F., Lund, J. R., and Howitt, R. E. (2005). Modeling irrigated agricultural production and water use decisions under water supply uncertainty. *Water resources research*, 41(8).

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4):303–328.

Melesse, M. B. and Bulte, E. (2015). Does land registration and certification boost farm productivity? Evidence from Ethiopia. *Agricultural Economics*, 46(6):757–768.

Mesfin, H., Tessema, Y. M., Tirivayi, N., and Nillesen, E. (2022). The impact of agricultural extension service on the uptake of various agricultural technologies in Ethiopia. *Africa Development*, 47(4):77–105.

Mohammed, S. and Abdulai, A. (2022). Impacts of extension dissemination and technology adoption on farmers' efficiency and welfare in Ghana: Evidence from legume inoculant technology. *Frontiers in Sustainable Food Systems*, 6:1025225.

Moore, M. R., Gollehon, N. R., and Carey, M. B. (1994). Multicrop production decisions in western irrigated agriculture: the role of water price. *American Journal of Agricultural Economics*, 76(4):859–874.

Moreno, G. and Sunding, D. L. (2005). Joint estimation of technology adoption and land allocation with implications for the design of conservation policy. *American Journal of Agricultural*

Economics, 87(4):1009–1019.

Nkamleu, G. B. and Adesina, A. A. (2000). Determinants of chemical input use in peri-urban lowland systems: bivariate probit analysis in Cameroon. *Agricultural Systems*, 63(2):111–121.

Nuru, S. A. and Seebens, H. (2008). The impact of location on crop choice and rural livelihood: evidences from villages in northern ethiopia. *ZEF discussion papers on development policy*, (123).

Ojukwu, C., Omoluabi, C., Samba, I., Bisset, P., Pitamber, S., and Birhane, G. (2001). Ethiopia - Koga Irrigation and Watershed Management Project - Appraisal Report.

Pan, Y., Smith, S. C., and Sulaiman, M. (2018). Agricultural extension and technology adoption for food security: Evidence from Uganda. *American Journal of Agricultural Economics*, 100(4):1012–1031.

Rosenzweig, M. and Udry, C. R. (2013). Forecasting profitability. Technical report, National Bureau of Economic Research.

Rudnick, D. R., Stockton, M., Taghvaeian, S., Warren, J., Dukes, M. D., Kremen, A., Henry, C. G., Aguilar, J., Ortiz, B., Andales, A., et al. (2020). Innovative extension methods in the US to promote irrigation water management. *Transactions of the ASABE*, 63(5):1549–1558.

Seo, S. N. and Mendelsohn, R. O. (2007). Climate change adaptation in Africa: a microeconomic analysis of livestock choice. *World Bank Policy Research Working Paper*, (4277).

Shangshon, B. B., Polthanee, A., Wongsamun, C., and Suriya, P. (2023). Impact of agricultural extension services on cereal production among rural farmers in Bhutan. *Journal of Agricultural Extension*, 27(1):13–26.

Shitu, K. and Almaw, M. (2020). Review on hydrological and environmental challenges for irrigation agriculture development in Ethiopia. *Hydrology*, 8(4):86.

Shitu, K., Hymiro, A., Tesfaw, M., et al. (2022). Review on baste soil improving approaches for water logging and soil salinization problems in agricultural land of Ethiopia. *Soil Water Sci*, 6:229–235.

Shrestha, R. B. and Gopalakrishnan, C. (1993). Adoption and diffusion of drip irrigation technology: an econometric analysis. *Economic Development and Cultural Change*, 41(2):407–418.

Sumberg, J. (2005). Constraints to the adoption of agricultural innovations: Is it time for a re-think? *Outlook on Agriculture*, 34(1):7–10.

Taffesse, A. S., Dorosh, P., and Gemessa, S. A. (2012). Crop production in Ethiopia: Regional patterns and trends. *Food and Agriculture in Ethiopia: Progress and Policy Challenges*, pages

53–83.

Taraz, V. (2017). Adaptation to climate change: Historical evidence from the Indian Monsoon. *Environment and Development Economics*, 22(5):517–545.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418(6898):671–677.

Wakeyo, M. B. and Gardebroek, C. (2013). Does water harvesting induce fertilizer use among smallholders? Evidence from Ethiopia. *Agricultural Systems*, 114:54–63.

Wang, J., Mendelsohn, R. O., Dinar, A., and Huang, J. (2008). How China's farmers adapt to climate change. *World Bank Policy Research Working Paper*, (4758).

Winer, N. (1989). Agriculture and food security in Ethiopia. *Disasters*, 13(1):1–8.

Zerfu, D. and Larson, D. F. (2010). Incomplete markets and fertilizer use: evidence from ethiopia. *World Bank Policy Research Working Paper*, (5235).

APPENDIX 3A

FIGURES

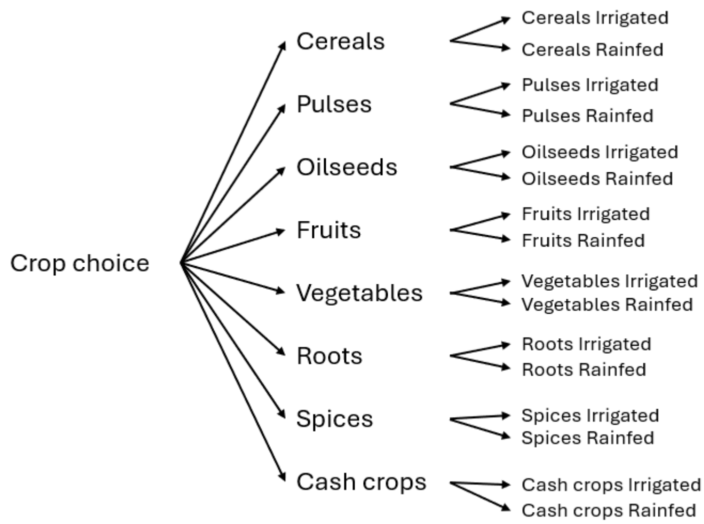

Figure 3A.1 Nested structure of irrigation adoption and crop choice

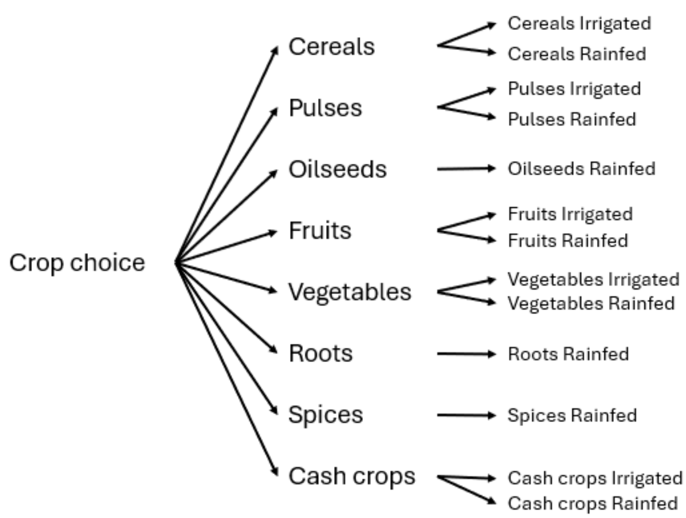

Figure 3A.2 Nested tree structure specified for the nested logit model with 13 alternatives

APPENDIX 3B

TABLES

Table 3B.1 Crop groups and frequency of crops in the dataset

Crop Group	Crop	Freq.	Percent
Cereals/Grains	Barley	4,352	4.86
	Maize	10,071	11.24
	Millet	1,770	1.97
	Oats	179	0.20
	Rice	66	0.07
	Sorghum	7,621	8.50
	Teff	6,593	7.36
	Wheat	4,326	4.83
Pulses	Chick Peas	821	0.92
	Fenugreek	573	0.64
	Field Peas	1,708	1.91
	Gibto	123	0.14
	Haricot Beans	2,387	2.66
	Horse Beans	3,090	3.45
	Lentils	660	0.74
	Vetch	526	0.59
Oilseeds	Cotton Seed	3	0.00
	Linseed	848	0.95
	Nueg	1,004	1.12
	Rape Seed	920	1.03
	Sunflower	211	0.24
Spices	Cardamon	241	0.27
	Chilies	53	0.06
	Coriander	163	0.18
	Ginger	166	0.19
	Kazmir	49	0.05
Roots	Beer Root	296	0.33
	Garlic	1,311	1.46
	Onion	617	0.69
	Potato	1,004	1.12
	Sweet Potato	1,595	1.78

Table 3B.2 Crop groups and frequency of crops in the dataset, continued

Crop Group	Crop	Freq.	Percent
Fruits	Apples	38	0.04
	Avocados	1,306	1.46
	Bananas	2,450	2.73
	Cactus	550	0.61
	Citron	28	0.03
	Gishita	78	0.09
	Godere	2,579	2.88
	Guava	311	0.35
	Mangos	1,287	1.44
	Oranges	422	0.47
	Papaya	704	0.79
	Peach	127	0.14
	Pinapples	48	0.05
	Watermelon	5	0.01
	Other Fruits	576	0.64
Vegetables	Cabbage	347	0.39
	Carrot	89	0.10
	Cauliflower	48	0.05
	Green Pepper	940	1.05
	Kale	3,382	3.77
	Pumpkins	1,592	1.78
	Red Pepper	1,666	1.86
	Roman	70	0.08
	Spinach	82	0.09
	Tomatoes	380	0.42
Cash crops	Chat	2,648	2.95
	Coffee	5,010	5.59
	Enset	4,580	5.11
	Gesho	1,373	1.53
	Sesame	1,052	1.17
	Sugar Cane	888	0.99
	Tobacco	253	0.28

Table 3B.3 Alternatives summary for crop groups

Crop	Cases	Frequency	Percent
Cereals	29,046	5,669	39.03
Pulses	29,046	1,286	8.85
Oilseeds	14,523	455	3.13
Vegetables	29,046	1,331	9.16
Roots	14,523	612	4.21
Fruits	29,046	2,184	15.04
Spices	14,523	331	2.28
Cash	29,046	2,655	18.28

Table 3B.4 Alternatives summary for crop-irrigation alternatives

Alternative	Cases	Frequency	Percent
Cereals Irrigated	14,523	137	0.94
Cereals Rainfed	14,523	5,532	38.09
Pulses Irrigated	14,523	9	0.06
Pulses Rainfed	14,523	1,277	8.79
Oilseeds Rainfed	14,523	455	3.13
Vegetables Irrigated	14,523	70	0.48
Vegetables Rainfed	14,523	1,261	8.68
Roots Rainfed	14,523	612	4.21
Fruits Irrigated	14,523	192	1.32
Fruits Rainfed	14,523	1,992	13.72
Spices Rainfed	14,523	331	2.28
Cash Irrigated	14,523	220	1.51
Cash Rainfed	14,523	2,435	16.77

Table 3B.5 Conceptual model variables and their corresponding empirical model variables

Variables	Conceptual Model	Empirical Model
Human Capital Assets	Education	Education
	Extension	Extension
Farm Assets	Land tenure	Land tenure
	Land area	Land area
	Livestock	Livestock
	Labor	HH members
	Oxen	Oxen
	Off-farm Business	Off-farm
	Access to credit	Credit
	Irrigation availability	Irr. Scheme
Physical Land Quality	Soil	Soil type
	Potential Wetness Index	PWI
	Climate/Elevation	Elevation
Economic Incentives	Crop prices	Price
	Access to markets	Dist. Road
	Subsistence constraint	HH consumption

Table 3B.6 Descriptive statistics

	Label	Count	Mean	SD	Min	Max
Education	Have you ever attended school?	84,036	0.359	0.479	0	1
Extension	Do you participate in the extension program?	118,841	0.565	0.495	0	1
Land tenure	Does your HH have a certificate for this Parcel?	54,351	0.582	0.493	0	1
Land area	GPS measured area of field (Square Meters)	66,439	1,450	17,932	0	4,257,091
Livestock	Do you own livestock?	85,578	0.421	0.494	0	1
Oxen	How many oxen do you have?	93,912	1.204	1.244	0	25
Off-farm	Has anyone in this HH owned an off-farm business?	59,127	0.317	0.465	0	1
Credit	Do you get credit services?	96,204	0.240	0.427	0	1
HH members	How many HH members work on the farm?	119,875	3.982	0.182	0	4
Dist. road	HH distance to nearest major road (kilometers)	79,360	14.7	14.9	0	241
Irr. scheme	Is there an irrigation scheme in this community?	78,596	0.621	0.485	0	1
Irrigation	Is this field irrigated?	98,510	0.06	0.24	0	1
Elevation	Plot Elevation (meters)	114,370	1,867	477	192	3,419
PWI	Plot Potential Wetness Index (percentage)	114,370	12.6	1.86	0	36
Prices (Birr/kg)						
Cereals		4,325	138.25	297.39	0.5	3,017
Pulses		1,107	177.75	579.23	0.05	10,000
Oilseeds		635	108.49	283.24	0.01	1,950
Vegetables		29	56.7	154.39	0.4	800
Roots		13	115.47	279.81	0.5	1,000
Fruits		818	4.31	3.84	0.2	30
Spices		18	17.73	18.35	3.8	60
Cash crops		1,857	106.64	421.65	0.02	11,250

Note: 1 square meter = 1/10,000 hectares.

Table 3B.7 Crop prices (Birr/kg) comparison

	Sample	FAO	Eth. Stat. Agency
Cereals	138.25	5.99	4.37
Pulses	177.75	9.16	2.86
Oilseeds	108.49	9.95	3.69
Vegetables	56.7	4.72	1.89
Roots	115.37	8.27	1.86
Fruits	4.31	4.32	3.10
Spices	17.73	9.25	1.69
Cash crops	106.64	11.04	5.53

Note: The reported FAO prices are calculated by converting FAO producer prices (FAOSTAT, 2024) in USD/ton for each crop to USD/kg, and multiplying by the average exchange rate of the corresponding year, then averaging across crops within a crop category and across years. Prices are averaged across three years: 2011, 2013, 2015, to match the sample. The corresponding exchange rates from USD to Birr are 16.97, 18.71, and 20.69, respectively. The reported prices from the Ethiopia Central Statistical Agency (CSA, 2011) are calculated by dividing the external trade values of each crop category by the quantity exported.

Table 3B.8 Soil types in the sample

Soil Type	Freq.	Perc.	Cum.Perc.
Leptosol	3,291	11.50	11.50
Cambisol	583	2.04	13.54
Vertisol	10,267	35.87	49.41
Luvisol	9,867	34.46	83.87
Mixed	4,237	14.80	98.67
Other	379	1.32	100

Table 3B.9 Descriptive statistics of variables on Type of Farm - Market oriented or subsistence

	Label	Mean	SD	Min	Max
Sell	Did you sell any of the harvest?	0.17	0.37	0	1
Dist. Market	HH Distance to nearest market (kilometers)	68.28	48.73	1	283
HH consumption	Proportion of crop used for HH consumption				
Cereals		68.94	42.57	0	100
Pulses		65.26	43.61	0	100
Oilseeds		48.92	45.71	0	100
Vegetables		64.25	47.86	0	100
Roots		67.45	46.79	0	100
Fruits		64.48	42.18	0	100
Spices		56.04	48.40	0	100
Cash crops		58.96	47.26	0	100

Table 3B.10 Comparing goodness-of-fit across models using the log likelihood function

	Nested L.	Nested L.	Nested L.	Cond. L	Cond. L	Cond. L
Alternatives	13	13	13	13	16	16
Price	included	excluded	adjusted	included	included	adjusted
Log Likelihood	-25,544.642	-25,563.943	-25,562.69	-27,789.79	-28,015.782	-28,035.299
LR test for IIA	313.07***	313.37***	314.00***			

Table 3B.11 Nested logit model estimation results (variable *price* included), part 1

Variables	Pulses	Oilseeds	Vegetables	Roots	Fruits	Spices	Cash
Education	0.199***	0.004	0.135**	0.137	0.431***	-0.129	0.313***
	(0.068)	(0.109)	(0.066)	(0.092)	(0.056)	(0.126)	(0.051)
Extension	0.095	-0.025	-0.032	0.065	-0.146**	0.015	0.205***
	(0.067)	(0.106)	(0.067)	(0.095)	(0.058)	(0.125)	(0.052)
Land tenure	0.174**	0.621***	0.006	0.134	-0.115**	-0.065	-0.182***
	(0.073)	(0.126)	(0.067)	(0.095)	(0.057)	(0.122)	(0.052)
Land area	-2.71e-09	-2.35e-06	-1.72e-04***	-2.36e-04***	-1.03e-04***	-3.85e-04***	-2.70e-05***
	(1.41e-06)	(3.28e-06)	(1.07e-05)	(1.99e-05)	(6.20e-06)	(3.86e-05)	(2.94e-06)
Livestock	0.003	-0.112	0.148	0.530***	0.034	-0.176	0.135*
	(0.097)	(0.145)	(0.093)	(0.146)	(0.078)	(0.151)	(0.072)
Oxen	0.012	0.062**	-0.022	-0.098**	-0.129***	-0.028	-0.201***
	(0.021)	(0.029)	(0.025)	(0.03)	(0.024)	(0.049)	(0.022)
Off-farm	0.132*	3.76e-04	-0.082	0.413***	0.164***	0.334***	0.172***
	(0.069)	(0.110)	(0.068)	(0.091)	(0.056)	(0.119)	(0.051)
Credit	-0.231***	0.219**	0.087	-0.083	-0.345***	0.173	-0.055
	(0.077)	(0.110)	(0.076)	(0.109)	(0.073)	(0.140)	(0.061)
HH members	-0.702***	-0.600***	-0.536***	-0.582***	-0.163	-0.553***	0.046
	(0.081)	(0.232)	(0.104)	(0.112)	(0.140)	(0.184)	(0.102)
Cambisol	-0.500**	-0.306	-0.293	-1.253***	-0.522**	-0.692	-0.812***
	(0.240)	(0.425)	(0.260)	(0.438)	(0.254)	(0.617)	(0.228)
Vertisol	0.158	0.266	0.589***	-0.059	0.803***	0.617***	0.514***
	(0.109)	(0.201)	(0.115)	(0.146)	(0.095)	(0.215)	(0.082)
Luvisol	0.025	0.793***	0.479***	-0.132	0.853***	0.542**	0.558***
	(0.110)	(0.191)	(0.116)	(0.146)	(0.095)	(0.216)	(0.083)
Soil Mixed	0.155	0.631***	0.644***	0.187	0.731***	0.633***	0.474***
	(0.124)	(0.214)	(0.128)	(0.162)	(0.108)	(0.239)	(0.095)
Soil Other	-0.233	0.871**	0.537**	-0.956**	-0.385	0.090	-0.759**
	(0.243)	(0.341)	(0.245)	(0.421)	(0.353)	(0.555)	(0.297)
Dist. road	0.006***	0.013***	0.004**	0.006**	-0.005***	-0.005	0.002
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)
Irr. scheme	0.149**	0.334***	-0.328***	-0.372***	-0.097	-0.081	-0.115**
	(0.076)	(0.125)	(0.069)	(0.098)	(0.059)	(0.132)	(0.054)
Elevation	0.001***	2.17e-04*	-3.75e-04***	7.41e-04***	-0.002***	-3.69e-04***	-3.72e-04***
	(7.07e-05)	(1.13e-04)	(7.34e-05)	(9.75e-05)	(6.80e-05)	(0.000136)	(5.79e-05)
Water Index	-0.048**	-0.091***	-0.056***	-0.152***	-0.100***	-0.160***	-0.065***
	(0.019)	(0.031)	(0.018)	(0.029)	(0.016)	(0.036)	(0.014)

Note: This table reports the estimated effects of farmer and farm-specific variables on the probability of choosing crop types with irrigation choice nested within crop choice, from the nested logit choice model specification that includes all eight crop groups, with constraints on degenerate nests, and using both household consumption as generic variables. Variable *price* is included in the specification with the coefficient reported in Table 3B.12. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.12 Nested logit model estimation results (variable *price* included), part 2

	Crop-Irrigation	Crop Type
Price	-0.027***	
	(0.004)	
HH consumption	0.087*	
	(0.049)	
Cereals_tau		0.849*
		(0.485)
Pulses_tau		0.882*
		(0.507)
Oilseeds_tau		1
		(constrained)
Vegetables_tau		0.787*
		(0.450)
Roots_tau		1
		(constrained)
Fruits_tau		0.999*
		(0.572)
Spices_tau		1
		(constrained)
Cash_tau		0.473*
		(0.271)

Note: This table reports the estimated effects of the crop-irrigation alternative-specific variables on crop choice, as well as the estimated crop-specific dissimilarity parameters (tau), from the nested logit choice model specification that includes all eight crop groups, with constraints on degenerate nests, and using both crop price and household consumption as generic variables. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.13 Nested logit model estimation results (variable *price* excluded), part 1

Variables	Pulses	Oilseeds	Vegetables	Roots	Fruits	Spices	Cash
Education	0.183***	0.00	0.144**	0.138	0.438***	-0.116	0.315***
	(0.067)	(0.109)	(0.066)	(0.092)	(0.056)	(0.126)	(0.051)
Extension	0.070	-0.015	-0.012	0.068	-0.126**	0.043	0.212***
	(0.067)	(0.106)	(0.067)	(0.094)	(0.058)	(0.126)	(0.052)
Land tenure	0.175**	0.624***	0.006	0.134	-0.114**	-0.066	-0.182***
	(0.073)	(0.126)	(0.067)	(0.095)	(0.056)	(0.122)	(0.051)
Land area	-7.39e-08	-2.26e-06	-1.71e-04***	-2.35e-04***	-1.02e-04***	-3.81e-04***	-2.68e-05***
	(1.43e-06)	(3.26e-06)	(1.06e-05)	(1.99e-05)	(6.17e-06)	(3.85e-05)	(2.94e-06)
Livestock	-0.024	-0.102	0.169*	0.532***	0.052	-0.157	0.142*
	(0.096)	(0.145)	(0.094)	(0.146)	(0.078)	(0.152)	(0.073)
Oxen	0.011	0.062**	-0.021	-0.098**	-0.127***	-0.026	-0.201***
	(0.021)	(0.029)	(0.025)	(0.038)	(0.024)	(0.049)	(0.022)
Off-farm	0.127*	0.001	-0.078	0.413***	0.169***	0.338***	0.173***
	(0.069)	(0.110)	(0.068)	(0.091)	(0.056)	(0.120)	(0.051)
Credit	-0.227***	0.219**	0.086	-0.084	-0.347***	0.170	-0.056
	(0.077)	(0.110)	(0.076)	(0.109)	(0.073)	(0.140)	(0.061)
HH members	-0.867***	-0.359	-0.073	-0.447***	0.629***	0.139	0.272***
	(0.074)	(0.251)	(0.076)	(0.109)	(0.080)	(0.167)	(0.100)
Soil: Cambisol	-0.517**	-0.299	-0.266	-1.250***	-0.498**	-0.650	-0.808***
	(0.239)	(0.425)	(0.260)	(0.438)	(0.254)	(0.618)	(0.228)
Soil: Vertisol	0.133	0.273	0.616***	-0.057	0.823***	0.654***	0.517***
	(0.107)	(0.202)	(0.116)	(0.146)	(0.095)	(0.219)	(0.083)
Soil: Luvisol	-0.003	0.802***	0.508***	-0.129	0.874***	0.582***	0.563***
	(0.108)	(0.192)	(0.117)	(0.146)	(0.096)	(0.220)	(0.083)
Soil: Mixed	0.127	0.641***	0.680***	0.190	0.765***	0.689***	0.481***
	(0.122)	(0.214)	(0.129)	(0.162)	(0.109)	(0.243)	(0.095)
Soil: Other	-0.245	0.875**	0.555**	-0.954**	-0.364	0.128	-0.758**
	(0.242)	(0.341)	(0.245)	(0.420)	(0.352)	(0.556)	(0.297)
Dist. road	0.006***	0.013***	0.004**	0.006**	-0.004***	-0.004	0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.001)
Irr. scheme	0.133*	0.340***	-0.317***	-0.370***	-0.086	-0.065	-0.111**
	(0.075)	(0.125)	(0.069)	(0.098)	(0.059)	(0.132)	(0.054)
Elevation	0.001***	2.32e-04**	-3.28e-04***	7.50e-04***	-0.002***	-3.03e-0.4**	-3.56e-04***
	(7.02e-05)	(1.13e-04)	(7.32e-05)	(9.74e-05)	(6.74e-05)	(1.36e-04)	(5.77e-05)
Water Index	-0.074***	-0.084***	-0.037**	-0.150***	-0.081***	-0.129***	-0.059***
	(0.019)	(0.031)	(0.018)	(0.029)	(0.015)	(0.036)	(0.014)

Note: This table reports the estimated effects of farmer and farm-specific variables on the probability of choosing crop types with irrigation choice nested within crop choice, from the nested logit choice model specification that includes all eight crop groups, with constraints on degenerate nests, and using only household consumption as a generic variable. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3B.14 Nested logit model estimation results (variable *price* excluded), part 2

	Crop-Irrigation	Crop Type
HH consumption	0.105*	
	(0.055)	
Cereals_tau		1.026*
		(0.536)
Pulses_tau		1.065*
		(0.559)
Oilseeds_tau		1
		(constrained)
Vegetables_tau		0.952*
		(0.498)
Roots_tau		1
		(constrained)
Fruits_tau		1.210*
		(0.634)
Spices_tau		1
		(constrained)
Cash_tau		0.572*
		(0.299)

Note: This table reports the estimated effects of the crop-irrigation alternative-specific variable on crop choice, as well as the estimated crop-specific dissimilarity parameters (tau), from the nested logit choice model specification that includes all eight crop groups, with constraints on degenerate nests, and using only household consumption as a generic variable. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.15 Nested logit model estimation results (variable *price* is adjusted), part 1

Variables	Pulses	Oilseeds	Vegetables	Roots	Fruits	Spices	Cash
Education	0.188***	0.00446	0.142**	0.140	0.438***	-0.115	0.317**
	(0.0678)	(0.109)	(0.0663)	(0.0924)	(0.0558)	(0.126)	(0.0508)
Extension	0.0798	-0.0203	-0.0148	0.0732	-0.126**	0.0455	0.215***
	(0.067)	(0.106)	(0.067)	(0.094)	(0.058)	(0.126)	(0.052)
Land tenure	0.175**	0.622***	0.006	0.134	-0.114**	-0.066	-0.182***
	(0.073)	(0.126)	(0.067)	(0.095)	(0.056)	(0.122)	(0.051)
Land area	-4.88e-08	-2.29e-06	-1.71e-04***	-2.35e-04***	-1.02e-04***	-3.81e-04***	-2.68e-05***
	(1.42e-06)	(3.27e-06)	(1.06e-05)	(1.99e-05)	(6.17e-06)	(3.85e-05)	(2.94e-06)
Livestock	-0.014	-0.107	0.166*	0.539***	0.052	-0.156	0.146**
	(0.096)	(0.145)	(0.094)	(0.146)	(0.078)	(0.152)	(0.073)
Oxen	0.012	0.062**	-0.021	-0.098**	-0.127***	-0.026	-0.201***
	(0.021)	(0.029)	(0.025)	(0.038)	(0.024)	(0.049)	(0.022)
Off-farm	0.129*	7.52e-04	-0.079	0.414***	0.169***	0.338***	0.174***
	(0.069)	(0.110)	(0.068)	(0.091)	(0.056)	(0.120)	(0.051)
Credit	-0.228***	0.219**	0.086	-0.084	-0.346***	0.170	-0.056
	(0.077)	(0.110)	(0.076)	(0.109)	(0.073)	(0.140)	(0.061)
HH members	-0.745***	-0.438*	-0.121	-0.335**	0.594***	0.202	0.418***
	(0.107)	(0.237)	(0.081)	(0.131)	(0.082)	(0.170)	(0.136)
Soil: Cambisol	-0.512**	-0.306	-0.271	-1.248***	-0.499**	-0.649	-0.805***
	(0.240)	(0.425)	(0.260)	(0.438)	(0.254)	(0.618)	(0.228)
Soil: Vertisol	0.141	0.265	0.609***	-0.0531	0.822***	0.656***	0.519***
	(0.108)	(0.201)	(0.116)	(0.146)	(0.096)	(0.219)	(0.083)
Soil: Luvisol	-0.006	0.793***	0.502***	-0.124	0.873***	0.583***	0.566***
	(0.108)	(0.192)	(0.117)	(0.146)	(0.096)	(0.220)	(0.083)
Soil: Mixed	0.137	0.632***	0.673***	0.196	0.764***	0.691***	0.485***
	(0.122)	(0.214)	(0.129)	(0.162)	(0.109)	(0.243)	(0.095)
Soil: Other	-0.243	0.868**	0.550**	-0.951**	-0.366	0.129	-0.756**
	(0.242)	(0.341)	(0.245)	(0.420)	(0.352)	(0.556)	(0.297)
Dist. road	0.006***	0.014***	0.004**	0.006**	-0.005***	-0.005	0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.001)
Irr. scheme	0.139*	0.337***	-0.318***	-0.367***	-0.086	-0.065	-0.108**
	(0.076)	(0.125)	(0.069)	(0.098)	(0.059)	(0.132)	(0.054)
Elevation	0.001***	2.32e-04**	-3.28e-04***	7.50e-04***	-0.002***	-3.03e-0.4**	-3.56e-04***
	(7.02e-05)	(1.13e-04)	(7.32e-05)	(9.74e-05)	(6.74e-05)	(1.36e-04)	(5.77e-05)
Water Index	-0.064***	-0.089***	-0.040**	-0.144***	-0.081***	-0.127***	-0.056***
	(0.019)	(0.031)	(0.018)	(0.029)	(0.015)	(0.037)	(0.014)

Note: This table reports the estimated effects of farmer and farm-specific variables on the probability of choosing crop types with irrigation choice nested within crop choice, from the nested logit choice model specification that includes all eight crop groups, with constraints on degenerate nests, and using both adjusted crop prices and household consumption as generic variables. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.16 Nested logit model estimation results (variable *price* is adjusted), part 2

	Crop-Irrigation
Price_adjusted	0.014***
	(9.88e-05)
HH consumption	0.092***
	(5.98e-04)

Note: This table reports the estimated effects of the crop-irrigation alternative-specific variables on crop choice, from the nested logit choice model specification that includes all eight crop groups, with constraints on degenerate nests, using both adjusted prices and household consumption as generic variables. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, *$p<0.1$.

Table 3B.17 Conditional Logit model estimation results of farm-specific variables for cereals, pulses, oilseeds, and vegetables (13 alternatives)

Variables	Cereals R-fed	Pulses Irr.	Pulses R-fed	Oilseeds R-fed	Vegetables Irr.	Vegetables R-fed
Education	-0.544***	0.030	-0.359*	-0.586***	-0.258	-0.394**
	(0.186)	(0.742)	(0.193)	(0.211)	(0.309)	(0.192)
Extension	0.625***	-0.332	0.724***	0.607***	0.101	0.636***
	(0.199)	(0.795)	(0.206)	(0.222)	(0.326)	(0.206)
Land tenure	0.528***	-0.080	0.705***	1.119***	0.703**	0.485**
	(0.194)	(0.734)	(0.202)	(0.226)	(0.319)	(0.200)
Land area	4.78e-05***	2.63e-05	4.71e-05***	4.54e-05***	-5.01e-04***	-1.26e-04***
	(1.23e-05)	(4.84e-05)	(1.24e-05)	(1.27e-05)	(1.05e-04)	(1.60e-05)
Livestock	-0.787**	-1.819**	-0.783**	-0.905***	-0.960**	-0.579*
	(0.313)	(0.796)	(0.323)	(0.339)	(0.445)	(0.321)
Oxen	0.417***	-0.314	0.425***	0.482***	0.494***	0.373***
	(0.102)	(0.447)	(0.103)	(0.105)	(0.139)	(0.104)
Off-farm	-0.454**	-1.661	-0.277	-0.446**	0.0636	-0.562***
	(0.183)	(1.086)	(0.191)	(0.209)	(0.302)	(0.191)
Credit	0.245	-0.733	0.003	0.458*	-0.0752	0.346
	(0.250)	(1.105)	(0.257)	(0.269)	(0.402)	(0.256)
HH members	-0.348	-1.127	-1.010***	-0.888***	-1.206***	-0.838**
	(0.359)	(0.961)	(0.361)	(0.304)	(0.357)	(0.372)
Soil: Cambisol	0.518	-14.79	0.078	0.221	0.746	0.158
	(0.549)	(2.882)	(0.587)	(0.685)	(0.850)	(0.598)
Soil: Vertisol	0.638***	0.105	0.850***	0.925***	0.534	1.305***
	(0.231)	(0.971)	(0.248)	(0.300)	(0.434)	(0.250)
Soil: Luvisol	0.882***	0.645	0.919***	1.678***	0.666	1.423***
	(0.266)	(0.936)	(0.281)	(0.321)	(0.463)	(0.282)
Soil: Mixed	0.699**	-13.470	0.926***	1.358***	0.927*	1.420***
	(0.299)	(822.8)	(0.315)	(0.359)	(0.502)	(0.316)
Soil: Other	-0.409	-15.170	-0.655	0.562	-15.26	0.173
	(0.718)	(3,556)	(0.747)	(0.779)	(1,593)	(0.742)
Dist. road	0.004	0.016	0.010*	0.018***	0.025***	0.006
	(0.005)	(0.018)	(0.006)	(0.006)	(0.006)	(0.005)
Irr. scheme	-3.268***	0.296	-3.129***	-2.916***	-2.461***	-3.611***
	(0.577)	(3.056)	(0.580)	(0.588)	(0.661)	(0.579)
Elevation	0.003***	0.003***	0.004***	0.003***	0.001***	0.002***
	(2.42e-04)	(8.10e-04)	(2.48e-04)	(2.63e-04)	(3.60e-04)	(2.48e-04)
Water Index	-0.114***	0.025	-0.163***	-0.188***	0.059	-0.195***
	(0.038)	(0.167)	(0.041)	(0.047)	(0.054)	(0.041)

Note: This table reports the estimated marginal effects of farmer and farm-specific variables on the probability of choosing crop-irrigation alternatives, from the conditional logit choice model specification that includes only 13 choice combinations. Both crop price and household consumption are included as generic variables. The results are split to fit the page. This table reports the results for the crop types cereals, pulses, oilseeds, and vegetables. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.18 Conditional Logit model estimation results of farm-specific variables for roots, fruits, spices, and cash crops (13 alternatives)

Variables	Roots R-fed	Fruits Irr.	Fruits R-fed	Spices R-fed	Cash Irr.	Cash R-fed
Education	-0.380*	-0.486**	-0.100	-0.678***	-0.348	-0.189
	(0.200)	(0.235)	(0.188)	(0.217)	(0.230)	(0.187)
Extension	0.629***	0.127	0.492**	0.627***	0.153	0.878***
	(0.214)	(0.245)	(0.202)	(0.229)	(0.242)	(0.201)
Land tenure	0.634***	0.362	0.379*	0.489**	0.227	0.325*
	(0.209)	(0.238)	(0.196)	(0.222)	(0.234)	(0.195)
Land area	-1.83e-04***	-7.66e-05***	-4.37e-05***	-3.62e-04***	-7.56e-05***	2.05e-05
	(2.17e-05)	(2.31e-05)	(1.33e-05)	(4.06e-05)	(2.29e-05)	(1.25e-05)
Livestock	-0.113	-0.451	-0.700**	-1.008***	-0.943***	-0.606*
	(0.339)	(0.371)	(0.315)	(0.338)	(0.352)	(0.315)
Oxen	0.294***	0.383***	0.246**	0.368***	0.282**	0.199*
	(0.107)	(0.120)	(0.103)	(0.112)	(0.120)	(0.103)
Off-farm	-0.042	-0.150	-0.232	-0.0730	-0.226	-0.248
	(0.198)	(0.228)	(0.186)	(0.211)	(0.225)	(0.185)
Credit	0.170	0.041	-0.131	0.375	0.346	0.122
	(0.266)	(0.305)	(0.255)	(0.280)	(0.291)	(0.253)
HH members	-0.868**	-1.210***	-0.410	-0.809**	-0.846***	-0.145
	(0.377)	(0.333)	(0.399)	(0.373)	(0.315)	(0.331)
Soil: Cambisol	-0.652	-0.315	-0.0752	0.0320	-0.505	-0.194
	(0.673)	(0.747)	(0.596)	(0.760)	(0.746)	(0.582)
Soil: Vertisol	0.669**	0.425	1.646***	1.240***	0.273	1.404***
	(0.262)	(0.285)	(0.240)	(0.302)	(0.285)	(0.237)
Soil: Luvisol	0.853***	0.408	1.919***	1.395***	0.524*	1.642***
	(0.293)	(0.324)	(0.274)	(0.330)	(0.316)	(0.271)
Soil: Mixed	0.915***	-0.253	1.658***	1.245***	0.0214	1.413***
	(0.329)	(0.396)	(0.307)	(0.370)	(0.377)	(0.304)
Soil: Other	-1.456*	-1.257	-0.657	-0.422	-15.10	-1.094
	(0.820)	(1.200)	(0.773)	(0.891)	(907.1)	(0.761)
Dist. road	0.010*	0.005	-0.002	-0.002	0.005	0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)
Irr. scheme	-3.722***	-0.133	-3.436***	-3.356***	-0.883	-3.510***
	(0.582)	(0.728)	(0.577)	(0.589)	(0.662)	(0.577)
Elevation	0.003***	3.74e-04	0.001***	0.002***	0.001***	0.002***
	(2.55e-04)	(2.89e-04)	(2.44e-04)	(2.71e-04)	(2.86e-04)	(2.43e-04)
Water Index	-0.273***	-0.009	-0.252***	-0.270***	0.046	-0.218***
	(0.046)	(0.045)	(0.039)	(0.051)	(0.044)	(0.039)

Note: This table reports the estimated marginal effects of farmer and farm-specific variables on the probability of choosing crop-irrigation alternatives, from the conditional logit choice model specification that includes only 13 choice combinations. Both crop price and household consumption are included as generic variables. The results are split to fit the page. This table reports the results for the crop types roots, fruits, spices, and cash crops. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3B.19 Conditional Logit model estimation results of crop-specific variables (13 alternatives)

	Crop-Irrigation
Price	-0.028***
	(0.004)
HH consumption	0.107***
	(0.032)

Note: This table reports the estimated marginal effects of the alternative-specific variables on the probability of crop-irrigation outcomes, from the conditional logit choice model specification that includes only 13 alternatives. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.20 Conditional Logit model estimation results of farm-specific variables for cereals, pulses, oilseeds, and vegetables (16 alternatives)

Variables	Cereals R-fed	Pulses Irr.	Pulses R-fed	Oilseeds Irr.	Oilseeds R-fed	Vegetables Irr.	Vegetables R-fed
Education	-0.539***	0.040	-0.352*	5.272	-0.579***	-0.216	-0.389**
	(0.185)	(0.742)	(0.193)	(3.867)	(0.211)	(0.311)	(0.192)
Extension	0.618***	-0.324	0.717***	-5.404	0.603***	0.127	0.629***
	(0.199)	(0.797)	(0.206)	(4.081)	(0.222)	(0.329)	(0.206)
Land tenure	0.523***	-0.0861	0.699***	6.215	1.115***	0.678**	0.480**
	(0.194)	(0.735)	(0.202)	(5.498)	(0.226)	(0.322)	(0.200)
Land area	4.76e-05***	2.64e-05	4.70e-05***	-0.001	4.53e-05***	-5.09e-04***	-1.26e-04***
	(1.23e-05)	(4.83e-05)	(1.24e-05)	(0.001)	(1.27e-05)	(1.08e-04)	(1.60e-05)
Livestock	-0.780**	-1.809**	-0.776**	-1.697	-0.895***	-0.981**	-0.572*
	(0.313)	(0.797)	(0.323)	(2.484)	(0.339)	(0.447)	(0.321)
Oxen	0.414***	-0.316	0.422***	-15.01	0.479***	0.489***	0.370***
	(0.102)	(0.447)	(0.103)	(608.0)	(0.105)	(0.140)	(0.104)
Off-farm	-0.463**	-1.668	-0.288	-15.42	-0.455**	0.102	-0.572***
	(0.183)	(1.086)	(0.191)	(1.633)	(0.209)	(0.304)	(0.191)
Credit	0.242	-0.737	1.83e-04	-15.44	0.454*	-0.012	0.344
	(0.250)	(1.105)	(0.257)	(1.207)	(0.269)	(0.403)	(0.256)
HH members	-0.510	-1.106	-1.167***	-2.674	-0.965***	-1.221***	-1.006***
	(0.359)	(1.041)	(0.361)	(2.750)	(0.305)	(0.362)	(0.371)
Soil: Cambisol	0.504	-18.78	0.063	-11.36	0.211	0.738	0.143
	(0.549)	(20.748)	(0.587)	(56.504)	(0.685)	(0.851)	(0.598)
Soil: Vertisol	0.636***	0.103	0.847***	2.508	0.928***	0.549	1.302***
	(0.231)	(0.972)	(0.248)	(3.651)	(0.300)	(0.436)	(0.250)
Soil: Luvisol	0.883***	0.645	0.919***	-16.13	1.684***	0.602	1.422***
	(0.266)	(0.939)	(0.281)	(2.770)	(0.322)	(0.471)	(0.282)
Soil: Mixed	0.705**	-17.45	0.934***	-13.05	1.373***	0.963*	1.427***
	(0.299)	(6.061)	(0.315)	(1.477)	(0.360)	(0.504)	(0.316)
Soil: Other	-0.300	-19.07	-0.541	-18.44	0.677	-19.15	0.280
	(0.714)	(26.349)	(0.743)	(33.084)	(0.775)	(11.997)	(0.738)
Dist. road	0.004	0.016	0.010*	-0.398	0.018***	0.025***	0.006
	(0.005)	(0.019)	(0.006)	(0.384)	(0.005)	(0.006)	(0.005)
Irr. scheme	-3.338***	0.439	-3.198***	0.711	-2.983***	-2.531***	-3.680***
	(0.584)	(3.385)	(0.587)	(4.069)	(0.595)	(0.669)	(0.586)
Elevation	0.003***	0.003***	0.004***	9.70e-04	0.003***	0.001***	0.002***
	(2.41e-04)	(8.14e-04)	(2.48e-04)	(0.00426)	(2.62e-04)	(3.63e-04)	(2.47e-04)
Water Index	-0.117***	0.032	-0.166***	0.274	-0.187***	0.063	-0.199***
	(0.038)	(0.169)	(0.041)	(0.591)	(0.047)	(0.055)	(0.041)

Note: This table reports the estimated marginal effects of farmer and farm-specific variables on the probability of choosing crop-irrigation alternatives, from the conditional logit choice model specification that includes all 16 choice combinations. Both crop price and household consumption are included as generic variables. The results are split to fit the page. This table reports the results for the crop types cereals, pulses, oilseeds, and vegetables. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.21 Conditional Logit model estimation results of farm-specific variables for roots, fruits, spices, and cash crops (16 alternatives)

Variables	Roots Irr.	Roots R-fed	Fruits Irr.	Fruits R-fed	Spices Irr.	Spices R-fed	Cash Irr.	Cash R-fed
Education	-0.643	-0.375*	-0.483**	-0.098	-18.70	-0.672***	-0.339	-0.184
	(0.453)	(0.200)	(0.235)	(0.187)	(4.855)	(0.217)	(0.230)	(0.187)
Extension	-1.270**	0.622***	0.123	0.486**	0.449	0.621***	0.159	0.872***
	(0.556)	(0.214)	(0.245)	(0.202)	(0.929)	(0.229)	(0.242)	(0.201)
Land tenure	0.431	0.629***	0.361	0.375*	2.773**	0.484**	0.214	0.321
	(0.440)	(0.209)	(0.238)	(0.196)	(1.317)	(0.222)	(0.234)	(0.196)
Land area	-9.66e-05	-1.83e-04***	-7.66e-05***	-4.39e-05***	-0.002**	-3.62e-04***	-7.61e-05***	2.03e-05
	(5.94e-05)	(2.17e-05)	(2.31e-05)	(1.33e-05)	(8.91e-04)	(4.06e-05)	(2.29e-05)	(1.25e-05)
Livestock	-0.633	-0.105	-0.440	-0.694**	-1.739**	-1.000***	-0.940***	-0.598*
	(0.661)	(0.339)	(0.371)	(0.315)	(0.880)	(0.338)	(0.352)	(0.316)
Oxen	0.360*	0.290***	0.380***	0.244**	-0.088	0.365***	0.284**	0.197*
	(0.188)	(0.107)	(0.120)	(0.103)	(0.457)	(0.112)	(0.120)	(0.103)
Off-farm	0.087	-0.052	-0.154	-0.239	-0.052	-0.083	-0.227	-0.257
	(0.416)	(0.198)	(0.228)	(0.186)	(0.845)	(0.211)	(0.225)	(0.185)
Credit	-0.699	0.167	0.037	-0.134	0.322	0.373	0.350	0.119
	(0.604)	(0.265)	(0.304)	(0.254)	(0.976)	(0.280)	(0.291)	(0.252)
HH members	-1.437***	-1.035***	-1.290***	-0.595	-1.640*	-0.955**	-0.937***	-0.275
	(0.421)	(0.376)	(0.335)	(0.395)	(0.903)	(0.372)	(0.317)	(0.331)
Cambisol	1.334	-0.669	-0.324	-0.079	5.333*	0.014	-0.501	-0.203
	(1.023)	(0.673)	(0.747)	(0.596)	(3.048)	(0.760)	(0.747)	(0.581)
Vertisol	0.363	0.666**	0.421	1.646***	-16.34	1.239***	0.287	1.403***
	(0.630)	(0.262)	(0.285)	(0.240)	(6.717)	(0.302)	(0.285)	(0.237)
Luvisol	0.341	0.851***	0.409	1.922***	3.275	1.396***	0.541*	1.644***
	(0.664)	(0.293)	(0.325)	(0.274)	(2.850)	(0.330)	(0.317)	(0.271)
Soil Mixed	1.109*	0.921***	-0.243	1.666***	3.761	1.255***	0.043	1.421***
	(0.655)	(0.329)	(0.397)	(0.307)	(2.846)	(0.370)	(0.378)	(0.305)
Soil Other	-20.44	-1.345*	-1.219	-0.564	5.102	-0.308	-19.01	-0.988
	(13.812)	(0.817)	(1.198)	(0.768)	(3.228)	(0.888)	(6.685)	(0.757)
Dist. road	0.011	0.009*	0.004	-0.002	-0.195**	-0.001	0.004	0.004
	(0.0136)	(0.005)	(0.006)	(0.005)	(0.088)	(0.006)	(0.006)	(0.005)
Irr. scheme	-2.976***	-3.792***	-0.160	-3.504***	-3.713***	-3.425***	-0.938	-3.578***
	(0.756)	(0.589)	(0.739)	(0.585)	(1.137)	(0.596)	(0.670)	(0.585)
Elevation	0.002***	0.003***	3.67e-04	0.001***	0.002**	0.002***	0.001***	0.002***
	(4.82e-04)	(2.54e-04)	2.89e-04)	2.43e-04)	8.60e-04)	(2.70e-04)	(2.86e-04)	(2.43e-04)
Water Index	0.154**	-0.276***	-0.007	-0.258***	0.041	-0.272***	0.047	-0.221***
	(0.059)	(0.046)	(0.045)	(0.039)	(0.164)	(0.051)	(0.044)	(0.039)

Note: This table reports the estimated marginal effects of farmer and farm-specific variables on the probability of choosing crop-irrigation alternatives, from the conditional logit choice model specification that includes all 16 choice combinations. Both crop price and household consumption are included as generic variables. The results are split to fit the page. This table reports the results for the crop types roots, fruits, spices, and cash crops. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3B.22 Conditional Logit model estimation results of crop-specific variables (16 alternatives)

	Crop-Irrigation
Price	-0.028***
	(0.004)
HH consumption	0.129***
	(0.031)

Note: This table reports the estimated marginal effects of the crop-irrigation alternative-specific variables on crop choice, from the conditional logit choice model specification that includes all 16 alternatives. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3B.23 Irrigation scheme summary by crop-irrigation alternative

Irrigation scheme	Obs	Mean	SD
Cereals Irrigated	444	0.964	0.186
Cereals Rainfed	15,559	0.714	0.451
Pulses Irrigated	40	0.95	0.221
Pulses Rainfed	4,379	0.721	0.448
Oilseeds Irrigated	4	1	0
Oilseeds Rainfed	1,304	0.778	0.415
Vegetables Irrigated	150	0.9	0.301
Vegetables Rainfed	3,253	0.62	0.485
Roots Irrigated	110	0.918	0.275
Roots Rainfed	1,888	0.606	0.488
Fruits Irrigated	434	0.97	0.17
Fruits Rainfed	4,869	0.606	0.488
Spices Irrigated	12	0.75	0.452
Spices Rainfed	734	0.67	0.47
Cash Irrigated	478	0.968	0.174
Cash Rainfed	7,064	0.608	0.325

Note: This table reports the summary statistics for the variable *Irr. scheme* by crop-irrigation alternative, where all 16 alternatives are included. *Irr. scheme* is a binary variable that takes on the value 1 if the respondent answers "yes" to the question "Is there an irrigation scheme in this community?", and 0 otherwise.

Table 3B.24 Predictive margins of crop-irrigation alternatives

	Margin
Cereals Irrigated	0.009***
	(7.09e-04)
Cereals Rainfed	0.355***
	(0.004)
Pulses Irrigated	5.65e-04**
	(1.87e-04)
Pulses Rainfed	0.088***
	(0.002)
Oilseeds Irrigated	1.30e-04
	(8.61e-05)
Oilseeds Rainfed	0.029***
	(0.001)
Vegetables Irrigated	0.004***
	(5.26e-04)
Vegetables Rainfed	0.089***
	(0.002)
Roots Irrigated	0.002***
	(3.40e-04)
Roots Rainfed	0.045***
	(0.002)
Fruits Irrigated	0.014***
	(8.97e-04)
Fruits Rainfed	0.144***
	(0.003)
Spices Irrigated	5.04e-04**
	(1.74e-04)
Spices Rainfed	0.023***
	(0.001)
Cash Irrigated	0.015***
	(9.44e-04)
Cash Rainfed	0.181***
	(0.003)

Note: This table reports the predictive margins of the crop-irrigation alternatives, from the conditional logit choice model specification that includes all 16 alternatives. Rows represent the probability of choosing an alternative among the choice set. Delta-method standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.