ABSTRACT

Title of Dissertation: ESSAYS ON CLIMATE ADAPTATION AND ENVIRONMENTAL VALUATION

Haoluan Wang
Doctor of Philosophy, 2022

Dissertation directed by: Professor Erik Lichtenberg
Department of Agricultural and Resource Economics

Associate Professor David A. Newburn
Department of Agricultural and Resource Economics

Risk is an important component of the decision-making process. Often time, risk assessment is associated with either space or time. How agents perceive risk and how they respond to risk can have significant policy implications, especially when government programs are designed to either incentivize the provision of environmental amenities or reduce the production of environmental disamenities. This dissertation features three chapters that examine the role of risk, time, and space in evaluating environmental disamenities and amenities in the context of climate adaptation and ecosystem goods and services.

The first chapter studies the spillover effects of levee building in response to rising flood risks in the U.S. Mississippi. Using newly digitized data on levee locations and elevations with the Great Mississippi Flood of 2011 as a natural experiment, I show that a 1% increase in the upstream levee elevation increased the downstream levee elevation by 0.7%. A back-of-the-envelope

calculation suggests the external costs due to upstream levee building are at least $0.2 billion, reducing the net benefits of heightened levees by 55%. The results highlight the importance of regional coordination to manage large-scale natural disasters while mitigating inter-jurisdictional spillovers.

The second chapter uses a discrete choice experiment implemented in a farmer survey to elicit landowners' willingness to enroll in long-term payments for ecosystem services programs in Maryland. We address the issue of serial non-participation (SNP) when landowners always choose the status quo option and examine the role of time and risk preferences in landowners' enrollment decisions. We find that ceiling on program participation is evident when SNP is accounted for, pointing to an inherent limitation in voluntary programs. Failing to account for SNP can also lead to quantitatively different welfare measures. Landowners are responsive to program payments with low discount rates consistent with market interest rates. Risk-averse landowners are less likely to enroll in programs, suggesting that they perceive participation to increase income risk.

The third chapter proposes a novel extension of existing semi-parametric approaches to examine spatial patterns of willingness to pay (WTP) and status quo effects, including tests for global spatial autocorrelation, spatial interpolation techniques, and local hotspot analysis. We incorporate the statistical precision of WTP values into the spatial analyses using a two-step methodology and demonstrate this method using data from a stated preference survey that elicited values for improvements in water quality in the Chesapeake Bay and lakes in the surrounding watershed. Our proposed methodology offers a flexible way to identify potential spatial patterns of welfare impacts and facilitates more accurate benefit-cost and distributional analyses.

ESSAYS ON CLIMATE ADAPTATION AND ENVIRONMENTAL VALUATION

by

Haoluan Wang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Erik Lichtenberg, Co-Chair
Associate Professor David A. Newburn, Co-Chair
Professor Anna Alberini
Professor Lori Lynch
Professor Maureen Cropper

Foreword

The second chapter is jointly authored with Erik Lichtenberg and David A. Newburn. The third chapter is jointly authored with Dennis Guignet and Chris Moore. The Dissertation Committee acknowledges that Haoluan Wang made substantial contributions to relevant aspects of both chapters.

Dedication

To what I have received and to what I am still receiving.

Acknowledgements

The level of support I received during my study at the University of Maryland, both professionally and personally, can never be overstated. First and foremost, I am indebted to my advisors, Erik Lichtenberg and David A. Newburn, who have provided me with unparallel mentorship and research opportunities throughout my Ph.D. journey. They have demonstrated unbounded compassion and dedication towards their mentees, from which I have greatly benefited and will keep benefiting for a very long time. I especially appreciate their constant availability for two-way conversations and their ability to find a way forward. Having the chance of working with them has inspired me to be a better scholar and their superior guidance will continue to be a source of enlightenment in my professional life.

I am grateful to Anna Alberini for providing countless suggestions on multiple research projects I have worked on that helped shape the course of my dissertation. She has always been willing and eager to offer help and her positive energy has always been an inspiration to me. I would also like to thank Lori Lynch and Maureen Cropper for their willingness to serve on my committee and for providing very helpful feedback.

I also want to express appreciation for my fellow graduate students in the Department of Agricultural and Resource Economics for their friendship. Many thanks are given to the AREC staff who have always been there to help me navigate, including Pin-Wuan, Dany, Jeff, John, and Katherine.

Completing a Ph.D. degree is not cost-free, and I am lucky to have received financial support from multiple sources. Thanks to the Department of Agricultural and Resource Economics for the graduate assistantship, especially to Lori Lynch for

providing additional support that helped many of us survive the pandemic. I am also grateful for the teaching assistantship from the Department of Geographical Sciences in my final year. I also want to personally thank David A. Newburn, Robert Johnston, Colin Polsky, and Erik Lichtenberg for generously providing additional financial support through the research opportunities that I very much enjoyed working on.

Six years of living in College Park has been a slice and I am truly fortunate to have roommates with whom I have been able to share many beautiful moments of my life. Thanks to Joey, Greg, Abby, Jack, Jake, Halvor, Anouk, Nika, Jacques, Evan, Mario, and Mike. Finally, I have the privilege of dog sitting you for the past three years, Otis. You are always by my side with a goofy look and are an expert in keeping me sane when research does not always work out smoothly.

Table of Contents

List of Tables

List of Figures

Introduction

Risk is an important component of the decision-making process. Often time, risk assessment is associated with either space or time. For example, inter-temporal decision-making is inherently risky, and the role of both risk and time preferences has been long acknowledged in contracting (Stiglitz, 1974; Allen and Lueck, 1995), land rental markets (Fischer and Wollni, 2018; Lloyd-Smith et al., 2021), and new technology adoption (Brick and Visser, 2015; Michler and Wu, 2020). In the meantime, risk spillover across space and its consequences have been increasingly examined in the context of environmental disamenities such as water and air pollution (Sigman, 2005; Cai et al., 2015; Lipscomb and Mobarak, 2017; Yang et al., 2017).

How agents perceive risk and how they respond to risk can have significant policy implications, especially when government programs are designed to either incentivize the provision of environmental amenities or reduce the production of environmental disamenities. When implementing such government programs, policymakers often face multiple challenges to achieve cost-effectiveness. For instance, risk aversion has been found to be a barrier to participation in government subsidy programs that incentivize the adoption of conservation practices (Canales et al., 2015; Ramsey et al., 2019). Not accounting for risk can substantially overestimate the effect of monetary incentives on program uptake. As another example, unintended consequences may occur when actions taken for self-protection in response to risk can induce external risks and costs to other agents (Anderson and Auffhammer, 2014; Lakdawalla and Zanjani, 2005). As a result, external costs of self-protective actions need to be measured and considered when evaluating the viability of these actions.

This dissertation features three chapters examining the role of risk, time, and space in evaluating environmental disamenities and amenities in the context of climate adaptation and

ecosystem goods and services. Specifically, I document how these aspects influence the evaluation of environmental disamenities and amenities and estimate the magnitudes of these effects. The results highlight the importance of taking these aspects into account when designing and implementing government programs to achieve cost-effectiveness.

The first chapter studies the spillover effects of levee building in response to rising flood risks in the U.S. Mississippi. Without clear compensation for potential damage from extra risk exposure, self-protective actions taken by some agents may force other agents to increase their own protective actions through risk spillovers, a scenario often referred to as arms race. Levee building along the Mississippi River appears to be a case in point. A levee is a man-made structure, usually an earthen embankment, that is commonly constructed along the Mississippi River to divert water flow and reduce the risk of temporary flooding. Despite the purpose of flood protection, growing concerns have been raised that levees heightened in some areas in response to catastrophic flood events increased flood risks downstream, causing downstream to build higher levees.

I provide both theoretical and empirical evidence of the relationship between levee building and flood risks, with a particular focus on spatial spillovers from upstream levee construction. Using the Great Mississippi Flood of 2011 as a natural experiment and applying a difference-in-differences approach to newly digitized data on levee locations and elevations, I establish three main results. First, I find that levee building in the flooded counties increased by over 2 ft after the 2011 flood. Second, I show that a 1% increase in the upstream levee elevation increased the downstream levee elevation by 0.7%. A back-of-the-envelope calculation indicates that spatial spillovers from upstream levee building can amount to at least $0.2 billion external costs for downstream jurisdictions. Third, a cost-benefit analysis further shows that considering the external costs of spillovers due to upstream levee building, the net benefits of heightened levees can be reduced by 55%.

Given the substantial spatial spillovers from upstream levee building, this study highlights the importance of inter-jurisdictional coordination and call for a more centralized regulation in levee construction for flood control to reduce negative externalities from heightened levees. I discuss alternative policy recommendations, including spatially explicit compensation schemes in terms of financial transfers across jurisdictions to correct the negative externalities. Future flood-control projects, or management of large-scale natural disasters in general, that require an evaluation of cost-effectiveness must take into account the potential external costs so that the implementation of those projects can be economically justified.

The second chapter studies the role of risk and time in landowners' willingness to enroll in long-term payments for ecosystem services (PES) programs. PES programs have been widely used in both developing and developed countries to address negative environmental externalities from agriculture (Wunder et al, 2008; Alix-Garcia and Wolff, 2014; Ribaudo and Shortle, 2019). In most settings, willing landowners voluntarily enroll in PES programs, through which they receive a variety of payments (e.g., upfront signing bonus, recurring annual payment) in exchange for practice adoption on farmland or changes in land use. Since policymakers are looking at expanding enrollment in PES programs, stated preference (SP) surveys offer a way to design future programs and have been increasingly used to evaluate landowners' willingness to accept (WTA) for program participation (see Villanueva et al., 2017 for a literature review).

We use a discrete choice experiment implemented in a farmer survey in Maryland to elicit landowners' WTA for government payment programs that incentivize the adoption of riparian buffers on working farmland. We particularly address the issue of serial non-participation (SNP) when landowners always choose the status quo option using single- and double-hurdle variants of the discrete choice model. The key distinction between a standard logit and the two hurdle models is the extent to which the preferences of serial non-participants

can be extrapolated from the preferences of participants who chose at least one proposed program. The standard logit model assumes that serial non-participants and participants share the same utility functions so that the preferences of serial non-participants can be linearly extrapolated from the observed preferences of participants. Allowing serial non-participants' preferences to be fundamentally different from those of participants, the single-hurdle logit model assumes that no such extrapolation is possible while the double-hurdle logit model assumes that some extrapolation is possible.

We establish several important results in this paper. First, whether SNP is accounted for in the econometric model is crucial for the empirical analysis. Our post-estimation calculations of the probability of enrollment, average WTA for program participation, and implicit discount rates indicate that different assumptions about serial non-participants embedded in alternative models can make a significant difference in inferences about policy. Our simulated probabilities of enrollment show that ceiling on PES program participation becomes evident when SNP is explicitly modeled using a hurdle framework that allows serial non-participants to behave differently from participants. This result points to an inherent limitation in most voluntary programs that aim to provide ecosystem goods and services, suggesting that simply making programs more lucrative may not raise that ceiling very much. Failing to account for SNP can also lead to quantitatively different WTA estimates for program participation. Second, we find that risk-averse landowners are less likely to enroll in proposed programs, indicating that they perceive program participation to increase income risk. This result suggests that guaranteed government payments may not offset risks and uncertainty to farm production from potential changes in existing farm operations due to riparian buffer adoption. Third, our estimates of landowners' implicit discount rates are quite low with weak evidence of hyperbolic discounting. The estimated supply-side discount rates are substantially

lower than those estimated for the demand of environmental amenities but are consistent with market interest rates such as loans offered by the U.S. Department of Agriculture.

The third chapter proposes a novel extension of existing semi-parametric approaches that consider the statistical precision to examine spatial patterns of willingness to pay (WTP) and status quo (SQ) effects. A better understanding of the spatial distribution of welfare impacts is necessary for conducting accurate benefit-cost and distributional analyses, both in terms of defining the appropriate extent of the market and in interpolating values within that market. To examine spatial heterogeneity in WTP, most SP studies have applied the distance decay paradigm, where WTP is hypothesized to diminish with distance from the resource (Bateman et al., 2006). While there is a growing number of SP studies using traditional econometric methods to control for "distance decay" or other forms of spatial heterogeneity (e.g., Hanley et al., 2003; Rolfe and Windle, 2012; Olsen et al., 2020), these parametric approaches can sometimes fail to identify existing spatial patterns.

Alternatively, some recent studies have employed local indicators of spatial association (LISAs) or hotspot analysis to identify non-continuous, local clusters of systematically higher or lower WTP values (i.e., hot and cold spots, respectively) (Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015). However, WTP values are estimates and not observed values. No study to date has formally accounted for the statistical precision of those estimates when conducting spatial analyses, despite its potential importance in drawing policy implications. We therefore borrow techniques from meta-analytic methods and formally account for the statistical precision of WTP estimates by incorporating the variance of these estimates into the subsequent spatial analyses, including tests for global spatial autocorrelation, spatial interpolation techniques, and local hotspot analysis.

We apply our proposed two-step methodology using data from a SP survey that elicited values for improvements in water quality in the Chesapeake Bay and lakes in the surrounding

watershed. Our semi-parametric results of the spatial interpolation suggest distinct local patterns in marginal willingness to pay (MWTP) estimates for all program attributes, and evident spatial heterogeneity across the study area. The hotspot analysis confirms statistically significant spatial clusters of high and low MWTP values. Comparison of the conventional spatial analyses to our variance-adjusted results that take into account the statistical precision reveals some differences. In general, accounting for the variances of the MWTP estimates diminishes spatial variation, suggesting that not accounting for the statistical precision of the first-stage MWTP estimates could lead analysts to falsely identify patterns of global and local spatial heterogeneity in the second stage. Our analysis of spatial variation of individual-specific SQ effects reveals substantial differences when accounting for the statistical precision of the estimates. Particularly, our proposed variance-adjustment leads to an increased identification of clusters of individuals exhibiting "warm glow" or other biases for a policy option. Lastly, although differences in local patterns are revealed, our policy simulations suggest that accounting for local spatial heterogeneity (with or without our variance-adjusted extension) may not yield substantial differences in terms of broader welfare implications, at least not in this specific application of water quality and ecosystem improvements in an iconic and well-known estuary.

Chapter 1: Flood Your Neighbors: Spillover Effects of Levee Building

1.1 Introduction

Economists have long acknowledged the problem of social cost (Coase, 1960) and the role of property rights in addressing externalities (Demsetz, 1967). However, when property rights are not well defined or Coasean bargaining is too costly, the level of externality can be higher than socially desired. One such example is measures taken to protect against natural catastrophes such as floods, hurricanes, or wildfires. These measures have been widely recognized to increase the risk of damage by incentivizing agents to increase economic activities in catastrophe-prone areas (Shafran, 2008; Pinter et al., 2016). Without clear compensation for potential damage from extra risk exposure, self-protective actions taken by some agents may force other agents to increase their own protective actions through risk spillovers, a scenario often referred to as arms race.[1]

Levee building along the Mississippi River appears to be a case in point. A levee is a man-made structure, usually an earthen embankment, that is commonly constructed along the Mississippi river to divert water flow and reduce risk from temporary flooding. Despite the purpose of flood protection, growing concerns have been raised that levees heightened in some areas in response to catastrophic flood events increased flood risks downstream, inducing transboundary spillovers.[2] Nevertheless, jurisdictions along the Mississippi River keep raising

[1] Some other cases of arms race include investment in high-frequency communication between brokerages and trading floors (Budish et al., 2015), intercollegiate athletic department expenditure decisions (Hoffer et al., 2015), and purchasing increasingly large vehicles for road safety (White, 2004).
[2] See an animation of how overbuilt levees along the upper Mississippi River push floods onto others: https://projects.propublica.org/graphics/rockisland.

levee heights to avoid increased flooding, creating the so-called "levee war" (Hersher, 2018; Maher, 2019).

This paper provides both theoretical and empirical evidence of the relationship between levee building and flood risks, with a particular focus on spatial spillovers from upstream levee construction. Specifically, I seek to address the following three questions. First, how does a jurisdiction respond to exogenous flooding shocks that change one's perceived flood risks with levee building? Second, how large are the spatial externalities when levees are heightened in response to rising flood risks? Third, what possible policy instruments can be designed and implemented to address the spatial externalities induced by levee building?

I build a stylized model to derive a jurisdiction's privately optimal levee building in response to flood risks[3] and establish two model predictions. First, I show that the optimal levee height increases with the perceived probability of flooding. Second, the privately optimal levee height is higher in the case when there are spatial externalities from upstream levee building, as opposed to the case when there is no such influence. To empirically test for the case of levee building in response to flood risks, I take advantage of the Great Mississippi Flood of 2011 as a plausibly exogenous shock. The flood event creates two sources of exogenous variation: spatial variation in flood coverage and temporal variation pre- and post-flood. I leverage this exogenous variation by comparing levee elevations between flooded and unflooded counties before and after the flood event in a difference-in-differences (DID) framework. Combining county variation in levee building and the water flow in the Mississippi that identifies the exogenous upstream and downstream relationship, I estimate the spillover effects of upstream levee building.

[3] The focus on private optimum instead of social optimum stems from the fact of highly decentralized planning on levee building in the Mississippi River Basin. Section 1.2.3 discusses the institutional background in more detail. The comparison between private optimum and social optimum is discussed in Appendix A.1.

I assemble several data sources to construct a unique dataset that records all the levees constructed along the Mississippi River and its tributaries, covering eight states in both the upper and lower Mississippi River Basins (i.e., Arkansas, Iowa, Illinois, Kentucky, Louisiana, Mississippi, Missouri, and Tennessee). To fully recover the historical levee elevations before and after the Great Mississippi Flood of 2011, I combine the newly digitized levee data from the U.S. Army Corps of Engineers (USACE)'s National Levee Database with the U.S. Geological Survey (USGS) topographic maps and National Elevation Dataset. Taking advantage of the geo-referenced levee data, I identify whether a jurisdiction faces any levees from an immediate upstream neighbor and the corresponding levee elevation. To proxy for the values protected by levee construction, I obtain the county-level data on farm acreage, farmland value, and total agricultural sales from the U.S. Department of Agriculture (USDA)'s Census of Agriculture and median housing value from the American Community Survey covering both pre- and post-flood periods. Flood count data at the county level and levee-specific flood zone designations are collected from the Federal Emergency Management Agency (FEMA) to further control for flood risks that could also affect levee construction.

My results show that, consistent with my model predictions, the Great Mississippi Flood of 2011 increased levee building in flooded counties by over 2 ft. Notably, I find that a 1% increase in levee elevation upstream increased levee elevation downstream by 0.7%. This finding implies that a jurisdiction's levee building for self-protection can induce extra flood risks and external costs to its downstream neighbors. A back-of-the-envelope calculation indicates that spillovers from upstream levee building can amount to at least $0.2 billion external costs for downstream jurisdictions. I further show that considering the external costs of transboundary spillovers, the net benefits of heightened levees (defined as the expected benefits minus the costs) can be reduced by 55%. This suggests that current levee building under highly decentralized planning is likely not cost effective. My results therefore highlight

the importance of regional coordination to manage the damage from large-scale natural disasters while mitigating inter-jurisdictional spillovers. Taking advantage of recent advances in remote sensing data that can monitor and measure levee building over time, I discuss relavent policies through spatially explicit compensation schemes in terms of financial transfers across jurisdictions to address the negative externalities from upstream levee building.

This paper connects multiple strands of literature and extends them in several ways. First, inter-jurisdictional spillovers in relation to decentralized planning have been increasingly examined in the setting of environmental regulations on water and air pollution. Prior studies generally find that decentralized planning leads to the degradation of water quality downstream (Sigman, 2005; Cai et al., 2015; Lipscomb and Mobarak, 2017) and air quality downwind (Yang et al., 2017). Such externalities have raised the question of whether a centralized policy can outperform decentralized planning in the presence of inter-jurisdictional spillovers (Banzhaf and Chupp, 2012; Fell and Kaffine, 2014) and the plausibility of spatially explicit taxation in addressing externalities (Kuwayama and Brozovic, 2013; Brown et al., 2020). This paper adds to this line of literature by providing new evidence that decentralized management of natural disasters can create inter-jurisdictional spillovers and such externalities are likely not internalized without regulation from upper-level authorities. In particular, the spatially explicit compensation schemes in terms of financial transfers across jurisdictions proposed in this paper offer one possible solution to correct this market failure in a relatively inexpensive way by using the remote sensing data that can monitor and measure levee building over time.

Second, self-protective actions that may induce external risks and costs to other agents have been studied in different contexts. For example, Lakdawalla and Zanjani (2005) focus on terrorism insurance subsidies, arguing that public investments in self-protection against terror can increase the risks faced by other agents. Anderson and Auffhammer (2014) show that raising vehicle weight by 1,000 pounds to keep occupants safer generated a 40%–55% increase

in fatality risk for other passengers. Private property owners' coastal erosion protection lowered the value of neighboring land by 8% due to altered shoreline wave dynamics (Dundas and Lewis, 2020). This paper focuses on another self-protective behavior, levee building in the Mississippi, that causes external flood risks and costs to downstream neighbors. Moreover, this paper takes a further step and contributes to the literature by estimating how agents respond to such spatial spillovers and quantifying the external costs of arms race.

Third, this paper contributes to the broader literature on spatial spillovers in natural resource management (Cunningham et al., 2016; Costello et al., 2017) and technology adoption (Rode and Weber, 2016; Burlig and Stevens, 2019; Missirian, 2020). The hydrology literature has extensively investigated levee building associated with flooding both theoretically (Yen, 1995; Pitlick, 1997) and empirically (Heine and Pinter, 2012; Remo et al., 2018). This paper is the first to quantify spatial externalities from levee construction and provides economically meaningful evidence on the external costs of levee building that are often overlooked during the decentralized decision-making process. Incorporating spatial externalities from levee building advances our understanding of the economic aspect of flood-levee system planning and policies (Vousdoukas et al., 2020), especially in face of rising flood occurrences due to climate change.

Fourth, the relationships between natural disasters, risk perceptions, and risk management behaviors have been increasingly documented. For instance, Lawrence et al. (2014) find that previous flood experience contributes to a heightened perception of flood risks and increased preparedness of households in face of future flood events. However, risk management behaviors are highly variable and may not always be rational. McCoy and Zhao (2018) show that homeowners are more likely to invest in a damaged building located in flood risk areas after a hurricane event. While most studies aim to capture flood risk perception by estimating housing prices (e.g., Gibson and Mullins, 2020; Hennighausen and Suter, 2020),

this study advances this line of literature by examining how municipal agents respond to catastrophic flood events through building up levees along the rivers as an adaptation to increased flood risks.

Last but not the least, this paper speaks to the vast literature on a variety of long-term socioeconomic impacts of natural disasters, floods and hurricanes in particular, on a larger spatial scale. Prior studies generally find that floods and hurricanes have negative impacts on economic growth (Belasen and Polachek, 2009; Strobl, 2011) and property values (Kousky, 2010; Bin and Landry, 2013; Beltran et al., 2019), with a regional focus on coastal areas. One exception is Hornbeck and Naidu (2014) that examine the impact of the Great Mississippi Flood of 1927 on black out-migration and subsequent agricultural development in the lower Mississippi River basin. This paper adds to this strand of literature by studying a more recent catastrophic flood event in non-coastal areas (i.e., the upper and lower Mississippi River Basins) and particularly estimating the extra social cost of flooding resulting from overbuilt levees.

1.2 Background and Study Area

1.2.1 Flooding history in the Mississippi

Floods have been part of the earliest recorded history of the Mississippi that can date back to the mid-1500s. Since the 20th century, a growing number of significant flood events have occurred in the Mississippi, with catastrophic flooding in 1927, 1936, 1973, 1993, 2011, and 2019.[4] Most of these flood events resulted from regional rainfall and snowmelt that caused raises on rivers, inundating surrounding land for days or even weeks and inducing substantial socioeconomic consequences. For example, the Great Flood of 1927 in the lower Mississippi

[4] See more details on the dates, deaths, impacts, flood inundations, and costs of Mississippi River floods since the 20th century here: https://www.weather.gov/media/jan/JAN/Hydro/Flood_History_MS.pdf.

devastated over 67,340 km^2 of landmass, took more than 200 lives, and cost over $10 billion in damage to landowners and municipalities (Barry, 1998). The flood in 1993 inundated the upper Mississippi and Missouri basins, with 38 deaths attributed directly to the flood. Estimates of damages ranged from $12 to $20 billion and over 6.6 million acres of land were flooded, with agriculture accounting for over half of these damages (Galloway, 1995).

The Great Mississippi Flood occurring in April and May 2011 is among the largest and most damaging recorded along the U.S. waterway in the past century, comparable in extent to the major floods of 1927 and 1993. Heavy springtime snowmelt combined with two major storm rainfalls across the eastern half of the Mississippi watershed resulted in one of the most powerful floods in the river's known history (USACE, 2011). The flood affected more than 1.2 million acres of agricultural land and cost over $2 billion across the entire river system. Areas along the Mississippi River and its tributaries that experienced flooding span Iowa, Illinois, Missouri, Kentucky, Tennessee, Arkansas, Mississippi, and Louisiana (see Figure 1.1 for flooded counties in 2011).

1.2.2 Flood control policies and regulation

Due to the collective nature of flood control, the earliest flood-control efforts were fairly local, where farmers constructed levees along the rivers to protect farmland from frequent flooding. Such cases in the U.S. can date back to the early 18th century when the very first levee was built in New Orleans between 1718–1727 as the first Europeans settled the region (Cowdrey, 1977). The increased floods in the 19th century and the successive collapse of local levee systems led to regional efforts in combating floods, especially in the lower Mississippi Basin (Humphreys, 1914). For instance, a major Mississippi River flood in 1874 inspired the creation of the Mississippi River Commission (MRC) in 1879, with the goal of developing and implementing plans to prevent destructive floods. In 1882, MRC adopted a policy of

coordinated local efforts in setting standards and specifications for levee construction and allocating federal funds to local levee districts that are financially constrained.

Despite extensive local and regional efforts, the federal government got involved when those efforts became incapable of providing enough flood control. At the federal level, there are multiple laws known as the Flood Control Act (FCA) in the U.S. The history of federal flood control regulations can originate from the Swamp Land Acts of 1849 and 1850, the first significant federal flood control laws that encouraged the reclamation of flood-prone swampland for agricultural uses. Severe floods on the Mississippi, Ohio, and other rivers in the Northeast between 1907 and 1913 led to the FCA of 1917, the first act aimed exclusively at controlling floods. The Great Mississippi Flood of 1927 changed MRC's mission and the consequent FCA of 1928 created the Mississippi River and Tributaries Project (MR&T), assigning to MRC the responsibility for developing and implementing the project (Camillo, 2012). A series of floods during 1935-1936 across the nation were critical in passing the FCA of 1936 by the U.S. Congress, authorizing flood-control projects such as levees and floodwalls through USACE and other federal agencies (Arnold, 1988).[5] The most recent act was passed in 1965, which granted USACE more authority to design and construct any water resource development projects for flood control. Since then, USACE has become the primary federal agency that regulates flood-related projects for flood control.

Although USACE plays an essential role in conducting federal flood control measures, inter-agency relations are part of larger efforts in flood damage reduction across the nation nowadays.[6] For instance, USACE actively cooperates with FEMA in the federal levee certification program within the National Flood Insurance Program. Specifically, FEMA identifies flood hazards and assesses flood risks in levee-affected areas. FEMA also determines

[5] Arnold (1988) provides more details on the origins of federal flood control activities and the evolution of the 1936 FCA.
[6] National Research Council (2000) provides more details on USACE and U.S. flood damage reduction planning, policies, and programs as well as the history of levee certification.

and establishes appropriate flood risk zone designations in areas landward of levees and reflects those zones on Flood Insurance Rate Maps (FEMA, 2016).

1.2.3 Levee building in the Mississippi

Due to the collective nature of flood control, the earliest flood-control efforts were fairly local, where farmers constructed levees along the rivers to protect farmland from frequent flooding. Since levees are relatively cheap and easy to construct while providing great protection up to their design standards, the implementation of levee projects has become a common practice among floodplain communities (Tobin, 1995). Draining about 40% of the continental landmass in the U.S., the Mississippi River and its tributaries are protected by thousands of miles of earthen levees. In the 1920s, private and public levees along the Mississippi River were extensively engineered to protect farmland and cities along the rivers against flooding from the Arkansas, Ohio, Missouri, and Mississippi rivers, all the way south to the Gulf of Mexico (see Figure 1.1 for the spatial distribution of levees and Mississippi River tributaries in the study area).

Levee construction has become the primary flood-control strategy when the Congress appointed USACE as the lead responsibility for protecting land along the Mississippi River. Formed in 1802, USACE has been playing an irreplaceable role in helping reduce flood damage in the U.S. In addition to constructing and maintaining thousands of miles of levees, USACE provides local levee or drainage districts with levee certification criteria. USACE also actively works with FEMA on federal flood insurance policies based on multiple factors that include uncertainties in the frequency of floods, changes in land use, climate change, and the structural and geo-technical performance of levee systems (National Research Council, 2000). In particular, USACE sets standards on levee heights to prevent a flood of a given recurrence

probability.[7] In detail, levee heights are determined by the National Economic Development criterion based on prescribed benefit calculation procedures rather than by a levee's ability to withstand a flood of a given magnitude (USACE, 1999).

Despite the federal involvement, the centralization of the flood-control policy was nevertheless impeded by political reasons, with the Jeffersonian ideology in the 19th century advocating that flood control was a local, or at best, a state function, even though rivers in the Mississippi span many states (Klein and Zellmer, 2007; Tarlock, 2012). Today, there is still no unified levee system in the U.S. and nearly 85% of levees recorded in the USACE Levee Safety Program[8] are locally owned and maintained (FEMA, 2016). Furthermore, there is no national policy related to the safety of levees, with responsibility for levee safety often assigned in an uncoordinated manner, distributed across all levels of government, and housed in different agencies (FEMA, 2016). Federal, state, and local agencies generally have varying policies and criteria concerning many aspects of levee design, construction, operation, and maintenance.

Currently, four types of levees exist with differing degrees of federal involvement: (1) Levees built, operated, and maintained by the federal government; (2) Levees built by the federal government and then turned over for operation and maintenance by a local sponsor; (3) Levees constructed by a non-federal entity and are enrolled in the USACE's Rehabilitation and Inspection Program (RIP); and (4) Levees that were neither built by the federal government nor are part of federal operations or maintenance programs. Although the majority of levees are locally constructed, USACE supplements local efforts in the repair of these levees damaged by flood through the Rehabilitation and Inspection Program. To be eligible for rehabilitation

[7] For example, a flood with a probability of 0.01 of being equaled or exceeded in any given year is commonly called the "100-year flood."

[8] Created by the USACE in 2006, the mission of the Levee Safety Program is to assess the integrity and viability of levees in the U.S. and provide technical support and recommendations to make sure that levee systems do not present unacceptable risks to the public, property, and environment.

assistance,[9] projects must be inspected, evaluated, and enrolled in the program prior to the onset of the flood. Violation of USACE's regulation on levee building (e.g., authorized levee height) may lead to USACE withholding federal funds for maintaining these levees.[10] Other than RIP through which USACE exerts regulatory authority, USACE has limited jurisdiction over locally constructed and maintained levees, although USACE regularly provides recommendations on levee construction and maintenance for local levee sponsors.

Specifically for the levees that are locally owned, the cost of constructing and maintaining these structures generally exceeds the capability of a single property owner to undertake the funding alone. Levee or drainage districts at the local level are therefore formed by a group of interested property owners to control river levels for agricultural and sanitary purposes. The General Assembly in each state along the Mississippi River has passed laws such as "levee law" or "drainage law" that allow property owners in any county or city within the state to form a levee or drainage district. These districts are authorized to tax beneficiaries behind levees for constructing and maintaining these levees to protect their land (Harrison, 1951).

As a political subdivision with the status of a corporation, county levee or drainage districts are typically governed by a multiple-member Board that is comprised of landowners within the district or qualified designees. In cases when a multicounty levee district is established, the Board shall be elected by the county judges or executives of the counties involved, with appointments apportioned among the counties in ratio to the portion of the levee to be constructed within each county. The Board members serve a multiple-year term and the

[9] Rehabilitation assistance for a non-federal project is cost shared between the non-federal entity and USACE, where the non-federal entity must provide 20% of the cost.

[10] One recent example is that in 2015, USACE announced that it would withhold future federal funds for maintaining the overbuilt levees in the Sny Island Levee and Drainage District in Illinois due to higher-than-designed levee elevations and modified slopes of the levees. See the announcement letter here: https://www.documentcloud.org/documents/4465957-Army-Decert.html.

length of the term varies across states.[11] The Board annually provides either the county court or Collector's Office with assessment information on the land directly benefiting from the levee. The county court or Collector's Office accordingly creates and sends out bill based on the assessed amounts (often referred to as levee taxes) and distributes the collections back to the district. Levee bills are handled similarly to real estate bills that collect property taxes from property owners.

1.2.4 Levee building and flooding

While the purpose of levee building is to protect farmland and properties from flooding, levees have been found to increase flood heights and flow rates for a given discharge (Yen, 1995; Pitlick, 1997) and potentially speed up downstream propagation of a flood wave (Jacobson et al., 2015). Heine and Pinter (2012) used stream gauge records that span pre-levee and post-levee periods to assess the impact of levees on flood levels in Illinois and Iowa. They found that for above bankfull (flood) conditions, stages at sites within leveed reaches were statistically distinguishable before and after levee construction, with the median station-averaged stage increase being nearly half a meter. Similarly, Remo et al. (2018) assessed the long-term discharge along the Mississippi River between St. Louis, Missouri and Vicksburg, Mississippi and revealed that levees significantly increased peak flood discharges by up to 25%.

The consequences from levee-induced flood can be substantial, however. Remo et al. (2012) quantified the balance between the benefits of levees for flood protection and their potential to increase flood risk in the Middle Mississippi River. They found that although agricultural levees along the river protect against small- to medium-size floods (up to the 100-

[11] For the Board that governs the levee or drainage district, for example, Illinois has three members for a term of three years, Kentucky has five members for a term of four years, and Missouri has five members for a term of five years.

year flood level), they cause more damage than they prevent during large floods such as the 500-year flood. Combining hydraulic and economic modeling, Pinter et al. (2016) documented that levee-related surcharge plus the residual risk of levee overtopping can lead to negative benefits, adding long-term flood risk. Although counterintuitive, structures at the margin of a leveed floodplain can incur negative benefits due to greater flood levels resulting from levees purportedly built to protect them.

The concern of increased flood heights and flow rates induced by levee building has brought scholars to investigate the extent to which levees in the Mississippi have been heightened. Flor et al. (2011) measured the changes in levee-crest elevations along 328 kilometers of Mississippi River levees between St. Louis, Missouri and Cairo, Illinois during 1998–2007. Combining the high-resolution digital elevation model and GPS-based levee-crest surveys, they find that about 11–18% of surveyed levees were distinguishably heightened, with increases of as much as 1.49 meters. The USACE surveyed roughly 200 miles of levees along the Mississippi in 2016 and found that 40% of these levees were higher than authorized by federal law (Maher, 2019). In particular, levees were found to be too high in Iowa, Missouri, and Illinois. While raising levee heights in response to more frequent flooding has been increasingly debated, no study to date has examined how jurisdictions react to each other's levee building due to flooding shocks nor the spillover effects of the heightened levees.

1.3 Theoretical Model

In this section, I illustrate a theoretical model with two agents to produce testable hypotheses that inform my subsequent empirical analyses. The two-agent model allows for spatial externalities from one agent to the other and derives the privately optimal levee height for each agent. The two-agent model with spatial externalities can be generalized to cases with multiple agents when estimating the cascading effect (see Appendix A.2).

Consider two adjacent counties k and $k-1$ along the Mississippi River such that county $k-1$ is the upstream county and county k is the downstream county (see Figure 1.2 for a geographical illustration).[12] Each county chooses to build a levee with a certain height $L_{.,t}$ to protect the county's value such as property or agricultural values in time period t. The perceived flood probability occurs with $\Phi_{.,t} \in [0,1]$ and $V_{.,t}$ is the value lost in each period if there is a flood. Let $\mathcal{L}_{.}$ be the maximum levee height a county can build due to exogenous technical constraints that is assumed to be constant in the model. $\left(\frac{L_{.,t}}{\mathcal{L}_{.}}\right)^{a} \in [0,1]$ can then be interpreted as the proportion of values saved by levee building if there is a flood, which is assumed to be concave with $0 < a < 1$. Levee building induces some cost $A{L_{.,t}}^{b}$ that is assumed to be convex, where $b > 1$ and $A > 0$ is the cost per level of levee building.

Levee building in the upstream county is assumed to increase the occurrence of flooding in the downstream county. As a result, the perceived probability of flooding occurrence in county k changes from $\Phi_{k,t}$ to $\Phi_{k,t}(1+\gamma L_{k-1,t})$ if there are any levees upstream, where $\gamma > 0$. This functional form assumes no spatial externalities if there is no levee building in the upstream county ($L_{k-1,t} = 0$), a reasonable assumption in this study. Each county chooses its own levee building $L_{.,t}$ to minimize the sum of the expected values lost due to flooding and the cost of levee building, $C_{.,t}$:

$$C_{k-1,t} = \Phi_{k-1,t}V_{k-1,t} - \Phi_{k-1,t}V_{k-1,t}\left(\frac{L_{k-1,t}}{\mathcal{L}_{k-1}}\right)^{a} + AL_{k-1,t}^{b} \tag{1.1}$$

$$C_{k,t} = \Phi_{k,t}(1+\gamma L_{k-1,t})V_{k,t} - \Phi_{k,t}(1+\gamma L_{k-1,t})V_{k,t}\left(\frac{L_{k,t}}{\mathcal{L}_{k}}\right)^{a} + AL_{k,t}^{b} \tag{1.2}$$

Take the first-order condition and re-arrange in terms of levee height for each county:

[12] The model setup is analogous if the two counties are cross-river neighbors. In this case, both counties' levee building can possibly induce spatial externalities and thus increase each other's occurrence of flooding. Model predictions remain the same (see Appendix A.3 for detailed derivations).

$$L_{k-1,t}^{*} = \left(\frac{a\Phi_{k-1,t}V_{k-1,t}}{bA\mathcal{L}_{k-1}^{a}}\right)^{\frac{1}{b-a}} \tag{1.3}$$

$$L_{k,t}^{*} = \left(\frac{a\Phi_{k,t}(1+\gamma L_{k-1,t})V_{k,t}}{bA\mathcal{L}_{k}^{a}}\right)^{\frac{1}{b-a}} \tag{1.4}$$

Equations (1.3) and (1.4) respectively define the privately optimal levee height $L_{k-1,t}^{*}$ and $L_{k,t}^{*}$ for the upstream and downstream county, in which the marginal cost of levee building is equal to the marginal increase in the values saved by levee building. Both equations predict the first relationship of interest: the optimal levee height is increasing in the perceived probability of flooding as $\frac{\partial L_{k-1,t}^{*}}{\partial \Phi_{k-1,t}} > 0$ and $\frac{\partial L_{k,t}^{*}}{\partial \Phi_{k,t}} > 0$. Substituting $L_{k-1,t}^{*}$ from equation (1.3) into equation (1.4), the privately optimal levee height in the downstream county can be re-written as:

$$L_{k,t}^{*} = \left(\frac{a\Phi_{k,t}V_{k,t}}{bA\mathcal{L}_{k}^{a}}\right)^{\frac{1}{b-a}} \left[1 + \underbrace{\gamma\left(\frac{a\Phi_{k-1,t}V_{k-1,t}}{bA\mathcal{L}_{k-1}^{a}}\right)^{\frac{1}{b-a}}}_{spatial\ externality}\right]^{\frac{1}{b-a}} \tag{1.5}$$

Equation (1.5) shows the second relationship of interest: Compared to the case without considering spatial externality, the privately optimal levee height in the downstream county is higher if there is any levee building in the upstream county. Take the log of both sides of equation (1.5):

$$\begin{aligned} log\left(L_{k,t}^{*}\right) = \frac{1}{b-a}\Bigg\{ & log\left(\frac{a}{bA\mathcal{L}_{k}^{a}}\right) + log\left(\Phi_{k,t}\right) \\ & + log\left[1 + \gamma\left(\frac{a\Phi_{k-1,t}V_{k-1,t}}{bA\mathcal{L}_{k-1}^{a}}\right)^{\frac{1}{b-a}}\right] + log\left(V_{k,t}\right)\Bigg\} \end{aligned} \tag{1.6}$$

Equation (1.6) serves as a reduced-form equation for empirically estimating the impact of flooding on levee building when there are transboundary spillovers from upstream levee building.

1.4 Empirical Strategy

Predictions from the theoretical model show that (i) the optimal levee height increases with the perceived probability of flooding; and (ii) the optimal levee height is higher if there are spatial externalities from upstream levee building. The coefficients estimated from the econometric model should capture these two relationships. I employ a DID framework to test for the predictions of the model to data. Taking the Great Mississippi Flood of 2011 as a plausibly exogenous shock that changed perceptions of flood risk and using levee elevation data before and after flooding, I estimate the following model specification to test for prediction (i):

$$\begin{aligned} log(L_{ikt}) = \beta_1 Flooded_k + \beta_2 Post_t + \beta_3 Flooded_k \times Post_t \\ + log(\boldsymbol{X'_{ikt}})\delta + \theta_k + \epsilon_{ikt} \end{aligned} \tag{1.7}$$

where $log(L_{ikt})$ is the logged elevation of levee i belonging to county k in time period t, $Flooded_k = 1$ if county k was flooded in 2011, $Post_t = 1$ if levee elevation was measured after 2011, $log(\boldsymbol{X'_{ikt}})$ are the logged county-level control variables to proxy for the values lost if flooding occurs as well as levee characteristics to control for levee construction, θ_k are county fixed effects, and ϵ_{ikt} is an idiosyncratic error term. Standard errors are clustered at the county level to allow for arbitrary error dependence over time between levees in the same county. To more formally allow for the possibility of spatial autocorrelation, I alternatively use the robust standard errors following Conley (1999) and specify the distance cutoff at 50 km. The coefficient of interest β_3 captures the impact of flooding on levee elevation.

The DID approach relies on the identifying assumption that there are no time-varying unobserved factors that are different between counties that were flooded and unflooded in 2011. This assumption seems reasonable in this study for at least four reasons. First, as described in Section 1.2.1, fairly exogenous factors (i.e., two major storm rainfalls) resulted in the occurrence of flood in 2011 that inundated part of the Mississippi River and its tributaries,

making flooded areas as good as random. Second, I show that historic flood events in the study area demonstrated similar trends in terms of flood frequency between flooded and unflooded counties in 2011 (to be discussed in more detail in Section 1.5.3). Third, the proportion of levees located in the 100-year floodplain pre-2011 is shown to be not statistically different between flooded and unflooded levees in 2011.[13] The above three reasons provide strong evidence of the exogeneity of the flood event in 2011. Lastly, Figure A.1 plots the kernel densities of the change in levee elevation between the two pre-flood periods for flooded and unflooded levees. Flooded levees trended similarly to unflooded ones in terms of elevation, satisfying the parallel-trend assumption. The difference in the mean change in levee elevations between flooded and unflooded levees is found to be statistically insignificant, indicating that levee building prior to 2011 was not correlated with the occurrence of flooding in 2011.

To test for prediction (ii), I compare the elevation differences between flooded and unflooded levees with varying levee locations, before and after the flooding shock in 2011. I estimate the following model specification to capture the impact of upstream levee building on one's own levee construction:

$$\begin{aligned} log(L_{ikt}) = \beta_1 Flooded_k + \beta_2 Post_t + \beta_3 Flooded_k \times Post_t \\ + \beta_4\, log\left(\overline{L_{k-1,t}}\right) + log(\boldsymbol{X'_{ikt}})\delta + \theta_k + \epsilon_{ikt} \end{aligned} \tag{1.8}$$

where $log\left(\overline{L_{k-1,t}}\right)$ is the logged average elevations of levee(s) in county k's immediate upstream county. The other variables remain as described in equation (1.7). β_4 is the coefficient of additional interest that captures the impact of immediate upstream levee elevation on one's own levee elevation.

To further address the concern of the endogeneity of upstream levee building, I use an instrumental variables (IV) approach to estimating equation (1.8). Specifically, I use the

[13] A two-sample t-test with equal variances shows that the difference in the proportion of levees located in the 100-year floodplain between flooded and unflooded levees in 2011 is not statistically significant.

flooded status in the upstream county as an instrument for upstream levee building. The exclusion restriction is that whether the upstream county was flooded should not directly affect downstream county's levee building unless through upstream levee building that increases downstream's flood risk perceptions. The first-stage equation for the IV approach is as follows:

$$log\left(\overline{L_{k-1,t}}\right) = \gamma_1 Flooded_{k-1} + \gamma_2 Post_t + \gamma_3 Flooded_{k-1} \times Post_t + log\left(\boldsymbol{X'_{k-1,t}}\right)\phi + \epsilon_{k-1,t} \tag{1.9}$$

1.5 Data

I use detailed data on individual levees and county-level agricultural measures, housing values, and flood counts to estimate the impact of flooding on levee building in the Mississippi, focusing on spillover effects of upstream neighbors' levee building. In detail, I use data from 1,249 levees along the Mississippi River and its tributaries that belong to 207 counties in eight states covering both the upper Mississippi River Basin (i.e., Iowa, Illinois, and Missouri) and lower Mississippi River Basin (i.e., Arkansas, Kentucky, Louisiana, Mississippi, and Tennessee) before and after the Great Mississippi Flood of 2011.

1.5.1 Levee building in the Mississippi

I acquire the geo-referenced levee data from the USACE's National Levee Database. For each recorded levee, the dataset includes the information on levee name, length of embankment, number of segments, geo-referenced location, year constructed, sponsors, USACE division and district, FEMA region, and additional notes on levee construction. The levee elevation data are derived from two main sources: the USGS topographic maps that provide levee elevations for the pre-flood period (1982–2010)[14] and the digital elevation

[14] One concern of using the elevation data back in the 1980s for the pre-flood period is the occurrence of the Great Flood of 1993 in the upper Mississippi that could potentially bias my estimation. If there is a positive impact of flooding on levee building, my result will likely be overestimated.

models (DEMs) from the USGS 3D Elevation Program (3DEP) that provide levee elevations for the post-flood period (2012–2018).

In detail, the historical topographic maps are provided in quadrangle units at different scale levels (e.g., 1:24,000, 1:62,500) based on certain years when imageries were taken. These maps are available from TopoView, an interface created by the National Geologic Map Database project in support of monitoring some natural and cultural features that have changed over time. Overlaying the geo-referenced levee file with these topographic maps, I extract levee elevations from these maps for the particular pre-flood year when the maps were created. Figure A.2 illustrates a topographic map that shows the levees with elevations, using the Granite City Quadrangle Illinois-Missouri 7.5-minute Series at the scale of 1:24,000 (the finest scale that is available) in 1998 as an example.

USGS DEMs are arrays of regularly spaced elevation values referenced horizontally either to a Universal Transverse Mercator projection or to a geographic coordinate system. The grid cells are spaced at regular intervals along south to north profiles that are ordered from west to east. USGS acquires bare-earth elevation source data through the 3DEP and resamples the data to several National Map DEM products for the U.S. and its territories. The DEM products include 2, 1, 1/3, and 1/9 arc-second layers, which are logically seamless terrain surfaces in their respective areas of coverage and are the highest quality elevation data held by USGS. Again, I overlay the geo-referenced levee file with 1/3 arc-second (about 10 meters) layers that cover the most of my study area and extract the levee elevations from these layers for the particular post-flood year when the DEMs were acquired.

I measure levee elevation (above sea level) instead of levee height in this study for two main reasons. First, although the National Levee Database contains the data on average levee height, this information is largely missing in the database. For those levee heights that are recorded, they are mostly the heights when levees were constructed without information on

current levee heights. To obtain the changes in levee height pre- and post-flood, levee elevation can be used assuming that changes in levee elevation are approximately the changes in levee height over time.[15] Second, levee elevations, the sum of levee height and riverbed elevation, indeed determine whether (and how much) extra water can be pushed from one area to another if there is a flood.

In my main analysis, I focus on the levees that were constructed before 1982,[16] which is the latest year when the elevation for some levees in my study area could be measured using the USGS topographic maps for the pre-flood period. This process ensures that all the levees in my data cover both the periods before and after the flooding shock in 2011 to form a two-period balanced panel (i.e., one pre-flood and one post-flood). Table A.1 lists the years when levee elevations were measured pre- and post-flood in this study. Note that although levee elevations could be measured for two pre-flood and two post-flood periods, the two pre-flood periods are only used to verify the pre-trend assumption in Figure A.1 but not used for empirical analyses due to the unavailability of elevation data that would lose about 33% of levees in my data sample.

1.5.2 Agriculture data

I draw the agriculture data from the USDA's Census of Agriculture to control for jurisdictions' levee building. Censuses were taken only once every five years, with data collection taking place in the fall during the Census year. Data on a complete count of U.S. farms and ranches and the people who operate them were initially collected. After collection,

[15] One concern of using levee elevation changes to proxy for levee height changes is the potential decrease in riverbed elevation due to soil erosion as a result of flooding. If this is the case, the results should be interpreted as the lower-bound estimates for the impact of flooding on levee building in my study. However, the decrease in riverbed elevation because of flooding shall unlikely confound identification.

[16] According to the National Levee Database, the majority of recorded levees were constructed during the 1920s, 1950s, and 1970s. Less than 10% of them were constructed after the 1980s.

the Census underwent a multi-stage quality control process, and the final dataset is available at the county level to keep the information provided by individual respondents private.

I use the Census data from 1982–2007 and 2012–2017, respectively for the pre- and post-flood period, to create a balanced panel of 207 counties along the Mississippi River and its tributaries that have any levees. The main variables of interest in this study from the Census of Agriculture include farm acreage, farmland value (i.e., estimated market value of land and buildings per acre), and total agricultural sales (i.e., market value of agricultural products sold), all of which are used to proxy for the values protected by levee building. Mean values are used across Census years (i.e., the pre-flood measures are the mean of Census years of 1982, 1987, 1992, 1997, 2002, and 2007; the post-flood measures are the mean of Census years of 2012 and 2017).

1.5.3 Housing value data

To further account for jurisdictions' levee building that protects non-agricultural values, I obtain the 5-year data on county-level median housing value from the American Community Survey. The median housing value was measured among the universe of owner-occupied housing units in each county over the course of 5 years. The 5-year data cover all areas, have the largest sample size, and are most reliable when compared to 1-year and 3-year estimates also provided by the American Community Survey. I use the 2006–2010 and 2012–2016, respectively for the pre- and post-flood period, to create a balanced panel of 207 counties along the Mississippi River and its tributaries that have any levees.

1.5.4 Flood count data

One may argue that areas flooded in 2011 may be more prone to other floods absent the Great Mississippi Flood of 2011. If this is the case, the assumption of the exogenous shock that changed perceptions of flood risk may be violated and my identification using the DID

approach could thus be contaminated and lead to biased causal estimation. To address this concern and control for the possible influence of other floods on levee building, I draw the data from the list of major disaster declarations (incident type: flood) posted by FEMA to create a consistent series of flood counts at the county level. For each flood, FEMA provides information on the geographic location (county) as well as the month and year of occurrence. I aggregate the yearly flood counts to the pre-flood period (1982–2010) and post-flood period (2012–2018) at the county level in this study. Controlling for additional flood events in the model also mitigates the concern that pre-flood levee elevations were not measured in the same year for all levees.

To further provide evidence that can support the assumption of exogenous flooding shock in 2011, I use the flood count data compiled by Boustan et al. (2020).[17] Figure A.3 plots the average and the total number of flood counts at the county level in each decade during the period 1960–2010 (the pre-flood period in this study) for counties that were flooded or unflooded during the Great Mississippi Flood of 2011. First, the figure shows strong evidence of similar pre-trend in terms of flood counts between flooded and unflooded counties in 2011. This suggests that should other floods also influence levee building, the effects are likely to be similar between flooded and unflooded counties. Second, a two-sample t-test with equal variances shows that the difference in mean flood counts between flooded and unflooded counties is statistically insignificant in all decades except for the decade of 1990.

1.5.5 Floodplain maps

Floodplain maps that show areas with different flood risks are obtained from FEMA's National Flood Hazard Layer Viewer. By overlaying levees with these floodplain maps, I identify whether each levee is located in a 100-year floodplain (an area inundated by 1% annual

[17] I appreciate that the authors shared this wonderful dataset when they published the article. See more details on data collection and process in Boustan et al. (2020).

chance flooding) or 500-year floodplain (an area inundated by 0.2% annual chance flooding). According to FEMA's flood zone designations, 100-year and 500-year floodplains are considered high-risk and low-risk areas, respectively.

1.5.6 Summary statistics

Table 1.1 presents the summary statistics for levee and county characteristics. All the statistics are the average over the two periods (i.e., one pre-flood and one post-flood). One immediate observation from the table is the substantial heterogeneity in terms of both levee and county characteristics across units. For example, the average levee elevation across the study area is about 427 ft, with the lowest levee elevation of less than two feet and the highest levee elevation of more than 1,100 ft. At the state level, Iowa has the highest average levee elevation of 810 ft and Louisiana has the lowest average levee elevation of 37 ft. Such a variation also reflects the geographical difference in terms of elevation across states. In general, states in the upper Mississippi River Basin have higher elevations while those in the lower Mississippi River Basin have lower elevations.[18]

Figure 1.3 shows Epanechnikov kernel density plots of flooded and unflooded levees, with the left graph comparing the pre-flood logged levee elevation and the right graph comparing the change in levee elevation pre- and post-flood. On average, unflooded levees saw a reduction of 0.12 ft (most likely due to erosion and degradation, see a similar finding from Flor et al., 2011) but flooded levees saw an increase of 0.72 ft. Two-sample t-test with equal variances further shows that the difference in mean change in levee elevations between flooded and unflooded levees is statistically significant. While Figure 1.3 seems to tell a compelling story that the 2011 flood increased the levee elevation, it does not control for other factors such

[18] For example, in the upper Mississippi River Basin, the mean elevation of Iowa, Missouri, and Illinois is respectively 1,100 ft, 800 ft, and 600 ft. While in the lower Mississippi River Basin, the mean elevation of Arkansas, Mississippi, and Louisiana is respectively 650 ft, 300 ft, and 100 ft.

as local economic activities and levee regulations that may affect levee building over time. To capture these possible confounding factors, my empirical strategy employs a DID approach to control for both levee and county characteristics with county and state-by-year fixed-effects.

1.6 Results and Discussion

1.6.1 What predicts levee building?

I first show the relationships between county characteristics and levee building in Table 1.2. Time-invariant levee characteristics are not included due to levee fixed effects used in the model. Column [1] shows the univariate relationship between each county characteristic and levee building while column [2] shows the results including all predictors in the one single regression. Consistent with my theoretical framework, I find a positive correlation between levee building and the county's value to be protected from flooding (e.g., farmland value, total agricultural sales). The total number of flood counts at the county level during the study period seems to be positively correlated with levee building. However, this effect becomes statistically insignificant when other county characteristics are accounted for. This result may suggest that local flood events, whose scale and economic damage are much smaller compared to the Great Mississippi Flood of 2011, will likely not change perceptions of flood risk and are weakly related to levee building at best in my case.

1.6.2 Impact of flooding on levee elevation

I then estimate the impact of flooding on levee building using four model specifications, each with a different set of controls. Model I is the most parsimonious specification, including only the interaction term of interest (Flooded × Post), a postflood dummy, and a "flooded" dummy. Model II adds three time-invariant levee controls (i.e., levee length, the number of segments, whether the levee is in a 100-year floodplain with the 500-year floodplain being the baseline) and Model III adds five time-varying county controls (i.e.,

farm acreage, farmland value, total agricultural sales, median housing value, and flood counts). Model IV includes both levee and county controls. To account for time-invariant and time-varying unobservables, I include county fixed effects and state-by-year fixed effects, respectively, in all the model specifications.

The results in Table 1.3 are consistent with my model prediction. As expected, levees that experienced flooding in 2011 were built higher post-flood than those who did not. These effects are statistically significant and economically meaningful. Across specifications, I find positive point estimates of "Flooded × Post", ranging from 0.7% to 1.1% additional levee elevation as a result of flooding. With the average pre-flood elevation of flooded levees being 292 ft, a 0.7%(1.1%) increase in levee elevation indicates approximately a 2.1(3.3) ft increase in levee building. This estimate largely lines up with prior findings from hydrological studies in the Mississippi (Flor et al., 2011).

1.6.3 Spillover effects and external costs

I next estimate the spillover effects from upstream levee building using the illustration in Figure 1.2 to identify upstream and downstream counties in this study. Again, four model specifications with a different set of controls in each specification are used. In addition to the interaction term of interest (Flooded × Post) that captures the impact of flooding on levee building, I add the logged average elevation of levee(s) in each county's immediate upstream county to the models. Table 1.4 presents the results from both the OLS and IV approaches. I find that, in addition to the positive impact of flooding on levee elevation, a 1% increase in the immediate upstream levee elevation increased the downstream levee elevation by 0.7%, suggesting substantial spillovers. The spillover effect estimated from the IV approach is larger. Notably, the first stage F-statistic, which is much larger than 10, indicates a strong instrument used for upstream levee building.

To further infer the magnitude of spillover effects, I conduct a back-of-the-envelope calculation of the costs of heightened levees induced by the flood event in 2011 as well as the external costs of spillovers that upstream levee building imposes on the downstream county. In the engineering literature, the costs of a floodproofing structure such as a levee are typically evaluated in terms of primary and secondary costs. Specifically, primary costs include the costs of the basic floodproofing elements and secondary costs include the costs of auxiliary materials required to assure that the primary floodproofing elements function properly. Based on a combination of the floodproofing literature, manufacturer's quotes, and information collected from the owners and builders, FEMA reports that the average unit cost for a 3-feet-height, 5-feet-height, and 10-feet-height levee is respectively \$13 per foot length, \$30 per foot length, and \$85 per foot length in 1985 dollars. Adjusted for inflation, these numbers become \$32 per foot length, \$73 per foot length, and \$207 per foot length in 2020 dollars. Using these three data points to fit a power function as assumed in the theoretical model yields $A(L) = 5.92 \times L^{1.55}$, where $A(L)$ is the cost per foot length for a levee with L feet in height. Equations (1.10) and (1.11) show the formulas to compute the costs of levee building induced by the flood event in 2011 with and without spillovers due to upstream levee building:

$$\Delta C_K^S = \sum_i \left[A\left(L_{ik}^{post-2011}\right) - A\left(L_{ik}^{pre-2011}\right)\right] \times l_{ik} \tag{1.10}$$

$$\Delta C_K^N = \sum_i \left[A\left(\hat{L}_{ik}^{post-2011}\right) - A\left(L_{ik}^{pre-2011}\right)\right] \times l_{ik} \tag{1.11}$$

where ΔC_k^S and ΔC_K^N are respectively the costs of levee building with and without spillovers for county k, $A(\cdot)$ is the cost per foot length for a levee with a specific height, $L_{ik}^{pre-2011}$ is levee i's pre-2011 height, $L_{ik}^{post-2011}$ is levee i's observed post-2011 height, $\hat{L}_{ik}^{post-2011}$ is levee i's predicted post-2011 height without spillover effects of upstream levee building, and l_{ik} is the levee length. Specifically, $\Delta C_K^S - \Delta C_K^N$ is the external costs of spillovers due to upstream levee building for county k.

Since the pre-2011 levee heights are not documented in the levee database, simply using changes in levee elevation instead of changes in levee height will lead to inaccurate calculation of costs because the marginal cost of levee building is non-linear and increases with levee height. To mitigate this issue, I use the average levee height reported in USACE (2014) as the pre-2011 levee height $L_{ik}^{pre-2011}$ for levees in each of the eight states in my study area. Specifically, I use the following formulas to back out the post-2011 levee height with and without spillovers:

$$L_{ik}^{post-2011} = L_{ik}^{pre-2011} + \Delta L_{ik}^{S} \tag{1.12}$$

$$\hat{L}_{ik}^{post-2011} = L_{ik}^{pre-2011} + \Delta L_{ik}^{N} \tag{1.13}$$

where ΔL_{ik}^{S} and ΔL_{ik}^{N} are the increases in levee building after 2011 with and without spillovers due to upstream levee construction, respectively. Given a total of 13,414 km (or 44 million ft) levee length in the study area, the external costs of spillovers due to upstream levee building can amount to at least $0.2 billion (Table 1.5).

1.6.4 Cost-benefit analysis

To better understand the magnitude of the external costs due to upstream levee building, I conduct a cost-benefit analysis of levee heightening induced by the Great Mississippi Flood of 2011. I first calculate the benefits change before and after flooding:

$$\Delta B_k = (F_k \delta^F + H_k \delta^H) S_k \Delta \Phi \tag{1.14}$$

where ΔB_k is the benefits (or the expected flood damage reduction by levees) in county k, which equals the sum of total farmland values, total agricultural sales, and total housing values at the county level $F_k \delta^F + H_k \delta^H$, weighted by the proportion of areas designated as the 100-year floodplain in each county S_k and the change in the probability of flooding due to levee

heightening $\Delta\Phi$.[19] In detail, F_k is the sum of total farmland values and total agricultural sales and H_k is the total housing values at the county level, with δ^F and δ^H as the discounting factors to acknowledge the fact that these values are likely lower within the 100-year floodplain in each county.[20]

I then calculate the net benefits of levee building by comparing the benefits to the costs of levee heightening with and without the external costs of spillovers:

$$\Delta NB_k^S = \Delta B_k - \Delta C_k^S \tag{1.15}$$

$$\Delta NB_k^N = \Delta B_k - \Delta C_k^N \tag{1.16}$$

$$\%\,\Delta NB_k = \frac{\Delta NB_k^N - \Delta NB_k^S}{\Delta NB_k^N} \tag{1.17}$$

where ΔNB_k^S and ΔNB_k^N are the net benefits of heightened levees with and without the external costs of spillovers, respectively, and $\%\,\Delta NB_k$ is the percent reduction in net benefits due to the external costs of spillovers.

Table 1.5 summarizes the cost-benefit analysis results. On average, benefits that can be gained by heightened levees are about $3.6 millions, with around $1.8 millions direct costs because of construction and $1.0 million external costs imposed on downstream. For the entire study area, the net benefits of heightened levees with and without considering spillovers can amount to $174 millions and $375 millions, respectively. Taking the external costs of spillovers into account, net benefits of heightened levees can be reduced by about 55%. This result implies that current levee building under highly decentralized planning in the Mississippi is likely not cost effective.

[19] Since the data on actual changes in the probability of flooding due to levee heightening is not available, I use flood probability changes due to levee design specifications (e.g., protecting against a100-year flood pre-2011 versus protecting against a 500-year flood post-2011) as a proximation. Specifically, I use $\Delta\Phi$ = 1% - 0.2% = 0.8% in this study.

[20] The literature shows that agricultural land values and housing values inside the 100-year floodplain can discount by 6.5% and 4.6%, respectively (Struyk, 1971; Beltran et al., 2018). Therefore, I use δ^F= 0.935 and δ^H = 0.954 in this study.

Figure 1.4 further plots the distribution of county-level net benefits of heightened levees due to the Great Mississippi Flood of 2011, where light blue bars denote those without considering the external costs of spillovers and orange bars denote those with spillovers. There are two take-aways from this figure. First, not all levee heightening generated net benefits even without considering spillovers. In fact, about 19% of the counties in the study area experienced net losses after levee heightening due to the low benefits accruing to heightened levees. This suggests that levee heightening may not be an optimal solution to manage increased flooding for some jurisdictions in the Mississippi. Second, some levee heightening can change from having net benefits to having net losses if we take the external costs of spillovers into account, making levee heightening in some areas not economically justified. This implies that without considering external costs due to inter-jurisdictional spillovers, decentralized planning on levee building must be inefficiently high,[21] causing deadweight loss.

Figure 1.5 shows the spatial distribution of the external costs of spillovers and the net benefits of heightened levees with and without such spillovers. As expected, counties with the highest external costs of spillovers are mostly along the Mississippi River since this is the area where levees are extensively constructed. We also observe that counties with the largest net benefits are mostly in the upper Mississippi River Basin, while counties with the largest net losses are mostly in the lower Mississippi River Basin. One possible reason for this spatial disparity in net benefits comes from the wide range of benefits that can be gained from heightened levees across the entire Mississippi River Basin. For areas with highly productive farmland such as Iowa and Illinois, levee heightening will likely generate net benefits. However, for areas such as Louisiana where values that can protected by levees are relatively

[21] Theoretical predictions show that jurisdictions tend to build lower levees if they internalize the negative externalities so as to achieve the social optimum as opposed to their private optimum (see proofs in Appendix A.1).

low (due partially to frequent natural disasters such as hurricanes and floods), levee heightening may not be economically justified.

Given the geo-referenced levee locations and recent advances in remote sensing data that can be used to monitor and measure levee building over time, policymakers can therefore design a spatially explicit compensation scheme in terms of financial transfers across jurisdictions based on the external costs of spillovers that upstream jurisdictions impose on downstream neighbors. Who to compensate and how much to compensate are rather clear given the flow of externalities (i.e., from upstream to downstream) and the extent of externalities (i.e., how much levees are heightened). Since most levees in the Mississippi are locally constructed and maintained, imposing a tax by upper-level authorities such as USACE to address spatial spillovers may not be feasible and a compensation scheme is likely preferred. Alternatively, federal agencies such as FEMA can withhold flood-related recovery assistance from jurisdictions that build excessively high levees. For jurisdictions who receive rehabilitation assistance from USACE to construct levees, maintenance funds can be withheld if local jurisdictions build levees higher than authorized.

1.6.5 The role of decentralization

To further investigate what drives the spillover effects from levee building, I split my data sample into the upper and lower Mississippi River Basins. Table 1.6 shows the spillover effects in the two regions, in addition to the impact of flooding on levee elevation, using the preferred model specification in which both levee and county controls are included. I find that a 1% increase in the upstream levee elevation increased the downstream levee elevation by 0.6% for levees in the upper Mississippi River Basin, suggesting substantial spillovers. However, such spillover effects from upstream levee building become statistically insignificant for levees in the lower Mississippi River Basin.

One possible reason for this regional difference is that levee regulation is fairly decentralized in the upper Mississippi River Basin as opposed to the lower Mississippi River Basin. In specific, the MR&T project in the lower Mississippi valley largely unified the levee system in the area while there is no such level of coordination in the upper Mississippi. In contrast, jurisdictions alongside the upper Mississippi River such as Illinois and Missouri keep building levees and raising heights (Hersher, 2018; Maher, 2019), causing a "levee battle" between jurisdictions.

1.6.6 Robustness checks and falsification tests

First of all, to ensure that the positive impact of flooding on levee elevation is not spurious, I conduct a randomization inference exercise as a placebo test. Randomization inference is designed to assess whether observed outcomes in a given sample are likely to be observed by chance even if treatment had no effect. This tool has been increasingly used in the literature both inside and outside the realm of randomized control trials (see Cohen and Dupas, 2010; Cattaneo et al., 2017; Burlig and Stevens, 2019 for examples).

In detail, I randomly re-assign the 114 flooded counties to treatment 1,000 times, imposing a null hypothesis of no effect of flooding on levee elevation. For each iteration, I estimate every model specification in Table 1.3 and store the estimated $\widehat{\beta_3}$. The results of this test are obtained following Heß (2017) and displayed in Figure A.4. The yellow histograms show the estimated coefficients from these 1,000 random draws, with the dashed vertical red lines denoting the impact of flooding on levee elevation using the actual assignment vector. The randomization inference test also provides exact p-values (i.e., the fraction of permutated coefficients that are larger than the actual coefficient), which are valid for small samples. Across model specifications, I find that the impact of flooding on levee elevation is statistically significant at the 5% level or lower, suggesting that my results are not an artifact of random chance.

Second, to further test if the impact of flooding on levee elevation is robust, Panel A and B in Table A.2 respectively shows the results from estimating the impact of flooding on levee elevation using the data in the upper and lower Mississippi River Basin only, where the upper Mississippi River Basin includes Iowa, Illinois, Missouri and the lower Mississippi River Basin includes Arkansas, Kentucky, Louisiana, Mississippi, and Tennessee. The results from the upper Mississippi River Basin show a consistent estimate of 0.3% across model specifications. With the average pre-flood elevation of flooded levees being 453 ft in the upper Mississippi River Basin, a 0.3% increase in levee elevation indicates approximately a 1.36 ft in levee building. For the lower Mississippi River Basin, I find a 2% increase in levee elevation (approximately 1.90 ft increase in levee building) given the 95 ft average pre-flood elevation of flooded levees in the lower Mississippi River Basin. Despite the differences in landscape, levee regulation, and flood control policies between the upper and lower Mississippi, the positive impact of flooding on levee elevation seems to be consistent.

Third, as discussed in Section 1.5.1, one concern of using the levee elevation data back in the 1980s for the pre-flood period is the occurrence of the Great Flood of 1993 in the upper Mississippi that could potentially bias my estimation. To address this concern, I re-run equation (1.7) using the data sample that excludes any levees whose pre-flood elevations were measured prior to 1993. Table A.3 shows that across model specifications, I find positive point estimates of "Flooded × Post" ranging from 0.4% to 0.8% additional levee elevation as a result of flooding. With the average pre-flood elevation of flooded levees being 252 ft (after dropping the levees whose pre-flood elevations were measured prior to 1993), a 0.8% increase in levee elevation indicates approximately a 2–ft increase in levee building. This number is slightly lower than the one I obtain from using the full sample (i.e., 2.1–3.2 ft). Such a finding may indicate that if the Great Flood of 1993 increased levee elevation in the upper Mississippi, my

estimates using the full sample may be overestimated. However, the impact of flooding on levee elevation remains statistically significant and economically meaningful.

Fourth, the impact of flooding found in this study is also robust to different clustering of standard errors. To control for both spatial and serial correlation, for example, the standard errors are clustered at the county by year level. As shown in Table A.4, the impact of flooding on levee elevation remains statistically significant and is virtually the same as the main results in Table 1.3.

One may suspect that the positive spatial spillovers from upstream levee building in this study may purely come from spatial autocorrelation. If this is the case, the causal inference then becomes unclear. To address this concern, I take advantage of water flow from upstream to downstream, but not the other way round, to conduct a falsification test. Specifically, I estimate the following model specification that aims to capture the impact of downstream neighboring levee elevations on one's own levee elevation:

$$\begin{aligned} log(L_{ikt}) = \beta_1 Flooded_k + \beta_2 Post_t + \beta_3 Flooded_k \times Post_t \\ + \beta_4\, log\left(\overline{L_{k+1,t}}\right) + log(\boldsymbol{X'_{ikt}})\delta + \theta_k + \epsilon_{ikt} \end{aligned} \tag{1.18}$$

where $log\left(\overline{L_{k+1,t}}\right)$ is the logged average elevations of levee(s) in county k's immediate downstream county. The other variables remain as described in equation (1.8). β_4 is the coefficient of additional interest that captures the impact of immediate downstream levee elevation on one's own levee elevation.

Table 1.7 reports the falsification test results. While there is a consistent positive impact of flooding on levee elevation with a point estimate of 0.7%, I find no evidence of spillover effects from downstream neighbors as the coefficient is statistically insignificant across model specifications. This result confirms that the spillover effects from upstream levee building are not spurious or due to the fact of spatial autocorrelation. As additional robustness checks, I further estimate the spillover effects from upstream, cross-river, and upstream-cross-

river levees and the results are shown in Table A.5. While the spillover effect from upstream levee building remains statistically significant, there is no evidence of spillover effects from cross-river or upstream-cross-river neighbors as the coefficients are statistically insignificant across model specifications.

1.6.7 Policy implications

According to the National Committee on Levee Safety, about 100,000 miles of levees stretch across the U.S., most of which are built of earth and covered with grass to protect the estimated 14 million people who live behind them. However, there is currently no single government agency that regulates or monitors all of them, or even sets safety rules. Of all those levees, only about 15% are part of the USACE Levee Safety Program with the vast majority of the nation's levees falling under local or state control. As most levees along the Mississippi River are locally constructed and maintained through county flood control districts or levee districts that are not subject to USACE's levee regulation, the results from this study can provide some significant policy implications.

The positive and substantial spillover effects from upstream levee building imply that as a self-protective behavior, a jurisdiction's levee building can induce extra flood risks and external costs to its downstream neighbors. From an economic standpoint, unregulated or uncoordinated levee building must be inefficiently high. While higher levees are safer for a jurisdiction, the extra height inevitably creates higher flood risks for its downstream jurisdictions. The safety benefits of heightened levees are internal, but the safety costs of these levees are external. Without considering the spillover effects from levee building, levees that are aimed for flood control are likely not cost effective. Government programs that offer financial assistance for flood-control projects such as levees may need to take into account the potential external costs when evaluating the cost-effectiveness of those projects, despite the purpose of self-protection. This is especially important for future program evaluation given that

existing levees need continuous maintenance and that new levees may be constructed in response to more frequent and severe floods in the Mississippi.

In addition, given the decentralized planning in levee construction (especially in the upper Mississippi River Basin), my results also imply that regional coordination on levee building is desired for managing flood damage while mitigating interjurisdictional spillovers. This may require active regulatory involvement by upper-level authorities (e.g., USACE local districts) or the creation of a forum at the regional level (e.g., upper Mississippi River Basin) for negotiation, especially if compensation schemes in terms of financial transfers across jurisdictions are practically and politically feasible. In fact, one of the most important recommendations from the Interagency Floodplain Management Review Committee after the Great Flood of 1993 is that the upper Mississippi River Basin should be treated as a single, integrated system to coordinate flood damage reduction (IFMRC, 1994).

While this paper provides new evidence of inter-jurisdictional spillovers due to decentralized management of natural disasters, one question comes up naturally is what possible policy instruments can be designed and implemented so as to address negative externalities from upstream levee building. When it comes to interjurisdictional spillovers, the central debate in the literature is whether a centralized policy can outperform decentralized planning (Banzhaf and Chupp, 2012; Fell and Kaffine, 2014). While decentralization will likely fail to internalize externalities, central governments often face many hurdles such as high costs of enforcement and lack of accurate monitoring that make the implementation of tailored policies practically impossible. The spatially explicit compensation through financial transfers across jurisdictions offers one possible solution to correct the market failure in a relatively inexpensive way by largely reducing the monitoring costs, thanks to the recent advances in remote sensing data that can monitor and measure levee building over time.

1.7 Concluding Remarks

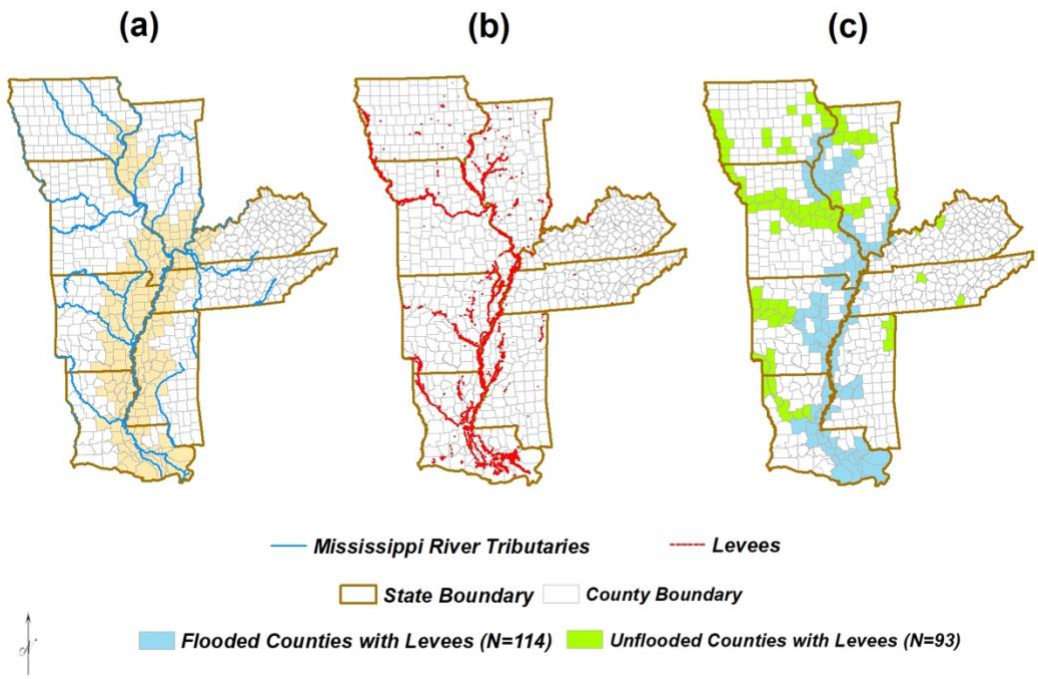
While self-protective actions have been widely acknowledged to induce external risks and costs to other agents (Lakdawalla and Zanjani, 2005; Shafran, 2008; Anderson and Auffhammer, 2014), the existing literature remains scarce on whether these actions can force other agents to increase their own protection through risk spillovers. Levee building along the Mississippi River provides a case in point to study such spillover effects of self-protective measures for flood control. Given the growing evidence of heightened levees in the aftermath of catastrophic flooding together with more frequent flooding in the Mississippi in the past decades, jurisdictions alongside the Mississippi River have expressed their great concerns about the so-called "levee war." Nevertheless, the federal government bears little responsibility for resolving this issue. Such a situation naturally poses the question of whether a centralized policy should be preferred over decentralized planning in dealing with large-scale natural disasters when inter-jurisdictional spillovers are present.

This paper provides both theoretical and empirical evidence of levee building in response to rising flood risks along the Mississippi River, with a particular focus on spillover effects. Using the Great Mississippi Flood of 2011 as a plausibly exogenous shock and applying a DID approach to newly digitized data on levee locations and elevations, I establish three main results. First, consistent with my model predictions, I find levee building in the flooded counties increased by over 2 ft after the 2011 flood. This result is robust across model specifications and is not spurious. Second, spillover effects from upstream levee building are evident and substantial. My results show that a 1% increase in the upstream levee elevation increased the downstream levee elevation by 0.7%. However, such spillover effects become statistically insignificant in areas where levee construction is largely coordinated among jurisdictions and that I find no evidence of spillover effects from cross-river or downstream neighbors' levee building. Third, a back-of-the-envelope calculation shows that the external costs due to

upstream levee building can amount to at least $0.2 billion. A cost-benefit analysis further shows that considering the spillovers from upstream levee building, the net benefits of heightened levees can be reduced by 55%, suggesting that current levee building under decentralized planning is likely not cost effective.

Given the positive and substantial spillover effects from upstream levee building, the results from my study highlight the importance of inter-jurisdictional coordination and call for a more centralized regulation in levee construction for flood control to reduce negative externalities from heightened levees. Future flood-control projects, or management of large-scale natural disasters in general, that require an evaluation of cost-effectiveness need to take into account the potential external costs so that the implementation of those projects can be economically justified.

Figure 1.1: Levees and Flooded Counties during the Great Mississippi Flood of 2011



Note: Map (a) shows the Mississippi River and its tributaries (blue lines) and all the flooded counties in 2011 (shaded yellow areas). Map (b) shows all the recorded levees in the study area (red dots). Map (c) shows the flooded counties with levees (N=114) and unflooded counties with levees (N=93), both of which are used in the empirical analysis.

Figure 1.2: Illustration of Upstream and Downstream Relationships

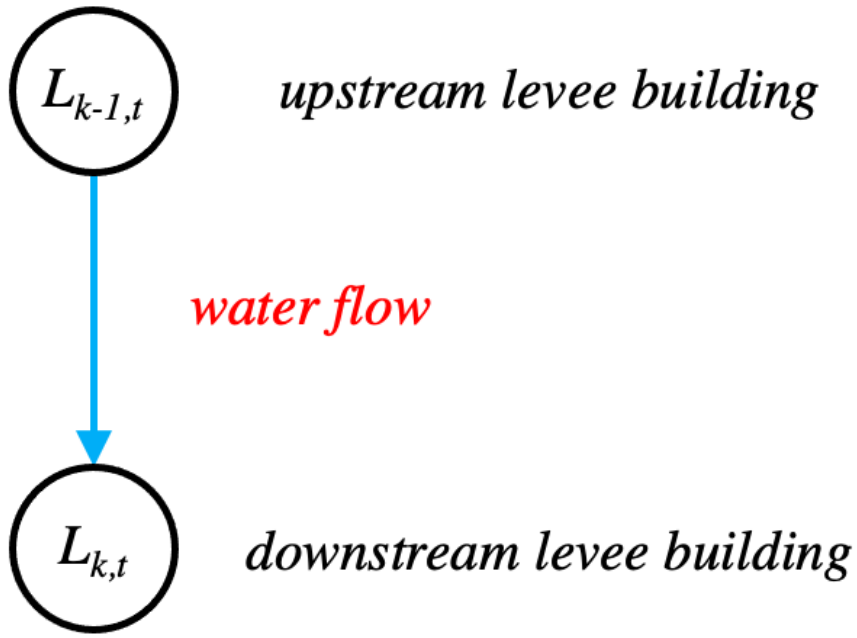
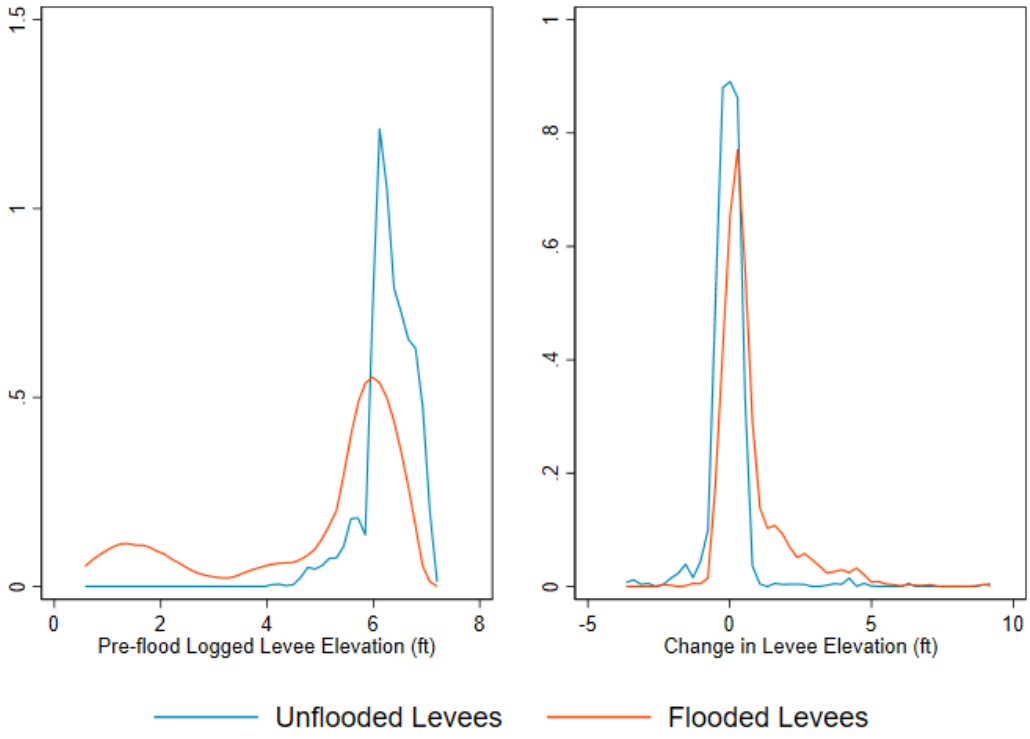
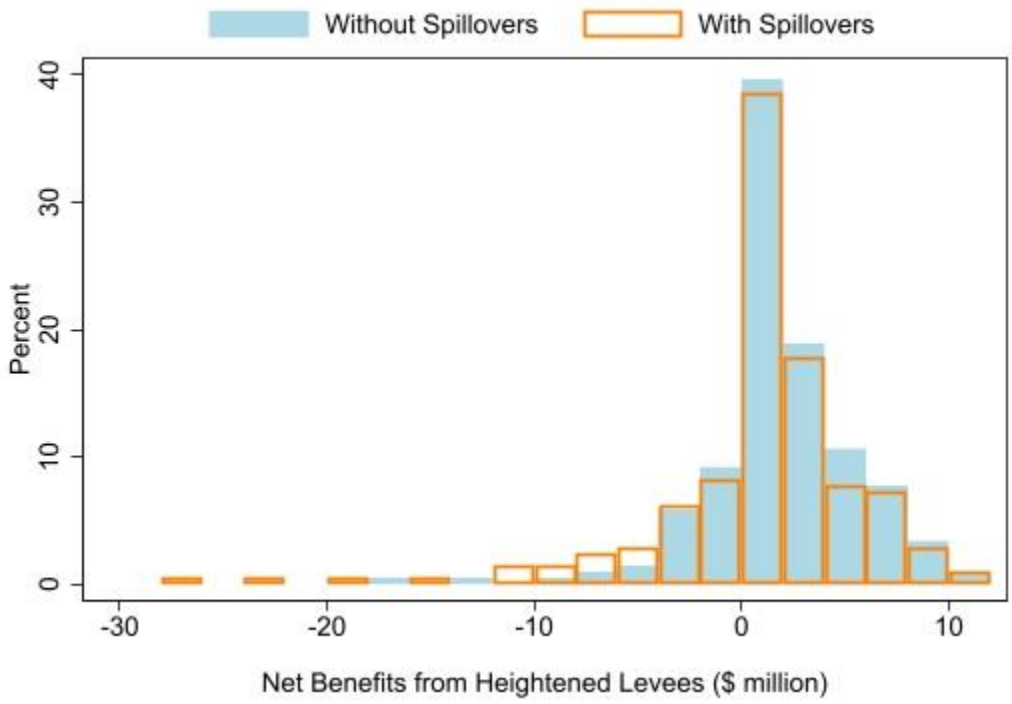


Figure 1.3: Kernel Density Plots of Flooded and Unflooded Levees



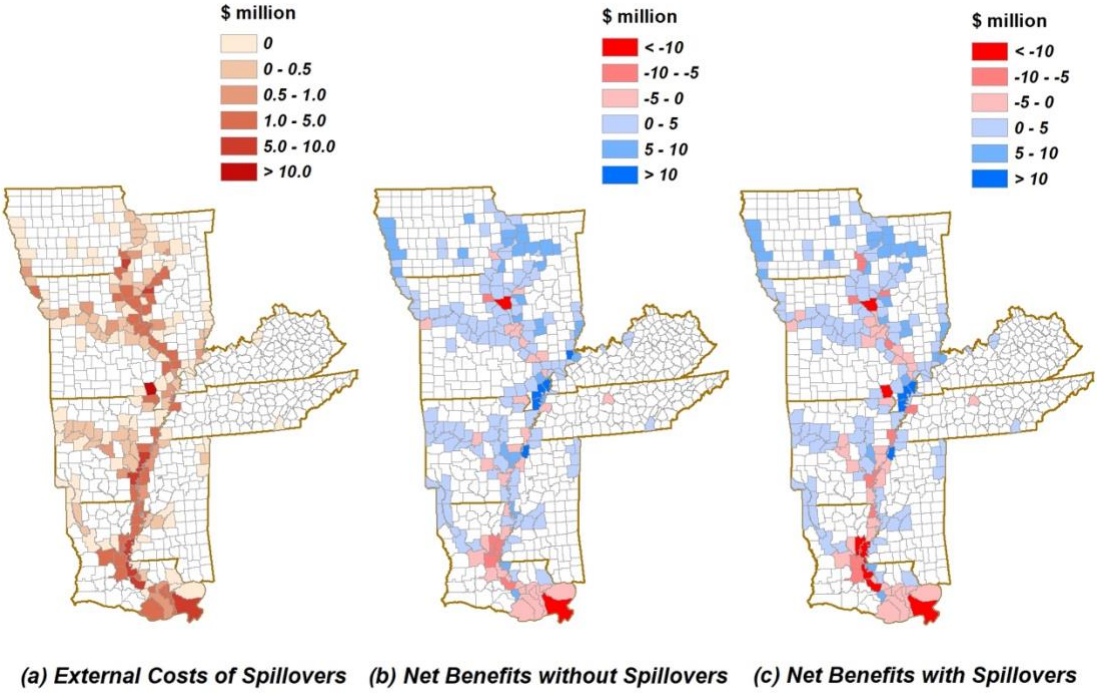
Notes: This figure plots Epanechnikov kernel densities of flooded and unflooded levees, with the left graph comparing the pre-flood logged levee elevation and the right graph comparing the change in levee elevation pre- and post-flood. For the change in levee elevations, the mean (S.D.) of the unflooded distribution is -0.12 (2.44), and the mean (S.D.) of the flooded distribution is 0.72 (1.20). A two-sample t-test with equal variances shows that the difference in mean change in levee elevation between flooded and unflooded levees is statistically significant.

Figure 1.4: Net benefits Comparison with and without Considering Spillovers



Note: This figure plots the distribution of county-level net benefits of heightened levees due to the Great Mississippi Flood in 2011, where light blue bars denote those without considering spillovers and orange bars denote those considering spillovers. Two take-aways from this figure: (1) Not all levee heightening generated net benefits even without considering spillovers; (2) Taking spillovers into account, some levee heightening can change from having net benefits to having net losses.

Figure 1.5: Spatial Distribution of External Costs of Spillovers and Net Benefits with and without Considering Spillovers



Note: The three maps in the figure respectively show the spatial distribution of (a) external costs of spillovers (b) net benefits of heightened levees without considering external costs of spillovers (c) net benefits of heightened levees considering external costs of spillovers. Counties in white are those not in the analysis either because they are not adjacent to any rivers or there are no levees in the county.

Table 1.1: Summary Statistics for Levees and Counties

Variable	Mean	S.D.	Min.	Max.	Obs.
	[1]	[2]	[3]	[4]	[5]
Panel A: Levee characteristics					
Levee elevation (ft)	427.3	259.7	1.560	1,193	2,498
Levee length (km)	10.74	17.14	0.015	144.6	2,498
Number of segments	2.269	4.289	1	89	2,498
100-year floodplain (yes=1)	0.715	0.452	0	1	2,498
Flooded in 2011 (yes=1)	0.542	0.498	0	1	2,498
Panel A: County characteristics					
Farm acreage (1,000 acres)	233.8	114.6	0.111	616.9	2,472
Farmland value ($1,000 per acre)	2.924	1.852	0.866	15.49	2,476
Total agricultural sales ($ million)	91.85	77.30	0.520	464.5	2,472
Median housing value ($ million)	0.109	0.036	0.044	0.209	2,498
Total flood counts	0.815	1.041	0	6	2,498

Notes: Farmland value, total agricultural sales, and median housing value are converted to 2020 $.

Table 1.2: What Predicts Levee Building?

Variable	Logged levee elevation	
	[1]	[2]
Logged farm acreage	0.035	259.7
	(0.026)	(0.017)
Logged farmland value	0.004**	0.019**
	(0.002)	(0.008)
Logged total agricultural sales	0.012**	0.031***
	(0.005)	(0.011)
Logged median housing value	0.682	0.535
	(0.518)	(0.501)
Flood counts	0.003**	-0.001
	(0.001)	(0.001)
Covariates regressed separately	✓	
Levee fixed effects	✓	✓

Notes: In both columns, the dependent variable is the logged levee elevation (ft). In column [1], each row is a separate regression with the variable in that raw as the only control variable in the regression. In column [2], all covariates listed are included in one single regression. In all cases, the specification is a levee fixed-effects regression (equivalent to a first-difference regression in a two-period model). As a result, any time-invariant levee characteristics are not included. Standard errors, clustered at the county level, are in parentheses. *** p<0.01; ** p<0.05

Table 1.3: Impact of Flooding on Levee Elevation

Variable	Logged levee elevation			
	Model I	Model II	Model III	Model IV
Flooded × Post	0.007***	0.007***	0.011***	0.011***
	(0.002)	(0.002)	(0.003)	(0.003)
	[0.002]	[0.002]	[0.003]	[0.003]
Mean of dependent variable	5.520	5.520	5.549	5.549
Number of counties	207	207	206	206
Number of observations	2,498	2,498	2,498	2,498
Levee controls		✓		✓
County controls			✓	✓
County fixed effects	✓	✓	✓	✓
State-by-year fixed effects	✓	✓	✓	✓

Notes: This table shows the results from estimating equation (1.7). The dependent variable is the logged levee elevation (ft). Flooded × Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county level, are in parentheses. Standard errors, adjusted to allow for spatial autocorrelation as modeled in Conley (1999) with a distance cutoff at 50 km, are in brackets. *** $p<0.01$

Table 1.4: Impact of Flooding on Levee Elevation with Spillover Effects

Variable	Logged levee elevation			
	Model I	Model II	Model III	Model IV
Flooded × Post	0.005***	0.004***	0.008***	0.005***
	(0.002)	(0.002)	(0.003)	(0.002)
	[0.002]	[0.002]	[0.003]	[0.002]
log(upstream levee elevation)	0.694***	1.027***	0.737***	01.558***
	(0.264)	(0.553)	(0.235)	(0.658)
	[0.281]	[0.578]	[0.264]	[0.697]
Mean of dependent variable	5.520	5.520	5.549	5.549
Number of counties	207	207	206	206
Number of observations	2,498	2,498	2,498	2,498
Levee controls			✓	✓
County controls			✓	✓
County fixed effects	✓	✓	✓	✓
State-by-year fixed effects	✓	✓	✓	✓
First stage F-statistic		160.6		89.95

Notes: This table shows the results from estimating equation (1.8). The dependent variable is the logged levee elevation (ft). Flooded × Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. log(upstream levee elevation) is the logged average elevation of levee(s) in the county's immediate upstream county. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county level, are in parentheses. Standard errors, adjusted to allow for spatial autocorrelation as modeled in Conley (1999) with a distance cutoff at 50 km, are in brackets. *** p<0.01

Table 1.5: Cost-benefit Analysis of Levee Building

Variable	County-level estimates				Total
	Mean	S.D.	Min.	Max.	
ΔB_k ($ million)	3.578	2.674	-0.910	12.85	740.7
ΔC_k^S ($ million)	2.805	4.600	0	29.93	580.6
ΔC_k^N ($ million)	1.809	2.967	0	23.10	374.6
$\Delta C_k^S - \Delta C_k^N$ ($ million)	0.995	2.445	0	26.64	206.1
ΔNB_k^S ($ million)	0.841	4.949	-26.27	10.34	174.2
ΔNB_k^N ($ million)	1.813	3.656	-17.91	10.94	375.3
% ΔNB_k	–	–	–	–	55%

Notes: ΔB_k is the benefits of heightened levees, ΔC_k^S is the costs of heightened levees with spillovers, ΔC_k^N is the costs of heightened levees without spillovers, $\Delta C_k^S - \Delta C_k^N$ is the external costs of spillovers due to upstream levee building, ΔNB_k^S is the net benefits with external costs of spillovers, ΔNB_k^N is the net benefits without external costs of spillovers, and % ΔNB_k is the percent reduction in net benefits due to external costs of spillovers.

Table 1.6: Impact of Flooding on Levee Elevation with Spillover Effects for Lower and Upper Mississippi River Basins

Variable	Logged levee elevation	
	Lower Mississippi	Upper Mississippi
Flooded × Post	0.022***	0.003***
	(0.006)	(0.001)
	[0.006]	[0.001]
log(upstream levee elevation)	0.003	0.580***
	(0.067)	(0.203)
	[0.082]	[0.215]
Mean of dependent variable	3.935	6.285
Number of counties	92	114
Number of observations	770	1,698
Levee controls	✓	✓
County controls	✓	✓
County fixed effects	✓	✓
State-by-year fixed effects	✓	✓

Notes: This table shows the results from estimating equation (1.8) separately using the data sample in the lower and upper Mississippi River Basin. The dependent variable is the logged levee elevation (ft). Flooded × Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. log(upstream levee elevation) is the logged average elevations of levee(s) in the county's immediate upstream county. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county level, are in parentheses. Standard errors, adjusted to allow for spatial autocorrelation as modeled in Conley (1999) with a distance cutoff at 50 km, are in brackets. *** $p<0.01$

Table 1.7: Falsification Tests on Spillover Effects

Variable	Logged levee elevation			
	Model I	Model II	Model III	Model IV
Flooded × Post	0.007***	0.007***	0.007***	0.007***
	(0.002)	(0.002)	(0.003)	(0.003)
	[0.002]	[0.002]	[0.003]	[0.003]
log(downstream levee elevation)	0.264	0.264	0.427	0.427
	(0.318)	(0.318)	(0.278)	(0.278)
	[0.331]	[0.331]	[0.290]	[0.290]
Mean of dependent variable	5.520	5.520	5.549	5.549
Number of counties	207	207	206	206
Number of observations	2,498	2,498	2,498	2,498
Levee controls		✓		✓
County controls			✓	✓
County fixed effects	✓	✓	✓	✓
State-by-year fixed effects	✓	✓	✓	✓

Notes: This table shows the results from estimating equation (1.16). The dependent variable is the logged levee elevation (ft). Flooded × Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. log(downstream levee elevation) is the logged average elevations of levee(s) in the county's immediate downstream county. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county level, are in parentheses. Standard errors, adjusted to allow for spatial autocorrelation as modeled in Conley (1999) with a distance cutoff at 50 km, are in brackets. *** $p<0.01$

Chapter 2: Evaluating Payments for Ecosystem Services Programs: Serial Non-participation, Time, and Risk

2.1 Introduction

Payments for ecosystem services (PES) programs have been widely used in both developing and developed countries to address negative environmental externalities from agriculture. In most settings, willing landowners voluntarily enroll in PES programs, through which they receive a variety of payments (e.g., upfront signing bonus, recurring annual payment) in exchange for practice adoption on farmland or changes in land use (Wunder et al, 2008; Alix-Garcia and Wolff, 2014; Ribaudo and Shortle, 2019). Since policymakers are looking at expanding enrollment in PES programs, stated preference (SP) surveys offer a way to design future programs and have been increasingly used to evaluate landowners' willingness to accept (WTA) for program participation (see Villanueva et al., 2017 for a literature review). However, the issue of serial non-participation (SNP) occurs when significant shares of respondents state that they are not interested in any of the offered programs in the survey. The presence of SNP raises the question of whether it is possible to infer what kinds of program features might attract serial non-participants: Can one extrapolate from the choices that participants make? Or are the preferences of serial non-participants so different from those of participants that no inference is possible? This problem is analogous to estimating labor supply functions in the presence of unemployment (Ham, 1982; Blundell et al., 1987). Individuals who record zero hours of work can consist of those who do not want to work at their market wage or those who want to work but remain unemployed. Ignoring or not accounting for unemployed workers can bias parameters in labor supply functions.

Accounting for SNP in the WTA context is of significant policy relevance and especially important due to a fundamental difference in welfare measures between the supply-

and demand-side of the market. On the demand-side of the market, WTP is well defined for serial non-participants and can be arguably treated as zero, making it possible to estimate the average WTP for the population (von Haefen et al., 2005). Whereas on the supply side of the market, treating WTA as either zero or infinity for serial non-participants is not conceptually defensible, creating the issue of whether it is possible to infer WTA for serial non-participants. This issue further points to the question of how high incentive payments would have to be to increase program participation given that serial non-participants' unobserved reservation prices may lie outside the reasonable range offered by government programs.

Empirically, three WTP studies focusing on the demand side of the market formally account for SNP by introducing a different data generating process using hurdle variants of discrete choice models (von Haefen et al., 2005; Burton and Rigby, 2009; Chen et al., 2020). The standard discrete choice model such as a logit model implicitly assumes that the utility functions of serial non-participants are approximately the same as those of participants. By contrast, hurdle variants of discrete choice models explicitly predict SNP using a probabilistic model and allow for serial non-participants, who may not be "playing the game," to behave fundamentally differently from participants. Accounting for SNP using hurdle variants of discrete choice models has been found to yield significant improvements in statistical fit and quantitatively different WTP estimates than standard discrete choice models (von Haefen et al., 2005).

Regarding the supply side of the market, Villanueva et al. (2017) are the first to examine how SNP affects landowners' WTA for PES program participation. Instead of formally accounting for SNP, they use subsamples by dropping identified non-participants for WTA estimation. While this ad-hoc approach sheds some light on whether the inclusion or exclusion of serial non-participants in the analysis can influence WTA estimates, little is known about what factors contribute to SNP. More importantly, serial non-participants' preferences

for proposed programs cannot be inferred from participants' choice behaviors for any post-estimation policy simulations because serial non-participants are excluded from the empirical analysis.

One important feature about PES programs is that they often feature long-term contracts between the government and individual landowners. Contract theory has long acknowledged the role of risk and time preferences in individual decision-making when it comes to long-term contracts between the principal and agents (Stiglitz, 1974; Allen and Lueck, 1995). Since policymakers often use increased upfront signing bonuses and/or annual payments to boost program enrollment, how individuals perceive discounting and risk over the contract period can affect how effective are these two incentives. This is especially true when PES programs provide guaranteed payments that may serve as an attractive means of insurance against income risk. So far, empirical PES studies tend to abstract away from the role of risk or time preferences and focus mostly on how various payment incentives and transaction costs impact contracting (Suter et al., 2008; Peterson et al., 2015; McCann and Claassen, 2016; Palm-Forster et al., 2016; Johnston et al., 2022). Nevertheless, both discounting and risk have been shown to play a significant role in similar cases such as land rental contracts (Fischer and Wollni, 2018; Lloyd-Smith et al., 2021) and agricultural technology adoption (Brick and Visser, 2015; Michler and Wu, 2020).

This paper focuses on landowners' enrollment in long-term PES programs using data from a discrete choice experiment (DCE) implemented in a farmer survey. We make our first contribution to the literature by adopting hurdle variants of discrete choice models to account for SNP in the WTA context and explore the differences between hurdle models and the standard discrete choice model that ignores SNP. This departs from the existing SP studies on estimating WTA for providing ecosystem goods and services that use subsamples only by excluding identified serial non-participants from estimation (Villanueva et al., 2017). While

the standard discrete choice model assumes serial non-participants' utility functions to be approximately the same as those of participants, hurdle variants of discrete choice models allow for serial non-participants to behave fundamentally differently from participants. Assumptions about choice behaviors of serial non-participants in alternative models can lead to quantitatively different welfare measures and qualitatively different policy implications.

We also fill the gap in the literature by examining the role of time and risk preferences in landowners' enrollment in long-term PES programs. We incorporate individuals' risk preferences into our theoretical and empirical framework that jointly estimates individuals' time and risk preferences in PES program enrollment. While implicit discount rates associated with intertemporal decisions have been estimated in a growing number of demand-side SP studies (see Howard et al., 2021 and Vasquez-Lavín et al., 2021 for literature reviews), to our best knowledge, the acknowledgment of risk preferences in PES contracts remains scarce, although intertemporal decision-making on long-term conservation practices is inherently risky. Unlike most previous studies that estimate the implicit discount rates with WTP for various ecosystem goods and services (e.g., Bond et al., 2009; Lew, 2018; Grammatikopoulou et al., 2020; West et al., 2021), we focus on the supply side of the market through the provision of ecosystem goods and services. Specifically, we examine how time and risk preferences affect landowners' enrollment decisions.

We designed a survey for rural landowners to examine their willingness to enroll in government payment programs that incentivize the adoption of riparian buffers on working farmland in Maryland. Program attributes were specified to reflect those for riparian buffer adoption under the Conservation Reserve Enhancement Program (CREP), including the buffer type, levels of one-time signing bonus and recurring annual payment, and contract length. Our survey sample includes 8,923 rural landowners in Maryland drawn from the complete spatially explicit parcel-level tax assessor using an address-based sampling approach. Interestingly, 46%

of our survey respondents chose the status quo option in all four DCE questions, suggesting that serial non-participation is a non-trivial issue to address in WTA estimation. To identify implicit discount rates, we followed the SP literature (Lew, 2018; Grammatikopoulou et al., 2020; West et al., 2021) by varying the time when the landowner would receive the one-time signing bonus. To measure landowner's risk aversion parameter, we complemented our survey with risk preference elicitation questions adapted from Falk et al. (2016).

We estimate three discrete choice model specifications including a standard logit model as well as single- and double-hurdle logit models. The key distinction among the three alternative models is the extent to which the preferences of serial non-participants can be extrapolated from the preferences of participants who chose at least one proposed program. The standard logit model assumes that serial non-participants and participants share the same utility functions so that the preferences of serial non-participants can be linearly extrapolated from the observed preferences of participants. Allowing serial non-participants' preferences to be fundamentally different from those of participants, the single-hurdle logit (SHL) model assumes that no such extrapolation is possible while the double-hurdle logit (DHL) model assumes that some extrapolation is possible.

We establish several important results in this paper. First, whether SNP is accounted for in the econometric model is crucial for the empirical analysis. Our post-estimation calculations of the probability of enrollment, average WTA for program participation, and implicit discount rates indicate that different assumptions about serial non-participants embedded in alternative models can make a significant difference in inferences about policy. Our simulated probabilities of enrollment show that ceiling on PES program participation becomes evident when SNP is explicitly modeled using a hurdle framework that allows serial non-participants to behave differently from participants. This result points to an inherent limitation in most voluntary programs that aim to provide ecosystem goods and services,

suggesting that simply making programs more lucrative may not raise that ceiling very much. Failing to account for SNP can also lead to quantitatively different WTA estimates for program participation. Second, we find that risk-averse landowners are less likely to enroll in proposed programs, indicating that they perceive program participation to increase income risk. This result suggests that guaranteed government payments may not offset risks and uncertainty to farm production from potential changes in existing farm operations due to riparian buffer adoption. Third, our estimates of landowners' implicit discount rates are quite low with weak evidence of hyperbolic discounting. The estimated supply-side discount rates are substantially lower than those estimated for the demand of environmental amenities but are consistent with market interest rates such as loans offered by the U.S. Department of Agriculture.

2.2 Theoretical Model

We develop a theoretical model to explain landowners' program enrollment decisions and the occurrence of SNP. The model formalizes landowners' intertemporal decision making in face of a specific riparian buffer program with specified contract terms under a random utility framework. Assume that the utility of landowner i enrolling in program j depends on program attributes and the present value of the stream of payments from each program and farm income. If utility is additively separable, the certainty equivalent of the expected discounted utility of enrolling in program j for a risk-averse landowner i is:

$$\begin{aligned} EV_{ij}^{1} &= \sum_{t=1}^{T_j} D(t) Eu\left(\boldsymbol{X_i}, z_j, y_{it}, p_{jt}, s_{jt}\right) \\ &= a^1(\boldsymbol{X_i}) + z_j + b\left[\sum_{t=1}^{T_j} D(t) p_{jt} + \sum_{t=1}^{T_j} D(t)\left(\bar{y}_{it}^{1} - \omega\sigma_i^1\phi_i\right)\right] + bD(d)s_j + \epsilon_{ij}^{1} \end{aligned} \tag{2.1}$$

where superscript 1 denotes the state with program enrollment, $a^1(\boldsymbol{X_i})$ is the background utility of landowner i under program j with $\boldsymbol{X_i}$ denoting observable landowner and parcel characteristics that vary across landowners, z_j is the utility of a specified buffer type (e.g.,

forest, grass), s_j is the one-time signing bonus, p_{jt} is the annual payment recurring for T_j years, $\bar{y}_{it}^1$ is the mean value of annual farm income, σ_i^1 is the variance of farm income, ϕ_i is the coefficient of risk aversion, $D(t)$ and $D(d)$ are discounting factors respectively for the annual income and the one-time signing bonus with d the number of years of delay in receiving the one-time signing bonus, b is the marginal utility of income, and ϵ_{ij}^1 is an error term denoting any unobserved utility. Correspondingly, the certainty equivalent of the expected discounted utility with no program enrollment is:

$$EV_{ij}^0 = \sum_{t=1}^{T_j} D(t) Eu(\boldsymbol{X_i}, y_{it}) = a^0(\boldsymbol{X_i}) + b\left[\sum_{t=1}^{T_j} D(t)\left(\bar{y}_{it}^0 - \omega\sigma_i^0\phi_i\right)\right] + \epsilon_{ij}^0 \qquad (2.2)$$

where superscript 0 denotes the state without program enrollment. Landowner i is willing to take a contract under program j if her expected discounted utility with enrollment is greater than that without enrollment. The probability of landowner i enrolling in program j then equals the probability that program j provides landowner i with a greater expected discounted utility than no enrollment:

$$P_{ij} = f\left(EV_{ij}^1 > EV_{ij}^0\right) \qquad (2.3)$$

where the functional form for $f(\cdot)$ is determined in the subsequent empirical estimation.

In a repeated discrete choice framework, landowners may always choose no program enrollment because their status quo utility is so high due to various costs of program participation (e.g., transaction costs, conversion costs of land uses, costs associated with increased pest presence due to riparian buffer adoption) that none of the offered programs in the survey provided enough monetary benefits. Based on landowners' observed choice outcomes, the probability of landowner i being a serial non-participant can be modeled as:

$$\Pr(SNP_i) = g\left(EV_{ij}^1 < EV_{ij}^0, \forall\, j \in J\right) = g(\boldsymbol{X_i}, \boldsymbol{Z_i}, \varepsilon_i) \qquad (2.4)$$

where $SNP_i = 1$ if landowner i chooses the status quo option in all given DCE questions, J is the total number of DCE questions each landowner faced in the survey, and $\boldsymbol{Z_i}$ denotes some additional landowner characteristics to explain SNP. The functional form for $g(\cdot)$ is determined in the subsequent empirical estimation.

2.3 Econometric Model and Estimation

Our econometric specification is grounded in the theoretical model above. The DCE presented each respondent with a set of four binary (yes/no) choice situations, in which they were asked whether they would enroll, compared to a status quo of no enrollment. Each landowner considers enrollment choices over four distinct programs, with the observable component of utility from each program determined by a set of attributes that vary over programs. Each program is considered independently relative to the status quo. To reflect how landowners discount payments over time, we include the interaction terms of the one-time signing bonus with different time delays following Viscusi et al. (2008). Given the three discrete contract length options in the program, we additionally interact the annual payment with contract length to capture landowners' preferences for the duration of program enrollment. Building upon the theoretical model and the program attributes in our DCE design, landowner i's indirect utility of program j is:

$$\begin{aligned} V_{ij} = \alpha_i + \boldsymbol{X_i}\lambda + \delta z_{ij} + \beta_0 s_{ij} + \beta_2 s_{ij} \times D_2 + \beta_5 s_{ij} \times D_5 \\ + \gamma_5 p_{ij} + \gamma_{10} p_{ij} \times C_{10} + \gamma_{15} p_{ij} \times C_{15} + \epsilon_{ij} \end{aligned} \tag{2.5}$$

where α_i is a constant term denoting the status quo utility and $\boldsymbol{X_i}$ is a vector of landowner and parcel characteristics to allow for heterogeneity in the status quo utility and capture the possible differentiated utility between buffer adoption and no adoption due to transaction costs, opportunity costs, and other factors. Specifically, we include the parcel-specific soil rental rate

to proxy for any farm income loss due to riparian buffer adoption.[22] The risk aversion dummy is used to capture the difference in the perceived riskiness of enrollment in the riparian buffer program compared to the status quo of no enrollment. Additionally, z_{ij} is a dummy variable indicating forest buffers (grass buffers as the baseline), D_2 and D_5 are dummy variables respectively indicating 2 and 5 years of delay in receiving the one-time signing bonus (no delay as the baseline), and C_{10} and C_{15} are dummy variables respectively indicating a 10-year and 15-year contract length (a 5-year contract as the baseline).

Following the general formulation of a logit model and if the error term ϵ_{ij} is an *i.i.d.* draw from the type I extreme value distribution, the probability that landowner i chooses program j in any given DCE question is written as:

$$P_{ij} = \frac{exp\ (V_{ij})}{1 + exp\ (V_{ij})} \tag{2.6}$$

The log-likelihood function is then:

$$ln\mathcal{L} = \log\left[\prod_i \prod_j (P_{ij})^{y_{ij}}\right] \tag{2.7}$$

where $y_{ij} = 1$ if landowner i chooses program j in any given DCE question and zero otherwise. Equation (2.7) shows the log-likelihood function for a logit model in which the utility functions of serial non-participants are approximately the same as those of participants.

To account for SNP and to allow serial non-participants' choice behaviors to differ from participants, we introduce a probability model as a "hurdle" framework that specifies landowners who always choose the status quo option as serial non-participants following von Haefen et al. (2005):

$$Pr(SNP_i) = \Phi(\boldsymbol{X_i}\eta, \boldsymbol{Z_i}\phi) \tag{2.8}$$

[22] To install riparian buffers, farmland needs to be taken out of production. The CREP riparian buffer program requires cropland with an established cropping history in pasture, commodity crops, or hay to be eligible for enrollment.

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. While landowner and parcel characteristics in $\boldsymbol{X_i}$ can partially explain SNP, we particularly include two variables denoting landowners' attitudes toward the government and government programs in $\boldsymbol{Z_i}$ to capture landowners whose reservation prices lie outside the reasonable range offered by government programs as our exclusion restrictions.[23] We argue that landowners who have dislikes for the government accessing landowners' private property or taxes used for government farm support programs will have high status quo utility and thus require compensation outside the reasonable payment range to even contemplate participating. However, attitudes about the government monitoring activities or farm support programs will have little effect on the specific program choices for landowners who do not harbor such feelings.[24]

As proposed in von Haefen et al. (2005), two types of "hurdle" frameworks can be introduced using the probability model shown in equation (2.8), depending on the assumptions made. In a single-hurdle framework, the probability model solely explains serial non-participants who choose the status quo in all given DCE questions because their reservation prices lie outside the reasonable range offered by government programs. By contrast, a double-hurdle framework assumes that the probability model augments the probability of SNP by allowing for the possibility that some utility-maximizing landowners, whose preferences are similar to participants, happen to choose the status quo option in all proposed programs given in the survey. The log-likelihood function for the SHL model is then:

[23] Note that the two variables denoting landowners' attitudes toward the government and government programs can also be included in the logit model. In this case, they mainly have the effect of increasing the WTA for serial non-participants whose preferences are the same as participants and thus decrease the likelihood of enrollment.

[24] Our exclusion restrictions are plausible as we find that landowners' attitudes toward the government and government programs do not seem to affect their preferences for program enrollment if they are participants (see Table 2.3). However, these attitudes can significantly explain SNP (see Table 2.4).

$$ln\mathcal{L} = \sum_i \log\left[-\Phi(\boldsymbol{X_i}\eta, \boldsymbol{Z_i}\phi)(1 - SNP_i)\prod_j (P_{ij})^{y_{ij}} + \Phi(\boldsymbol{X_i}\eta, \boldsymbol{Z_i}\phi)SNP_i\right] \quad (2.9)$$

The log-likelihood function for the DHL model is then:

$$n\mathcal{L} = \sum_i \log\left[-\Phi(\boldsymbol{X_i}\eta, \boldsymbol{Z_i}\phi)\prod_j (P_{ij})^{y_{ij}} + \Phi(\boldsymbol{X_i}\eta, \boldsymbol{Z_i}\phi)SNP_i\right] \quad (2.10)$$

Equations (2.9) and (2.10) illustrate the main difference between SHL and DHL models. While $\Phi(\boldsymbol{X_i}\eta, \boldsymbol{Z_i}\phi)$ is assumed to only explain landowners whose reservation prices lie outside the reasonable range offered by government programs in the SHL framework., it additionally explains landowners whose preferences are similar to participants in the DHL framework. Notably, in contrast to a standard logit model that assumes serial non-participants to share similar preferences with participants, both SHL and DHL models introduce a different data generating process through the first-stage hurdle using a probability model to explain landowners whose reservation prices lie outside the reasonable range offered by government programs at least in part.

Note that although the logit model can allow for heterogeneity in the expected utility by including landowner and parcel characteristics to potentially distinguish between participants and serial non-participants, an implicit assumption is that the utility functions of serial non-participants are approximately the same as those of participants. The two hurdle models characterize heterogeneity differently by assuming that there is a qualitative difference in utility functions between serial non-participants and participants. Such a difference becomes important in making policy inferences such as estimating potential program uptake, WTA, etc. The logit model assumes that program uptake and WTA of serial non-participants can be estimated from the parameters of the logit model. By contrast, the two hurdle models assume that uptake has a ceiling defined by the hurdle model and that WTA can only be estimated for

participants as there is insufficient information (i.e., no variation in choice outcomes) to identify the preferences of (some) serial non-participants.

2.3.1 Policy simulation on probabilities of enrollment

To provide a better understanding of how landowners respond to financial incentives when they decide to enroll in programs, we conduct a series of policy simulations to estimate supply curves using our regression results. The predicted probability of enrolling in a specific riparian buffer program can be computed when monetary incentives (signing bonus, annual payments), features of those monetary incentives (signing bonus differentiated by payment delay, annual payment differentiated by contract length), and buffer type (grass, forest) differ from scenario to scenario. The general calculation formula is as follows:

$$\Pr(enroll) = \underbrace{\Pr(participant)}_{hurdle\ equation} \times \underbrace{\Pr(enroll \mid participant)}_{choice\ equation} \tag{2.11}$$

where $\Pr(participant) = 1 - \Pr(SNP)$ is the predicted probability of being a participant to be estimated from equation (2.8) and $\Pr(enroll \mid participant)$ is the predicted probability of enrolling in a specified riparian buffer program conditional on being a participant to be estimated from equation (2.6). The distinction between the logit and the two hurdle models is the assumptions about serial non-participants inherent in alternative estimation methods. The logit model assumes that all serial non-participants' preferences can be extrapolated from observed choice behaviors of participants with $\Pr(participant) = 1$. The SHL model assumes that all serial non-participants are landowners whose reservation prices lie outside the reasonable range offered by government programs, while the DHL model assumes serial non-participants to be a mix of both.

2.3.2 Willingness to accept estimation

One important goal of this study is to estimate landowners' WTA for a contract under which they are willing to install a specified riparian buffer for certain years. Following Hanemann (1984), we derive WTA estimates based on our econometric specification using the

DCE data. Following our theoretical model, we can estimate the annualized WTA (i.e., a constant annual payment)[25] as the compensation that would make the landowner indifferent between enrolling and not enrolling in a specific program conditional on program participation:

$$\begin{aligned} & a^1(\boldsymbol{X_i}) + z_j + b\sum_{t=1}^{T_j} D(t)\left(\bar{y}_{it}^1 - \omega\sigma_i^1\phi_i + WTA_{jt}\right) + \epsilon_{ij}^1 \\ & = a^0(\boldsymbol{X_i}) + b\sum_{t=1}^{T_j} D(t)\left(\bar{y}_{it}^0 - \omega\sigma_i^0\phi_i\right) + \epsilon_{ij}^0 \end{aligned} \tag{2.12}$$

The average WTA conditional on program participation is thus:

$$\begin{aligned} & E\left[WTA_{jt}|participation\right] \\ & = \frac{a^0(\boldsymbol{X_i}) - a^1(\boldsymbol{X_i}) + b\sum_{t=1}^{T_j} D(t)\left[\bar{y}_{it}^0 - \bar{y}_{it}^1 - \omega\left(\sigma_i^0 - \sigma_i^1\right)\phi_i\right] - z_j}{b\sum_{t=1}^{T_j} D(t)} \end{aligned} \tag{2.13}$$

where the difference in error terms is mean zero. Since WTA estimates are specified in annualized terms, equation (2.13) indicates that we obtain a separate WTA estimate for each contract length T_j. For example, the annualized WTA for a forest buffer under a 5-year, 10-year, and 15-year contract is respectively calculated as follows:

$$\begin{aligned} E\left[WTA_{jt}^5|participation\right] &= \frac{-\widehat{\alpha_\iota} - \boldsymbol{X_i}\hat{\lambda} - \hat{\delta}}{\widehat{\gamma_5}} \\ E\left[WTA_{jt}^{10}|participation\right] &= \frac{-\widehat{\alpha_\iota} - \boldsymbol{X_i}\hat{\lambda} - \hat{\delta}}{\widehat{\gamma_5} + \widehat{\gamma_{10}}} \\ E\left[WTA_{jt}^{15}|participation\right] &= \frac{-\widehat{\alpha_\iota} - \boldsymbol{X_i}\hat{\lambda} - \hat{\delta}}{\widehat{\gamma_5} + \widehat{\gamma_{15}}} \end{aligned} \tag{2.14}$$

where $\widehat{\gamma_5}$, $\widehat{\gamma_5} + \widehat{\gamma_{10}}$, and $\widehat{\gamma_5} + \widehat{\gamma_{15}}$ are the marginal utility of income respectively for a 5-year, 10-year, and 15-year contract, the difference in the background utility and farm income is captured by $\widehat{\alpha_\iota}$ and $\boldsymbol{X_i}\hat{\lambda}$, and the utility of forest buffers (grass buffers as the baseline) is

[25] Note that the one-time signing bonus is not included in equation (2.12). As a result, our annualized WTA measure is essentially the sum of the recurring annual payment and the one-time signing bonus annualized by contract length.

captured by $\hat{\delta}$ according to our econometric specification in equation (2.5). Specifically, the difference in the mean value of annual farm income, $b\sum_{t=1}^{T_j} D(t)\left(\bar{y}_{it}^0 - \bar{y}_{it}^1\right)$, is proxied by the parcel-specific soil rental rate and the difference in the variance of farm income, $b\sum_{t=1}^{T_j} D(t)\omega\left(\sigma_i^0 - \sigma_i^1\right)\phi_i$, is proxied by the risk aversion parameter. The coefficient estimates for both variables are expected to be negative. A higher annual farm income means a higher opportunity cost of replacing farmland with riparian buffers, which decreases the likelihood of program enrollment. Since the crop return is highly variable and riparian buffers could change existing farm operations with uncertainty that induces larger variance in farm income, risk averse landowners are therefore less likely to enroll, despite a fixed guaranteed annual payment provided in the program.

Equation (2.14) can be used to calculate annualized WTA estimates for different programs (varied by the buffer type) and different types of landowners (varied by landowner and parcel characteristics). As WTA estimates are conditional on program participation, the coefficient estimates used to calculate WTA are from the choice equation (i.e., equation 2.5). Given that different landowners contribute to the choice portion of the log-likelihood function due to different assumptions about serial non-participants inherent in the logit, SHL, and DHL models, we report the average WTA estimates for the subsample of participants across three models for comparisons.

2.3.3 Discount rate estimation

Like several previous studies (Lew, 2018; Grammatikopoulou et al., 2020; West et al., 2021), our experimental design varying the time delay in receiving the one-time signing bonus allows us to identify respondents' average implicit discount rate by using the ratio of the coefficients of the signing bonus with and without time delays. The standard discounting formulation in the discrete time measure is defined as:

$$D(d) = \frac{1}{(1 + r_d)^d} \tag{2.15}$$

where $D(d)$ is the discounting factor for year d and r_d is the discount rate.

To estimate the discount rate, we note that the present utility of the one-time signing bonus that is delayed by d years is equivalent to the utility of the same payment in the year it occurs, then discounted back to the present value. Following this logic, we can derive a discount rate by setting these two equivalent terms equal and solving for the discount rate. Specifically, the coefficient on no delay in receiving the signing bonus, β_0, captures the current utility of increasing the signing bonus (i.e., today). β_2 and β_5 respectively capture the current disutility of increasing the signing bonus if receiving the signing bonus is delayed for 2 and 5 years. Therefore, the current utility of increasing the signing bonus that is delayed for 2 and 5 years is $\beta_0 + \beta_2$ and $\beta_0 + \beta_5$, respectively. We can then use the two following equations to solve for the two discount rates r_2 and r_5:

$$\beta_0 + \beta_2 = \frac{\beta_0}{(1 + r_2)^2}, \qquad \beta_0 + \beta_5 = \frac{\beta_0}{(1 + r_5)^5} \tag{2.16}$$

This approach is comparable to many other SP studies that use the ex-post estimation approach to estimating the implicit discount rate (see Howard et al., 2021 and Vasquez-Lavín et al., 2021 for a list of studies).

2.4 Survey Design and Implementation

2.4.1 Discrete choice experiment

We design a farmer survey in which a DCE is used to elicit landowners' preferences for government payment programs that vary over supported buffer type, one-time signing bonus, recurring annual payment, and contract length required for maintaining the buffer. The characterization of program attributes was designed to reflect the riparian buffer cost-share

program being implemented in Maryland.[26] The survey was developed in consultation with regional experts and riparian buffer program administrators from the USDA's Farm Services Agency Maryland State Office, Maryland Department of Agriculture, Maryland Department of Natural Resources, and Chesapeake Bay Foundation. Additional input was gained from the University of Maryland Extension agents who work closely with Maryland landowners on agricultural conservation practices. Multiple expert focus groups were convened to develop, test, and improve survey designs. Individual pretests were also conducted among Maryland landowners (both who currently have and do not have riparian buffers on their land) to ensure clarity and comprehension of survey language and graphics.

To develop DCE questions, we followed the best practice guidelines in Johnston et al. (2017) and gave particular attention to the information presented in the questions and the selection of program attributes. To enhance consequentiality, before landowners answered DCE questions they were informed that a summary report of anonymous results would be shared with government agencies and farmer stakeholder groups to help them understand program features and thus inform potential decisions on future program design changes. Following the suggestions by recent studies to achieve incentive compatibility (Johnston et al., 2017; Weng et al., 2021), we use a two-alternative choice format (i.e., the proposed program against the status quo option) instead of a three-alternative choice format. We acknowledge that WTA elicitation may not be framed in a way that satisfies all criteria for incentive compatibility even when consequentiality holds (Lloyd-Smith and Adamowicz, 2018). Nonetheless, a WTA framework is conceptually appropriate when one seeks to predict decisions on voluntary program participation. Considering the guidance in Johnston et al.

[26] For more information on Maryland's conservation buffer initiative, we refer readers to https://mda.maryland.gov/resource_conservation/Pages/conservation-buffer-initiative.aspx.

(2017), we frame the WTA elicitation as it is theoretically and conceptually correct for our decision context, although the attendant lack of incentive compatibility is acknowledged.

Each landowner received four binary (yes/no) choice scenarios, in which they were asked whether to enroll in a proposed program and thus adopt a riparian buffer on their land. Landowners were instructed to consider each DCE question as an independent, non-additive choice. Each proposed program included four attributes characterizing the supported buffer type, one-time signing bonus, recurring annual payment, and contract length (see Table 1 for a complete list of program attributes and attribute levels). Reflecting the current structure of the CREP cost-share program for riparian buffers in Maryland (i.e., 100% cost-share rate), installation and maintenance costs are fully covered by each program, regardless of the buffer type offered in the program. Like most other agricultural conservation cost-share programs, our proposed programs did not mandate the size of riparian buffers that would be installed. However, a minimum buffer width of 35 feet was stipulated per the CREP buffer requirements, and landowners were further asked about the buffer length if they would install under the given program assuming a 35-feet buffer width.

The experimental design was developed using a D-efficiency criterion for the choice model covariance matrix (Scarpa and Rose, 2008), which minimizes the variance of a CL model where the parameters are all assumed to equal zero. The resulting design generated a total of 48 choice scenarios blocked into 12 groups, each with four choice scenarios. At least one of the four choice scenarios featured either an annual payment or a signing bonus well outside the range currently being offered for the riparian buffer program in Maryland. Landowners randomly received one of the 12 groups when facing DCE questions. Each DCE question presented a proposed program against the status quo option (i.e., no program enrollment). Figure 2.1 provides an example of a DCE question, which describes a program that would require the installation of a grass buffer in exchange for a one-time signing bonus

of $500/acre and a recurring annual payment of $250/acre for 10 years. Additionally, installation and maintenance costs are fully covered by the program and the landowner would receive the one-time signing bonus upon program enrollment.

To identify the implicit discount rate, we specified when each landowner would receive the one-time signing bonus in our experimental design following recent SP studies that delay the timing of payment (Lew, 2018; Grammatikopoulou et al., 2020; West et al., 2021). To do so, we randomly split our sample into three groups in which landowners would receive the one-time signing bonus at the time of program enrollment, two years after program enrollment, or five years after program enrollment. DCE questions are the same across three groups, with the only difference being when the one-time signing bonus would be received.

The survey also asked landowners about their land operation, socio-demographics, prior experience with various farm support programs, and their attitudes toward the government and farm support programs. To elicit landowners' risk perceptions, we adapted questions from Falk et al. (2016) using risk scales that range from "completely unwilling to take risks" to "very willing to take risks." Risk scale questions have been increasingly used in different types of surveys to elicit risk preferences[27] and are reported to yield quantitatively similar results when compared to estimates from using experimental methods such as incentivized choice experiments (Falk et al., 2016).

2.4.2 Sample construction and survey implementation

Our sample population was drawn from the spatially explicit parcel-level tax assessor database from the Maryland Tax and Assessment Office, covering a complete list of land parcels that have received agricultural use assessments. We first screened our sample to select all farmland parcels adjacent to different types of waterbodies (e.g., rivers, streams, lakes,

[27] See Iyer et al. (2020) for an overview of farmer surveys that use the risk scale assessment questions.

ponds, wetlands, shore), where riparian buffers could be installed based on physical constraints and CREP requirements in Maryland. We further limited to parcels that have at least 10 acres of land uses in major crops (e.g., corn, soybeans, wheat), other cultivated crops such as vegetables and fruits, or hay and pastureland. Land-use acreages were computed by overlaying parcel boundaries with the USDA's Cropland Data Layer. As it is common for landowners to own multiple parcels, we randomly selected one parcel if a landowner has more than one eligible parcel after screening. The above-mentioned procedures led to a total of 8,923 eligible farmland parcels as our survey sample.

The survey was administered between May and July 2021 using an address-based sampling approach. To invite eligible landowners in our survey sample to complete the online survey through Qualtrics, we mailed out three recruitment letters using landowners' addresses obtained from the same parcel-level tax assessor database. The initial invitation letter was followed, at weekly intervals, by two reminder letters. All letters included instructions for completing the online survey and provided the survey link (URL), along with unique identification numbers and passwords linked to each landowner. These unique numbers were also matched to each parcel address, allowing us to know the parcel's location for the inclusion of geospatial data in the analysis. The invitation letters also provided the option to have a paper copy of the survey mailed to them.

Of the 8,923 landowners in our survey sample, we received 1,420 online complete responses and 110 surveys were completed by mail, with a response rate of 17%.[28] Not all landowners who completed our survey faced DCE questions. In the survey, we specifically asked whether each landowner currently has any riparian buffers and if so, whether they have

[28] To explore if response bias is an issue in this study, we estimate landowners' probability of responding to the survey using a probit model, where the dependent variable equals one if a landowner responded to our survey and explanatory variables include a list of parcel-specific characteristics. The results, attached in the Appendix (Table B.1), show that most of these variables are not statistically significant except for the parcel location. We find that landowners with parcels in certain crop reporting districts are less likely to respond to our survey.

any cropland or pastureland for additional buffers. Those who currently have riparian buffers but do not have any cropland or pastureland for more buffers were not presented with DCE questions.[29] That said, the sample used in this study consists of landowners who have land that is eligible for new riparian buffers, and they are exactly those that any efforts to expand the program would target. Among all complete responses, we have 552 landowners who faced DCE questions, yielding a total of 2,208 DCE questions. Particularly, 254 out of 552 landowners who faced DCE questions (about 46% of our sample) chose the status quo option in all four DCE questions, who can be considered serial non-participants in this study.

2.4.3 Summary statistics

Table 2.2 presents variable descriptions and summary statistics for DCE variables and landowner and parcel characteristics used in our empirical analysis. The binary dependent variable is *Enroll*, which takes a value of 1 if the landowner indicated they would enroll under each proposed program and 0 otherwise. On average, landowners chose to enroll in only 35.8% of the proposed programs, even with payment rates substantially higher than what is offered by the current program in Maryland. Regarding landowner characteristics, about half of landowners in our sample currently rent out any of their land to others, 57.0% of them are seniors (age over 65), and 60.6% of them have college degrees or higher. About 28.6% of landowners are found to be risk averse. While 26.4% of them have riparian buffers without government support, only 5.8% of them reported that they are currently enrolled in riparian buffer programs. Notably, 60.5% of landowners agree that the government should not be

[29] To explore if sample selection is an issue in this study, we estimate landowners' probability of facing DCE questions using a probit model, where the dependent variable equals one if the landowner faced DCE questions and explanatory variables include a list of parcel and landowner characteristics. The results, attached in the Appendix (Table B.2), show that most of these variables are not statistically significant except for a few characteristics. Particularly, we find that landowners who enroll in current programs or self-fund riparian buffers are less likely to face DCE questions. Table B.3 further compares the summary statistics for a comprehensive list of variables collected from the survey between landowners who faced DCE questions and those who did not.

allowed to come onto their property and monitor their farm operations and 19% of them expressed that tax revenues should not be used for farm support programs.

It is worth noting that our survey sample is by no means representative of Maryland farmers as characterized by the 2017 Census of Agriculture (see Table B.4 for comparisons in summary statistics for farm and landowner characteristics) due to several reasons. First, we specifically targeted farmland owners in this survey since landowners, rather than farm operators, are eligible for enrolling in riparian buffer programs. Second, only farmland owners who have applied for agricultural use assessments are included in the parcel database we drew our sample from and not all farm operators did so. Third, not all farmland owners in Maryland would be included in the Census if they are not farm operators or producers. A large share of our respondents rented out all their farmland and would thus not be included in the Census. Lastly, as our DCE questions are designed for landowners who are eligible for future program enrollment, such a selection likely makes our sample different from the general farmer population. Summary statistics show that landowners used in this study have slightly higher percentages of crops and pasture operation, lower percentages of animal ownership, lower gross farm sales revenues, and are older when compared to Maryland farmers documented in the Census.

2.5 Results

2.5.1 Logit, SHL, and DHL model results

We report the regression results from the logit, SHL, and DHL models in Tables 2.3 and 2.4. To compare the relative statistical fits of the three models, we consider the log-likelihoods and three information criteria (i.e., AIC, BIC, and CAIC). We also adopt a Vuong non-nested hypothesis test as done in von Haefen et al. (2005), Burton and Rigby (2009), and Chen et al. (2020). Specifically, three separate pairwise null hypothesis tests are conducted,

including that the logit model is statistically equivalent to or better than the SHL model, the logit model is statistically equivalent to or better than the DHL model, and the SHL model is statistically equivalent to or better than the DHL model. Both the information criteria and the Vuong test results, shown in Table 2.4, indicate that the DHL model fits the data generating process best among the three models, which is considered the preferred model in this study.

Comparison between the logit and two hurdle models shows that accounting for SNP in a hurdle framework, we find less heterogeneity in the status quo utility among program participants as shown in the choice equation of the two hurdle models. The key difference is that the logit and hurdle models treat heterogeneity differently. While the logit model treats utility as completely linear in payments and landowner characteristics, the hurdle models introduce nonlinearity and assume that the preferences of (some) serial non-participants differ qualitatively from those who chose at least one proposed program. Substantial differences in parameters for program attributes from the choice equation between the logit and two hurdle models (and between the logit model with the full sample and the logit model with the subsample of participants as shown in Table 2.3) further suggest that serial non-participants' preferences for program attributes can be different from participants, highlighting the need for a separate data generating process to formally account for SNP.

Table 2.4 also shows the hurdle equation results from SHL and DHL models, revealing certain important factors that contribute to landowners being serial non-participants. For example, landowners aged over 65 are more likely to be serial non-participants, due possibly to future retirement and thus not being willing to involve in long-term contracts. Landowners who currently have riparian buffers regardless of enrollment in current programs are less likely to be serial non-participants. Particularly, landowners' attitudes toward the government and government programs can significantly explain SNP. As expected, if landowners agree that the government should not monitor farm operations or tax revenues should not be used for farm

support programs, they are more likely to choose the status quo option in all DCE questions. This finding indicates that ceiling on program participation may come from non-monetary factors and that simply making programs more lucrative is may not raise that ceiling very much.

However, we find small differences in the results between SHL and DHL models. This result indicates that only a small proportion of serial non-participants whose preferences are approximately the same as those of participants are in our sample and are captured in the DHL model but not in the SHL model. This is likely attributable to the fact that our DCE questions offered payment levels that are already much higher than current payment rates. Serial non-participants are likely landowners whose reservation prices are somewhat larger than payment levels government programs could offer.

2.5.2 Program preferences and heterogeneity

When it comes to PES programs, how the contract is structured such as payment types, when the payment is received, and the contract length can influence landowners' enrollment decisions. Additionally, heterogeneity in landowner and parcel characteristics will likely play a role in the likelihood of enrollment. Across alternative models, we consistently find that the coefficient estimates for the one-time signing bonus and recurring annual payments are statistically significant. As expected, higher payment levels increase the likelihood of enrollment. Any delay in receiving the signing bonus decreases enrollment, although the coefficient estimates for the two payment delay interaction terms are economically insignificant and statistically not significantly different from zero. Such a result suggests that discounting in receiving payments over the contract lifetime may not matter when landowners decide to enroll, at least for the one-time signing bonus. We further find that contract length does not seem to significantly affect landowners' probability of enrollment as the coefficient estimates for the two contract length interaction terms are not statistically significant.

We find evident heterogeneity in landowners' status quo utility, indicating certain factors playing an important role in enrollment decisions. Of our particular interest is that risk-averse landowners are less likely to enroll. This result suggests that, although riparian buffer programs offer steady and risk-free payments, landowners seem to perceive program participation as increasing income risk. This interesting finding suggests that conservation programs, at least those incentivize the adoption of long-lived conservation practices on working farmland, are not viewed as unadulterated risk-reduction measures. One reason is that conservation practices implemented on working farmland may impose risks and uncertainty to the remaining farm production due to potential changes in existing farm operations (Canales et al., 2015; Ramsey et al., 2019). This is especially true in our case as farmland needs to be taken out of production for riparian buffers that have to be maintained differently from surrounding farmland for a long period of time. Another reason is that enrolling in a long-term government contract involves a reduction in the option value of selling the land for development, which also makes the program less attractive.

Our results also show that landowners who rent out any land are less likely to choose the status quo option and thus more likely to enroll. Installing riparian buffers on working farmland may impose some changes to existing farm operations and may make future leasing less attractive to renters, inducing some opportunity costs. Manley and Mathias (2017) find that higher opportunity costs such as owner-operated farms are negatively correlated with CREP participation. Our results support their finding and suggest that opportunity costs associated with program participation are not negligible and could have a significant impact on program participation.

2.5.3 Willingness to accept estimates

To monetize landowners' preferences for program participation, we estimate the annualized WTA (i.e., a constant annual payment) as the compensation that would make the

landowner indifferent between enrolling and not enrolling in a specific program. Annualized WTA estimates for different types of contracts and different types of landowners can be computed and Table 2.5 presents the average WTA estimates of our interest using a forest buffer program as an example. Note that to calculate the average WTA for a riparian buffer program with a specified contract length, we first calculate individual WTA using equation (2.14), based on which we compute the average WTA estimates. To allow for comparisons in WTA estimates across alternative models, we focus on landowners who are considered potential enrollees (i.e., those who chose at least one proposed program) since our WTA estimates are conditional on program participation.

We find significant differences in WTA estimates across three models and that using the parameters estimated from the logit model can lead to quantitatively different welfare measures. In detail, WTA estimates from the logit model are substantially higher (more than three times larger) when compared to those from the two hurdle models. These WTA estimates are also unreasonably higher than current program payment levels while those estimated from the two hurdle models are within the reasonable payment range. Such a result suggests that linearly extrapolating serial non-participants' preferences from the observed preferences of landowners who chose at least one proposed program in the logit model can lead to an overestimation of welfare measures. By contrast, the two hurdle models, in which serial non-participants are assumed to behave differently from participants, can generate more reasonable WTA estimates. Our results thus highlight the importance of properly treating serial non-participants in stated preference studies for welfare analyses.

2.5.4 Discount rate estimates

Many voluntary government programs operate in an intertemporal dimension, with costs and benefits that accrue over different time horizons. Some programs may generate substantial initial benefits accompanied by costs that occur far into the future, while others may

entail large upfront costs and generate a long temporal stream of benefits. Understanding whether targeted population responds to long-term government programs and how targeted population perceives discounting in intertemporal decisions can help policymakers improve program design. One goal of this study is to estimate the implicit discount rate when landowners respond to government payments through enrollment in long-term PES programs.

Table 2.6 presents our discount rate estimates using the geometric (discrete time) discounting formulation with different time intervals. We find that using the parameters estimated from the logit model can also make a difference in the estimation of implicit discount rates due to the treatment of serial non-participants. Discount rates estimated from the logit model are much higher than those estimated from hurdle models, especially for the payment delay in 2 years. For our preferred model specification (i.e., DHL), the discount rate estimates are 4.6% and 1.7%, respectively for the payment delay in 2 and 5 years. These estimates are similar to the magnitudes found in Lloyd-Smith et al. (2021) in the context of landowners' participation in agroforestry programs in Panama. However, all these estimates turn out to be not statistically significant. Our low discount rate estimates suggest that landowners show little time preference for receiving the one-time signing bonus and do not seem to prefer receiving the payment sooner than later. Additionally, we find weak evidence of hyperbolic discounting as the two discount rate estimates are not statistically different from each other.

Multiple reasons may explain the low discount rates found in our study. First, as our choice scenarios are certain with guaranteed government payments, our estimated discount rates are thus risk-free, which should be lower than discount rates adjusted for risk. Second, in our experimental design, the signing bonus represented a small fraction of the total payments throughout the contract lifetime in most cases. On average, only 23.4% of the total payments was the signing bonus across all DCE questions in our study. Given that only the signing bonus was delayed in our experimental design, the relatively small portion of the signing bonus could

contribute to the low discount rates. Third, we conducted our survey during a time of historically low interest rates for farmers, with interest rates of 1.75% and 2.875% respectively for farm operating and ownership loans as of 2021. Lastly, our sample is more educated than the population at large with about 60.6% having college degrees or higher. Global evidence shows that higher educational attainment is positively related to patience (Falk et al., 2018) and as such landowners in our sample are likely more patient.

Notably, our low discount rates estimated from the supply side of the market contrast with the demand-side SP studies that find generally large discount rates using the same ex-post method for estimation.[30] The disparity between the two discount rates estimated from the supply and demand side has some important implications for policymakers to conduct the cost-benefit analysis, based on which they decide whether to fund programs for the provision of ecosystem goods and services. While the benefits of ecosystem goods and services seem to discount substantially over time, the costs of providing these goods and services may present little discounting. This finding is valuable since governments often operate under tight budget constraints and want to know the lowest-cost incentive possible to achieve desired policy outcomes. If landowners have very low discount rates and place a large weight on future payments, government agencies may not need to offer certain incentive payments upfront. Instead, these payments can be provided by the time when projects are complete without discouraging landowners from enrolling in the program.

[30] For example, Viscusi et al., (2008) report a discount rate ranging from 48% to 61% for policies that aim to improve water quality; Kim and Haab (2009) estimate a discount rate within the range of 20-131% for oyster reef restoration programs; Myers et al. (2017) report a discount rate of over 100% for protecting migratory shorebirds; and Howard et al. (2021) find a discount rate within the range of 14-31% for protecting invasive species.

2.6 Policy Simulation on Probabilities of Enrollment

Raising payment levels is often proposed as a means of increasing participation in PES programs. Our econometric results can be used to estimate the supply curves and thus shed light on how effective increasing payment levels could be to incentivize participation, especially in the presence of SNP. Policy simulations on probabilities of enrollment are of particular policy relevance in this study since policymakers in Maryland have been considering raising payment levels to comply with the EPA's total maximum daily load requirements for the Chesapeake Bay watershed.

To illustrate, we show the relationship between the predicted probability of enrollment and the annual payment rate, using a 15-year contract with a one-time signing bonus of $1,000 per acre with no delay in receipt as an example.[31] Figure 2.2 presents the average probabilities of enrollment for all landowners in our sample and compares the results across alternative models. As expected, the predicted probability of enrollment increases as the annual payment rate increases, with everything else kept the same. While the two hurdle models may not seem to generate significant differences when payment levels are low, the disparity in the predicted probability of enrollment between the logit and hurdle models diverges quickly as the annual payment rate rises. For example, when the annual payment is close to what the CREP currently offers in Maryland (e.g., $500 per acre), the average predicted probability of enrollment for a 10-year grass buffer program estimated from the logit, SHL, and DHL model is respectively 45.7%, 43.8%, and 44.1%. If the annual payment level increases to what may be considered for future programs (e.g., $1,000 per acre), the average predicted probability of enrollment for a grass buffer program estimated from the logit, SHL, and DHL model is respectively 54.1%,

[31] Since 2021, the Maryland Department of Agriculture has increased the one-time signing bonus to $1,000 per acre for eligible landowners who install forest buffers along qualifying stream corridors through CREP. Previously, the one-time signing bonus was at $200 per acre, with $100 per acre provided by the Maryland Department of Agriculture and the other $100 per acre provided by the USDA's Farm Service Agency.

49.5%, and 50.4%. The standard logit model can generate a noticeable difference in the probability of enrollment that's close to 10% higher than the hurdle models.

Notably, assumptions about SNP embedded in alternative models make a substantial difference if we keep raising the payment rate. Ceiling on program participation becomes evident in the two hurdle models when SNP is accounted for while program participation is linearly projected in the logit model. In contrast to the logit model that assumes all serial non-participants to share similar choice behaviors with participants, the hurdle models explicitly allow serial non-participants' preferences to behave differently from participants, pointing to an inherent limitation in most voluntary programs that aim to provide ecosystem goods and services. Not accounting for SNP in conducting policy simulations will significantly inflate the uptake rate and thus lead to an overestimation of the effect of incentive payments on program uptake. This finding highlights an important policy implication. Since policymakers have been considering increasing upfront signing bonuses and/or annual payments to boost program enrollment, our result suggests that simply making programs more lucrative may not raise that ceiling substantially due to the presence of landowners whose reservation prices lie outside the reasonable range offered by government programs.

2.7 Conclusions

SP surveys have been commonly used to evaluate landowners' preferences for PES programs and WTA for participation. The occurrence of SNP begs the question of whether we can extrapolate from participants' observed choice behaviors to serial non-participants' unobserved preferences. While there has been some discussion on how SNP should be treated on the demand side of the market when evaluating individuals' WTP for ecosystem goods and services, research on the supply side of the market focusing on WTA for PES program participation is relatively limited. This paper studies landowners' willingness to enroll in long-term PES contracts using data from a DCE implemented in a farmer survey and specifically

addresses SNP using hurdle-variants of discrete choice models. We also fill the gap in the literature by exploring the role of time and risk preferences in landowners' enrollment in long-term PES programs.

In contrast to the logit model which assumes serial non-participants and participants to share similar choice behaviors, the SHL and DHL models introduce a different data generating process to explain SNP and allow serial non-participants to behave fundamentally different from participants. We find that whether serial non-participants are accounted for in the econometric model is crucial for the empirical analysis and can make a substantial difference in inferences about policy. We highlight such differences in our post-estimation calculations of the predicted probability of enrollment, average WTA for program participation, and implicit discount rates from alternative models, in which different assumptions about serial non-participants are made.

Our results show substantial differences in parameters for program attributes from the logit and two hurdle models, leading to quantitatively different effects of raising payment levels on program uptake. Specifically, our simulated probabilities of enrollment indicate that ceiling on PES program participation is evident when SNP is explicitly modeled using a hurdle framework that allows serial non-participants to behave differently from participants. This result points to an inherent limitation in most voluntary programs that aim to provide ecosystem goods and services, suggesting that simply making programs more lucrative may not raise that ceiling very much. Failing to account for SNP can also result in quantitatively different WTA estimates for program participation, which may lead policymakers to draw misleading policy implications.

Table 2.1: Program Attributes and Attribute Levels

Program element	What it means
Buffer type	**Type of buffer to be installed**. Options include: Grass buffer, forest buffer
Bonus payment	**One-time bonus payment ($ per acre) for enrolling in the program**. Options include: $200, $500, $1,000, $1,500 per acre
Annual payments	**Recurring annual payments ($ per acre)**. Options include: $100, $250, $500, $750 per acre
Contract length	**Number of years to maintain the buffer**. Options include: 5, 10, 15 years

Table 2.2: Variable Descriptions and Summary Statistics

Variable	Description	Mean	S.D.
DCE variables			
Enroll	Dummy variable indicating enrollment in the proposed program	0.358	0.479
Grass	Dummy variable indicating a grass buffer	0.544	0.498
Forest	Dummy variable indicating a forest buffer	0.456	0.498
Bonus	One-time bonus payment rate ($1,000/acre)	0.835	0.485
Delay_0	Dummy variable indicating no delay in bonus payment	0.329	0.470
Delay_2	Dummy variable indicating 2 years of delay in bonus payment	0.327	0.469
Delay_5	Dummy variable indicating 5 years of delay in bonus payment	0.344	0.475
Annual	Recurring annual payment rate ($1,000/acre)	0.414	0.249
Contract_5	Dummy variable indicating a 5-year contract	0.327	0.469
Contract_10	Dummy variable indicating a 10-year contract	0.343	0.475
Contract_15	Dummy variable indicating a 15-year contract	0.330	0.470
Landowner and parcel characteristics			
Rent_out	Dummy variable indicating renting out any land	0.504	0.500
Farm_income	Share of household income from farming	0.154	0.268
Senior	Dummy variable indicating age over 65	0.570	0.495
College	Dummy variable indicating having a college degree	0.606	0.489
Risk_averse	Dummy variable indicating being risk averse	0.286	0.452
Enrollee	Dummy variable indicating a current buffer program enrollee	0.058	0.234
Self_funder	Dummy variable indicating currently self-funding any buffers	0.264	0.441
SRR	Parcel-specific soil rental rate ($1,000/acre)	0.047	0.044
Attitude_property	Dummy variable indicating agreement on the government should not be allowed to come onto my property and monitor my farmland operations	0.605	0.489
Attitude_tax	Dummy variable indicating agreement on tax revenues should not be used for farm support programs	0.190	0.393

Table 2.3: Logit Regression Results

Variable	**Logit (Full sample)**		**Logit (Participants only)**	
	Coefficient	S.E.	Coefficient	S.E.
Forest	-0.129	(0.100)	-0.177	(0.130)
Bonus	0.264**	(0.126)	0.404**	(0.167)
Bonus × Delay_2	-0.092	(0.127)	-0.005	(0.169)
Bonus × Delay_5	-0.061	(0.125)	-0.018	(0.170)
Annual	0.773***	(0.275)	1.447***	(0.366)
Annual × Contract_10	0.307	(0.255)	0.351	(0.360)
Annual × Contract_15	0.125	(0.262)	0.125	(0.363)
Rent_out	0.254**	(0.102)	0.331**	(0.132)
Farm_income	-0.685***	(0.202)	-0.429	(0.274)
Senior	-0.670***	(0.100)	0.032	(0.131)
College	0.141	(0.106)	-0.250*	(0.147)
Risk_averse	-0.761***	(0.118)	-0.351**	(0.158)
Enrollee	0.716***	(0.207)	0.061	(0.242)
Self_funder	0.759***	(0.113)	0.224	(0.144)
SRR	-3.602***	(1.143)	-0.310	(1.462)
Attitude_property	-0.676***	(0.102)	-0.182	(0.132)
Attitude_tax	-0.329**	(0.133)	0.228	(0.192)
Constant	-0.230	(0.199)	-0.120	(0.254)
Log-likelihood	-1,197		-702.9	
AIC	2,431		1,442	
BIC	2,532		1,533	
CAIC	2,562		1,563	
Number of variables	18		18	
Number of observations	2,021		1,148	

Note: For buffer type, grass buffer is the baseline. For the time delay in receipt of the one-time signing bonus, no delay is the baseline. For contract length, a 5-year contract is the baseline. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.4: SHL and DHL Model Results

Variable	**SHL**		**DHL**	
	Coefficient	S.E.	Coefficient	S.E.
Choice equation				
Forest	-0.176	(0.128)	-0.177	(0.123)
Bonus	0.399**	(0.158)	0.427***	(0.142)
Bonus × Delay_2	-0.024	(0.160)	-0.037	(0.160)
Bonus × Delay_5	-0.015	(0.160)	-0.035	(0.163)
Annual	1.422***	(0.341)	1.433***	(0.316)
Annual × Contract_10	0.352	(0.345)	0.324	(0.353)
Annual × Contract_15	0.140	(0.353)	0.117	(0.325)
Rent_out	0.354***	(0.128)	0.390***	(0.135)
Farm_income	-0.461*	(0.261)	-0.534*	(0.293)
Senior	0.035	(0.122)	0.037	(0.129)
College	-0.262*	(0.140)	-0.297**	(0.143)
Risk_averse	-0.358**	(0.161)	-0.398**	(0.169)
Enrollee	0.060	(0.245)	0.064	(0.255)
Self_funder	0.199	(0.143)	0.219	(0.145)
SRR	-0.219	(1.419)	-0.054	(1.445)
Constant	-0.164	(0.192)	-0.227	(0.201)
Hurdle equation				
Rent_out	-0.053	(0.064)	0.002	(0.130)
Farm_income	0.345***	(0.115)	0.323	(0.244)
Senior	0.650***	(0.051)	0.664***	(0.126)
College	-0.267***	(0.063)	-0.279**	(0.134)
Risk_averse	0.480***	(0.067)	0.463**	(0.142)
Enrollee	-0.738***	(0.141)	-0.809***	(0.320)
Self_funder	-0.618***	(0.076)	-0.632**	(0.144)
SRR	2.886***	(0.680)	2.791***	(1.394)
Attitude_property	0.537***	(0.067)	0.561***	(0.132)
Attitude_tax	0.386***	(0.075)	0.406***	(0.157)
Constant	-0.895***	(0.086)	-0.973***	(0.192)

Log-likelihood	-1,088	-988.9
AIC	2,235	2,032
BIC	2,422	2,202
CAIC	2,427	2,229
Number of variables	27	27
Number of observations	2,021	2,021
Vuong test	logit v. SHL (p=0.064)	logit v. DHL (p=0.001)
	SHL v. DHL (p=0.023)	

Note: For buffer type, grass buffer is the baseline. For the time delay in receipt of the one-time signing bonus, no delay is the baseline. For contract length, a 5-year contract is the baseline. The Vuong test results show that we reject the following three null hypotheses at p=0.1: the logit model (full sample) is statistically equivalent to or better than the SHL model; the logit model (full sample) is statistically equivalent to or better than the DHL model; the SHL model is statistically equivalent to or better than the DHL model. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.5: Mean Annualized WTA Estimates for a Forest Buffer for Participants Only

Contract length	**Logit**	**SHL**	**DHL**
5 years	1.213 (0.949)	0.276 (0.205)	0.324 (0.227)
10 years	0.868 (0.679)	0.221 (0.164)	0.265 (0.185)
15 years	1.044 (0.816)	0.251 (0.187)	0.300 (0.210)

Note: Mean annualized WTA estimates (in $1,000/acre) are calculated for the subsample of participants using the coefficient estimates from the choice equation in Table 2.3 (for the logit model with full sample) and Table 2.4 (for the SHL and DHL models). Standard deviations are in parentheses.

Table 2.6: Discount Rates by Time Interval Across Alternative Models

Model specification	**No delay versus 2 years of delay**		**No delay versus 5 years of delay**	
	Discount rate	S.E.	Discount rate	S.E.
Logit	0.238	(0.399)	0.054	(0.117)
SHL	0.031	(0.214)	0.007	(0.083)
DHL	0.046	(0.209)	0.017	(0.082)

Note: Mean discount rates in discrete time measures are calculated using the coefficient estimates from the choice equation in Table 2.3 (for the logit model with full sample) and Table 2.4 (for the SHL and DHL models). Standard errors, calculated using the delta method, are in parentheses.

Figure 2.1: Sample Discrete Choice Experiment Question

Would You Enroll?

Assume that **Program A** was offered as a possible program. Would you enroll in this program to install and maintain a riparian buffer on your property? If you choose "yes", we will ask you the length of the buffer (in feet) you would install under the chosen program.

Please remember that:

- Installation costs and maintenance costs will be fully covered by the program, regardless of the buffer type offered in the program
- You will receive the one-time bonus payment **at the time you enroll in the program**
- The program requires a minimum buffer width of 35 feet

Program element	**Program A**
Buffer type	Grass buffer
Bonus payment ($/acre)	$500
Annual payments ($/acre)	$250
Contract length (years)	10

The payment schedule for **Program A** will look like the following:

	Program A
Year 0 – Bonus payment ($/acre)	$500
Year 1 – Annual payment ($/acre)	$250
Year 2 – Annual payment ($/acre)	$250
Year 3 – Annual payment ($/acre)	$250
Year 4 – Annual payment ($/acre)	$250
Year 5 – Annual payment ($/acre)	$250
Year 6 – Annual payment ($/acre)	$250
Year 7 – Annual payment ($/acre)	$250
Year 8 – Annual payment ($/acre)	$250
Year 9 – Annual payment ($/acre)	$250
Year 10 – Annual payment ($/acre)	$250
	Contract ends

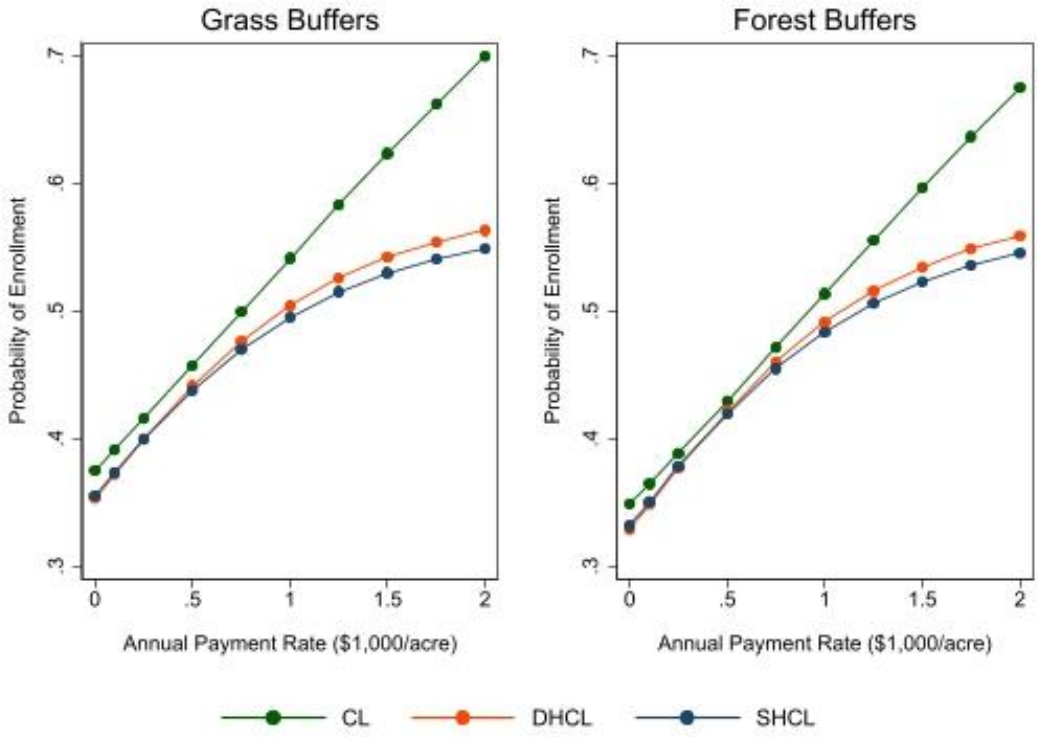
Would you enroll in **Program A**? (Choose one)

○ Yes – I would enroll

○ No – I would not enroll

Note: This figure provides an example of a DCE question, which describes a program that would require the installation of a grass buffer in exchange for a one-time signing bonus of $500/acre and a recurring annual payment of $250/acre for 10 years. Additionally, installation and maintenance costs are fully covered by the program and the landowner would receive the one-time signing bonus at the time of enrolling in the program.

Figure 2.2: Policy Simulations on Probabilities of Enrollment



Note: The two figures show the average predicted probabilities of enrolling in a 10-year riparian buffer program with a one-time signing bonus of $1,000/acre with no delay in receipt, respectively for a grass and forest buffer.

Chapter 3: Hot Spots, Cold Feet, and Warm Glow: Identifying Spatial Heterogeneity in Willingness to Pay

3.1 Introduction

A better understanding of the spatial distribution of welfare impacts is necessary for conducting accurate benefit-cost and distributional analyses, both in terms of defining the appropriate extent of the market and in interpolating values within that market. There is a rich body of literature on the spatial dimensions of stated preference (SP) studies, focusing on various analytical and statistical methods (see Glenk et al., 2020 for a review). These approaches include the incorporation of space within the survey design (Wang and Swallow, 2016; Badura et al., 2020), the combining of spatial variables with traditional econometric methods (Jørgensen et al., 2013; Schaafsma et al., 2013; Olsen et al., 2020), and the application of spatial econometrics and geo-statistics (Czajkowski et al., 2017; Budziński et al., 2017; Foelske and van Riper, 2020; Budziński and Czajkowski, 2021; Danley et al., 2021; Toledo-Gallegos et al., 2021).

To examine spatial heterogeneity in willingness to pay (WTP), most SP studies have applied the distance decay paradigm, where WTP is hypothesized to diminish with distance from the resource (Bateman et al., 2006). Generally, estimating the effect of distance on WTP depends on the nature of the distance measure (e.g., travel, Euclidean, or geodesic distance) and the econometric model specifications once a particular distance measure is chosen (e.g., linear or non-linear distance decay). While there is a growing number of SP studies using traditional econometric methods to control for “distance decay” or other forms of spatial heterogeneity (e.g., Hanley et al., 2003; Rolfe and Windle, 2012; Olsen et al., 2020), these parametric approaches can sometimes fail to identify existing spatial patterns.

Disciplines in the natural sciences employ more spatially-oriented analytical tools to examine spatial patterns. These tools include tests for global spatial autocorrelation (Getis, 2007), spatial interpolation techniques (Anselin and Gallo, 2006), and local cluster or hotspot analyses (Wang and Qiu, 2017). These tools have been increasingly applied in economics, and in particular, in the nonmarket valuation literature. For example, studies have tested for and generally found positive global spatial autocorrelation of individual-specific WTP values (Campbell et al., 2009; Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015; Toledo-Gallegos et al, 2021). SP studies have also employed local indicators of spatial association (LISAs) or hotspot analysis to identify local clusters of systematically higher or lower WTP values (i.e., hot and cold spots, respectively) (Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015). In general, these studies found non-continuous, local spatial patterns of WTP.

In contrast to applications of these spatial tools in the natural sciences, the measures under study by economists are often estimates (e.g., WTP), and not observed values. Although previous studies have qualitatively recognized this fact, and its potential importance, no study to date formally accounts for the statistical precision of those estimates when conducting spatial analyses. We are the first to do so by incorporating techniques borrowed from meta-analytic methods into our spatial autocorrelation, interpolation, and hotspot analyses.

We set out to accomplish three main research objectives. First, we develop a two-stage spatial econometric approach to account for the fact that economic analyses typically observe estimated values of WTP and other measures of interest. We use Bayesian modeling techniques to estimate not only household-specific WTP, but also household-level measures of the variances of those estimates. Doing so can be important because some households may be, for example, less knowledgeable of the environmental commodity, have less defined preferences, or even be less engaged when taking the SP survey. If such households are systematically

distributed over space, then not accounting for the statistical precision of the household-specific WTP estimates can confound subsequent spatial analyses. With the household-specific empirical WTP distributions in-hand, we are able to treat the variables of interest not as given data, but as statistically derived estimates. The household-specific variances of the WTP estimates are directly incorporated into the spatial weights used in the second-stage spatial analyses.

Our second objective is to demonstrate our proposed two-stage methodology using data from a SP survey that elicited values for improvements in water quality in the Chesapeake Bay and lakes in the surrounding watershed. Using our proposed variance-adjusted tests for global spatial autocorrelation, spatial interpolation techniques, and hotspot analyses, we examine the spatial distribution and clustering of marginal WTP (MWTP) for improvements in several environmental attributes. We also examine the spatial distribution of status quo (SQ) effects, which are intended to capture potential biases for (e.g., "warm glow") or against (e.g., "cold feet") a policy option that are not explained by the choice attributes defining that policy option. To our knowledge, this is the first study to examine the spatial distribution of respondents exhibiting potential biases associated with SP methods. Such an examination provides insights to improve SP methods, welfare analysis, and future survey designs.

The third objective is to illustrate the potential policy implications of our proposed variance-adjusted spatial analyses. We use our two-stage methodology to estimate total WTP for projected improvements resulting from the Chesapeake Bay Total Maximum Daily Loads (TMDLs). The total benefit estimates are compared to spatial interpolation approaches that do not account for the statistical precision of the WTP estimates, similar to those used in earlier studies (e.g., Campbell et al., 2009; Johnston and Ramachandran, 2014; Johnston et al., 2015), and to conventional models that assume homogeneity of WTP across the population or control

for heterogeneity parametrically based on observed household characteristics (Moore et al., 2018).

Our semi-parametric results of the spatial interpolation suggest distinct local patterns in MWTP estimates for all attributes, and evident spatial heterogeneity across the study area. The hotspot analysis confirms statistically significant spatial clusters of high and low MWTP values. Comparison of the conventional spatial analyses to our variance-adjusted results reveals some differences. In general, accounting for the variances of the MWTP estimates diminishes spatial variation, suggesting that not accounting for the statistical precision of the first-stage MWTP estimates could lead analysts to falsely identify patterns of global and local spatial heterogeneity in the second stage. Our analysis of spatial variation of individual-specific SQ effects reveals substantial differences when accounting for the statistical precision of the estimates. In particular, our proposed variance-adjustment leads to an increased identification of clusters of individuals exhibiting "warm glow" or other biases for a policy option. Lastly, although differences in local patterns are revealed, our policy simulations suggest that accounting for local spatial heterogeneity (with or without our variance-adjusted extension) may not yield substantial differences in terms of broader welfare implications, at least not in this particular application of water quality and ecosystem improvements in an iconic and well-known estuary.

3.2 Methods

3.2.1 Random utility models

Stated choice models are often estimated in a random utility framework, where v_{ij} denotes the deterministic component of utility that respondent i receives from alternative j in choice occasion t. Each respondent is given three choice questions in the application presented, but the choice occasion subscript t is omitted for notational ease. The random component of

utility is denoted as ε_{ij}. Assuming ε_{ij} is independently and identically distributed following a type I extreme value distribution allows the model to be estimated as a conditional or mixed logit (Maddala, 1983; Greene, 2003; Train, 2003).

The deterministic component of indirect utility v_{ij} is a function of the vector of environmental improvements $\boldsymbol{x}_j$, the cost-of-living increase incurred by the household $cost_j$, and a binary indicator denoting the status quo option SQ_j. To better capture preference heterogeneity, we interact $\boldsymbol{x}_j$ with a vector of dummy variables denoting whether individual i is a user of the corresponding resource $user_i$. We adopt a linear model and log-transform the environmental attributes to capture diminishing marginal utility, while also preserving more degrees of freedom than a model with higher order effects. Marginal utility of income is assumed constant across users and nonusers. Using this specification of $v(\cdot)$, the conditional probability that household i would choose alternative j is:

$$P_i\left(j \mid \boldsymbol{x}_q, cost_q, SQ_q, user_i\right) = \frac{exp\left\{ln\left(\boldsymbol{x}_j\right)\boldsymbol{\beta} + \left(ln\left(\boldsymbol{x}_j\right) \times user_i\right)\boldsymbol{\delta} + \gamma cost_j + \varphi SQ_j\right\}}{\sum_q exp\left\{ln\left(\boldsymbol{x}_q\right)\boldsymbol{\beta} + \left(ln\left(\boldsymbol{x}_q\right) \times user_i\right)\boldsymbol{\delta} + \gamma cost_q + \varphi SQ_q\right\}} \tag{3.1}$$

where q indexes all available alternatives in a given choice occasion. The coefficients to be estimated are $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, γ, and φ. The first three coefficients can then be used to derive the $MWTP_i$ vector for respondent i:

$$MWTP_i = \frac{\boldsymbol{\beta} + (user_i)\boldsymbol{\delta}}{-\gamma\tilde{\boldsymbol{x}}} \tag{3.2}$$

Given the natural log specification in the model, MWTP is nonlinear and varies at different levels of the environmental attribute. In equation (3.2), $\tilde{\boldsymbol{x}}$ denotes the environmental attribute reference levels from which $MWTP_i$ is calculated. In the empirical analysis, we set $\tilde{\boldsymbol{x}}$ equal to the baseline values shown in the survey.

3.2.2 Household-level MWTP distributions

To derive the household-level MWTP distributions that inform our spatial analyses, we adopt a random parameters framework to estimate the logit model characterized by equation (3.1). To simplify notation, we stack the variables $ln(x)$, $cost$, SQ, and the user interaction terms into a single M-element vector z and label the corresponding vector of coefficients λ. The typical exposition of mixed logit models estimates the distribution of the utility parameters over the population $g(\lambda|\upsilon)$, where υ are the parameters of the distribution such as mean and variance. When estimating household-level parameters, the central concept is the distinction between two distributions: the distribution of preferences in the population and the distribution of preferences in the sub-population who make particular choices (Train, 2003, p. 263). To that end, let $\eta(\lambda|\boldsymbol{y}, z, \upsilon)$ be the distribution of λ in the sub-population that chose a particular set of responses to the repeated choice experiments contained in the *t*-element vector $\boldsymbol{y}$.

The conditional distribution $\eta(\lambda|\boldsymbol{y}, z, \upsilon)$ can be estimated in a classical framework via maximum likelihood, and household-level parameters for that distribution can then be found by substituting in values of $\boldsymbol{y}$ and z (Revelt and Train, 2000). Generally, those expressions will be integrals without closed forms and require simulation. We adopt a Bayesian approach to estimation because the household-level distributions are more easily simulated as part of the Metropolis-Hastings (MH) algorithm, from which any moments of those distributions can be found. We provide a derivation of $\eta(\lambda|\boldsymbol{y}, z, \upsilon)$ and a detailed description of the estimation algorithm in Appendix C.1. There is nothing fundamentally different about our approach from the hierarchical Bayes' iterative estimation that has become standard in the literature (Train, 2003). What allows us to characterize household-level parameter distributions is the omission of a step in the standard algorithm that is usually included for computational efficiency.

Typically, when hierarchical Bayes is executed, the draws of the household-level parameters are only used to condition draws of the population-level parameters and then

discarded to preserve computational memory and improve processing speed. In this application, however, we store the household-level parameter estimates and treat them as draws from the empirical distributions of each respondent, $\eta(\lambda|\boldsymbol{y}, z, \upsilon)$. When the simulation is complete, we have a multivariate distribution of λ_i for each respondent i from which we can calculate draws from the distribution of MWTP via equation (3.2). Those vectors of MWTP values characterize the distribution for each individual respondent and can be used to calculate the mean and variance of MWTP at the household level. The latter gives us a measure of statistical precision for the MWTP estimates for each respondent.

Except for the coefficient on cost, all parameters are drawn from a multivariate normal distribution. We follow the common practice of treating the cost coefficient as fixed rather than random to ensure MWTP has defined moments (e.g., Revelt and Train, 1998; Layton and Brown, 2000), and we recognize this imposes a fixed marginal utility of income over the population. There are at least two alternatives to modelling the cost parameter that avoid this assumption. One is to choose a distribution with a strictly positive domain for the cost coefficient, such as log-normal. This is problematic in our application because our algorithm requires storing all draws of every coefficient, rather than just the mean value for each iteration. As a result, there can be draws of the cost coefficient small enough to result in very large MWTP values which create computational difficulties, requiring ad hoc solutions. The second alternative is to estimate the model in WTP space (Scarpa et al., 2008). This method brings its own computational issues and can result in less precise WTP estimates (Hole and Kolstad, 2012), so we opt for the simplest approach of fixing the cost parameter and acknowledge the implication on our results.

3.2.3 Household-level status quo effect distributions

In addition to household-level MWTP, we closely examine the SQ effects at the household level. Inclusion of the status quo indicator SQ_j allows us to capture a respondent's

tendency to vote for or against the SQ option, irrespective of the cost and attribute levels defining the alternative policy options. Such tendencies estimated by φ may capture respondents' consideration of omitted variables. It could also reflect "warm glow" if negative or "cold feet" if positive. For these reasons we omit φ from the MWTP calculations (see equation (3.2)). Nonetheless, the spatial distribution of φ could reveal important information about the validity and reliability of SP responses.

To provide an intuitive interpretation of the magnitude of the SQ effects, we express the effects in terms of probability differences. We compare the probability of choosing the SQ option to a constructed alternative that is identical with respect to the attribute levels but omits φ. We start with equation (3.1), which expresses the multinomial logit probabilities as a function of environmental attributes, the SQ effect, and the cost. We find the probability that each respondent would choose the SQ option ($j = SQ$) in a given choice occasion, $P_i(SQ \mid \boldsymbol{x}_q, cost_q, SQ_q)$. We then specify a second probability function for a constructed SQ alternative that omits the SQ constant, $\tilde{P}_i(SQ|\boldsymbol{x_q}, cost_q) = \frac{exp\{ln(\boldsymbol{x}_{SQ})\boldsymbol{\beta}+(ln(x_{SQ})\times user_i)\boldsymbol{\delta}\}}{\sum_q exp\{ln(\boldsymbol{x}_q)\boldsymbol{\beta}+(ln(\boldsymbol{x}_q)\times user_i)\boldsymbol{\delta}+\gamma cost_q\}}$, and estimate the SQ effect as the difference between the two probabilities:

$$SQ\ effect = P_i(SQ \mid x_q, cost_q, SQ_q) - \tilde{P}_i(SQ|x_q, cost_q) \tag{3.3}$$

We perform spatial analyses of the SQ effects and account for household-level variances in the same manner as the MWTP calculations, generating household-level distributions for the probability differences.

3.2.4 Spatial analyses

The main contribution of this study is the formal incorporation of the underlying statistical precision around each individual household's MWTP estimates into our spatial analyses. To our knowledge spatial clustering and interpolation studies examining spatial

variation to date, have treated the variables of interest as observed values, and not as statistically derived estimates (Campbell et al., 2009; Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015; Czajkowski et al., 2017; Toledo-Gallegos et al, 2021). Although these studies appropriately caveat their findings, none have formally accounted for the underlying statistical precision of the WTP estimates. We do so by borrowing techniques from meta-analytic methods. Conventional meta-analyses synthesize estimates from different primary studies, and in doing so often weight the primary study estimates according to their inverse variance (Borenstein et al., 2010). We incorporate this same idea into our three sets of spatial analyses – spatial interpolation, global autocorrelation, and hotspot analyses. The following discussion focuses on the MWTP estimates, but the same procedures are applied to the estimated SQ effects.

3.2.4.1 Variance-adjusted spatial interpolation

Spatial interpolation entails the creation of a raster (or grid) surface that depicts the distribution of household MWTP over space. The weighted household-specific MWTP values are used to predict MWTP for all locations in the study area, which in practice are identified as the centroid of each grid cell. The following equation is used to interpolate the MWTP value assigned to each cell l:

$$\widehat{MWTP}_l = \sum_{h=1}^{H} (\omega_{lh} \times MWTP_h) \tag{3.4}$$

where $\widehat{MWTP}_l$ is the predicted MWTP estimate at an unsampled location l, $MWTP_h$ is household h's estimated MWTP value, and ω_{lh} is the element from the spatial weights matrix that links locations l and h. We adopt the following functional form for our weighting equation to account for both the spatial relationships and statistical precision of the primary estimates:

$$\omega_{lh} = \begin{cases} \dfrac{\left(\frac{1}{d_{lh}}\right)^{\rho}\left(\frac{1}{v_h}\right)^{1-\rho}}{\sum_{h\in H_l}\left\{\left(\frac{1}{d_{lh}}\right)^{\rho}\left(\frac{1}{v_h}\right)^{1-\rho}\right\}}, & if\ h \in H_l \\ 0, & if\ h \notin H_l \end{cases} \tag{3.5}$$

where d_{lh} is the distance from the location of the centroid of cell l to household h, and v_h is the variance of the MWTP estimate for household h, which comes from the empirical distribution generated through the 10,000 iterations of our Bayesian modelling approach (after burn-in). The summation in the denominator is over the "K-nearest neighbors" to location l (denoted by the set H_l). Households at greater distances than the K-nearest neighbor are given a weight of zero.

Nelson and Boots (2008) have discussed several ways to define spatial weights matrices that include fixed distance, K-nearest neighbor, and shared boundaries. Following Johnston and Ramachandran (2014) and Johnston et al. (2015), we adopt the K-nearest neighbor method (K=8).[32] This spatial weighting scheme is appropriate for several reasons. First, K=8 is the number at which the permutation distribution of the test G_i^* statistic used in the later hotspot analysis approaches normality (Ord and Getis, 1995). Second, this method ensures that very far households across the large study area do not influence the MWTP estimates, so our results are not overly influenced by outliers. Third, the K-nearest neighbor method naturally adapts the neighborhood definition to account for different population densities in urban, suburban, and rural areas across our study area.

Of particular interest in equation (3.5) is the assumed value for the parameter ρ, which must satisfy $0 \leq \rho \leq 1$. The ρ parameter determines how much influence spatial proximity versus statistical precision of an estimate has on the spatially interpolated MWTP value for cell l. If $\rho = 1$, then equation (3.5) simplifies to the inverse distance weighting scheme commonly

[32] Johnston and Ramachandran (2014) conduct sensitivity analyses based on the assumed spatial relationships and found similar results across alternative assumptions.

used in past spatial analyses. If $\rho = 0$, then for the K-nearest neighbors, equation (3.5) is analogous to the common fixed effect size (FES) weighting scheme often utilized in meta-analyses (Borenstein et al., 2010). The choice of ρ is admittedly arbitrary but given our interests in accounting for both statistical precision and spatial patterns, we assume an equal influence of both factors and set $\rho = 0.5$. A sensitivity analysis is then conducted for alternative values of ρ, and most notably for the case where $\rho = 1$, which allows for comparison of our variance-adjusted weights to the conventional spatial weights used in previous studies.

3.2.4.2 Variance-adjusted global spatial autocorrelation and hotspot analyses

To test for global and local spatial autocorrelation, it is again necessary to define the neighborhood in which relationships across space are evaluated. In contrast to the interpolation exercise above, where the weights matrix denotes the spatial relationships between each interpolated cell centroid and the households in our sample, in the next set of analyses the spatial relationship defined is between each household and the K-nearest neighboring households, including the household itself (i.e., where $d_{ih} = d_{ii} = 0$). This mathematically prevents us from assuming an inverse distance relationship, as done in equation (3.5). Given our interests in identifying statistically significant high or low clusters of MWTP, a simple uniform weight of 1/K among the K-nearest neighbors (and zero otherwise) is assumed here.

Again, our novel contribution is to account for the statistical precision of the individual household MWTP estimates. More specifically, the weight given to each neighbor is re-distributed among the K neighbors, giving more weight to households where the observed MWTP value was estimated with greater statistical precision (i.e., a smaller variance). The weight used for household h in explaining the spatial relationship with household i is:

$$\omega_{ih} = \begin{cases} \dfrac{\left(\frac{1}{v_h}\right)^{1-\alpha}}{\sum_{h \in H_i}\left\{\left(\frac{1}{v_h}\right)^{1-\alpha}\right\}}, & if\ h \in H_i \\ 0, & if\ h \notin H_i \end{cases} \tag{3.6}$$

The summation in the denominator is over the K-nearest neighbors to household i (denoted by the set H_i). The assumed value for the parameter α must satisfy $0 \leq \alpha \leq 1$. Notice that equation (3.6) is a more general form of the usual K-nearest neighbor weighting scheme. When $\alpha = 1$, ω_{ih} simplifies to 1/K for those K-nearest neighbors. If $\alpha = 0$, then similar to before, for the K-nearest neighbors, equation (3.6) simplifies to the common FES weighting scheme used in meta-analyses (Borenstein et al., 2010).

To test whether the household-specific MWTP values are a result of a random spatial process, we apply Moran's I statistic to test for global spatial autocorrelation (Getis, 2010). The underlying expectation of Moran's I test is that the spatial process promoting the observed pattern of the attribute being analyzed is random. A rejection of the null hypothesis suggests that spatial autocorrelation, either spatial clustering or dispersion, exists. Moran's I statistic ranges from -1 to +1, with scores near +1 indicating spatial clustering and scores near –1 indicating spatial dispersion. Moran's I statistic is defined as:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{h=1}^{n}\omega_{ih}} \frac{\sum_{i=1}^{n}\sum_{h=1}^{n}\omega_{ih}(MWTP_i - \overline{MWTP})(MWTP_h - \overline{MWTP})}{\sum_{i=1}^{n}(MWTP_i - \overline{MWTP})^2} \tag{3.7}$$

where n is the number of households in the data sample, $MWTP_i$ and $MWTP_h$ are the household MWTP values, $\overline{MWTP} = \frac{\sum_{i=1}^{n} MWTP_i}{n}$ is the mean of all households' MWTP values in the sample, and ω_{ih} is the variance-adjusted spatial weight that links households i and h, as defined in equation (3.6). Notice that equation (3.7) is essentially the standard Moran's I statistic formula (Getis, 2010), with our adjusted spatial weights illustrated in equation (3.6).

While Moran's I test for global spatial autocorrelation provides a means to test for spatial patterns across the broader study area, little is revealed about local spatial heterogeneity

among households. Local indicators of spatial association (LISAs) have been developed to measure local spatial autocorrelation. Commonly referred to as hotspot analysis, the approach provides a statistical test for identifying spatial clusters of high values (hot spots) or low values (cold spots) of a variable of interest beyond what can be explained by random coincidence (Anselin, 1995). Among the most common LISAs is the Getis-Ord G^* statistic proposed by Getis and Ord (1992). Our proposed variance-adjusted Getis-Ord G^* statistic is calculated as follows:

$$G_i^* = \frac{\sum_{h=1}^{n}\{\omega_{ih} MWTP_h\} - \overline{MWTP}}{\sqrt{\frac{\sum_{h=1}^{n}\{MWTP_h{}^2\}}{n} - (\overline{MWTP})^2}\sqrt{\frac{n\sum_{h=1}^{n}\{\omega_{ih}{}^2\} - (\sum_{h=1}^{n}\omega_{ih})^2}{n-1}}} \tag{3.8}$$

Equation (3.8) is simply the standard Getis-Ord G^* statistic (Getis and Ord, 1992), but with our variance-adjusted spatial weights from equation (3.6).

3.3 Data

We demonstrate our two-step methodology by examining the spatial distribution of household MWTP for improvements in the Chesapeake Bay and freshwater lakes in the broader Chesapeake Bay Watershed. These MWTP values are estimated from data obtained in a stated choice study by Moore et al. (2018), which focused on reductions in nutrient and sediment pollution, and the resulting improvements in conditions for recreation and aquatic wildlife. Three choice questions were included in each survey. Each choice question presented respondents with a SQ alternative, showing current conditions and zero cost, and two policy alternatives with improvements in some or all of the attributes and a positive cost (see Table 3.1). The cost attribute was expressed as a permanent cost of living increase shown in annual terms for each household. The attributes defining each policy alternative were improvements in water clarity, striped bass population, blue crab population, and oyster abundance in the Bay. In addition to the Bay itself, freshwater lakes in the watershed benefit from the management

practices targeting the Bay. To capture these benefits, an additional attribute reflecting the number of lakes in the watershed with low algae growth was included in the alternatives presented in the survey. Through a series of ten focus groups, 72 one-on-one cognitive interviews, and an extensive pre-test, these attributes were found to be the most salient and important to the general population.

The survey was administered via mail to a geographically stratified random sample of households who reside in the District of Columbia or one of the 17 U.S. states that contain at least part of the Chesapeake Bay Watershed or lie within 100 miles of the Eastern coast of the U.S. The survey was sent to 2,829 households, and after adjusting for undeliverable addresses achieved a response rate of 31%. The resulting 671 surveys were screened for protest responses and hypothetical bias, leaving 559 completed surveys with which to estimate our models. Moore et al. (2018) provide more details on the study design and sample characteristics.

3.4 Results

3.4.1 Bayesian model results

There are three sets of results from the Bayesian mixed logit that are relevant to our objectives. The first are the summary statistics of the posterior distributions of our estimated model parameters, $g(\lambda|v)$. From a classical perspective these statistics can be interpreted as the coefficient estimates and standard errors, presented in the first two columns of results in Table 3.2. Given the inclusion of user-interaction terms in our model, the coefficients on the logged attribute levels are the marginal utilities for nonusers and the coefficients on the interaction terms are the differentials in marginal utilities for users of the corresponding resource. On average, water clarity in the Bay and striped bass population are not significant determinants of choice for non-users but they are for users that recreate in the Chesapeake Bay. Crab populations are significant for non-users, on average, while their contribution to marginal

utility is less for users but not significantly so. Water quality in watershed lakes is a significant attribute for the average nonuser and generates a significantly greater marginal utility for lake-users. The MWTP estimates are reported in the last column of Table 3.2. The MWTP values refer to a one-unit increase relative to the SQ quantities shown in Table 3.1; for example, a one-inch increase in clarity, a one-million increase in striped bass and crab populations, etc.

The second set of results describe how the household-level parameters are distributed in the population, shown in Table 3.3. In this case the mean values refer to the average household value and the standard deviation is an indication of how disperse the values in the population are, and not an indication of estimation precision. The corresponding MWTP values and the inner 90th percentile for each are shown in the right-hand side of Table 3.3. The subsequent spatial analyses examine how households' values in these distributions vary over space, after adjusting for statistical precision using the final set of Bayesian results.

The third and final set of Bayesian results relevant to our objectives are the household-level empirical distributions of MWTP that contain information on the central tendency and precision of our estimates at the household level, $\eta(\lambda_i|\boldsymbol{y_i}, z_i, \upsilon)$. Given the number of respondents in the data, it is not practical to report means and standard deviations for each.

3.4.2 Spatial interpolation analysis

Figures 3.1-3.6 present the MWTP results for Bass, Clarity, Crab, Lake, Oyster, and the SQ effects. In each figure we include two panels covering the results of the spatial interpolation (left) and hotspot analysis (right). Figure 3.7 shows the distribution of Getis-Ord G_i^* statistics across households in the hotspot analysis. For each set of maps, we present the results from the conventional spatial weights matrix and those from the variance-adjusted spatial weights matrix introduced in this study.

We first demonstrate the spatial interpolation results on "heat" maps, where darker shades identify higher estimates and lighter shades identify lower estimates. In general, visual

inspection of the interpolated MWTP surfaces suggest distinct spatial patterns in MWTP estimates for all attributes and spatial heterogeneity across the study area. For instance, we see some of the highest values for striped bass and water clarity among households living nearest to the Chesapeake Bay, or along the Atlantic Coast and just south of the Bay (left panels of Figures 3.1 and 3.2, respectively). We also observe relatively higher MWTP values for improvements in freshwater lakes in the Watershed (*Lake*) among households who live within the watershed (left panel, Figure 3.5). One unexpected spatial pattern, suggested by the left panel in Figure 3.3, is that households nearest the Chesapeake Bay hold lower values for improvements in the population of blue crabs (*Crab*), an iconic shellfish species in the Bay. Whereas households outside the watershed seem to hold higher values. As suggested by the Bayesian model results in Table 3.2, this could reflect the relatively large values among nonusers. Visually there is no clear global pattern in values for improved oyster abundance, as seen in the left panel of Figure 3.4. There are some areas near and just south of the Bay where residents hold relatively high values for increases in oyster populations, but there are other scattered clusters of high MWTP values (e.g., in New England, the gulf-side of Florida, and northwest Pennsylvania).

The interpolated surface of the SQ effects suggests an interesting spatial pattern. In West Virginia and in Northern New York near the Great Lakes and Finger Lakes, we see large positive values for the SQ option, suggesting that households in these areas generally hold a preference against any policy options leading to improvements in the Chesapeake Bay. One theory driving this result could be that households in these areas have their own substitute, and possibly equally as iconic, environmental amenities they care about, and thus have a bias against policy options leading to improvements in the Bay. Another possibility is that there is a protest or strategic response against increased regulations that was not completely eliminated by Moore et al.'s (2018) screening criteria based on responses to debriefing questions. In any

case, the potential biases captured by the SQ effects are controlled for and are excluded from the MWTP and welfare calculations.

Comparison of the conventional inverse distance interpolated MWTP maps ($\rho = 1$), to our variance-adjusted interpolations ($\rho = 0.5$) in the left panels of Figures 3.1-3.5 reveals some differences. In general, we find that the range of the interpolated MWTP estimates across the study area becomes smaller when we account for the variance of individual respondents' estimates. This suggests that accounting for the statistical precision of the first-stage estimates reduces the influence of outlying, often less precisely estimated, values. Although accounting for the variance of the first-stage estimates seems to diminish spatial variation, and reveals some local differences, the general spatial patterns appear similar.

One exception pertains to the interpolated SQ effect estimates. As shown in Figure 3.6, when accounting for the individual variances ($\rho = 0.5$), many of the areas exhibiting biases against a policy option remain, but there is now noticeably more evidence of respondents exhibiting a relatively high preference for a policy option (i.e., SQ effect < 0), irrespective of the improvements and costs defining the policy options. Such cold spots, for example, are now evident in the area around New York City, as well as in southern Florida and western Maryland. In any case, the interpolation exercise should be interpreted as suggestive at best. Although it visually depicts relevant spatial patterns, whether such patterns are statistically significant remains an open question. To answer that question, we turn to tests for global spatial autocorrelation and the hotspot analysis in sections 3.4.3 and 3.4.4, respectively.

3.4.3 Tests for global spatial autocorrelation

As can be seen in Table 3.4, the Moran's I tests for global spatial autocorrelation reveal broader spatial patterns for some of the MWTP estimates, but not all. MWTP for increases in striped bass and blue crab populations are both spatially correlated over the broader study area. However, such spatial trends are not generally revealed through parametric modelling of the

distance gradients (see Appendix C.2), thus highlighting the importance of considering potentially relevant spatial patterns revealed by non-parametric methods beyond just distance decay. The Moran's I test suggests that the SQ effects are also highly correlated over space.

These findings are robust as we vary the value of α. The strongest evidence of global spatial autocorrelation occurs when using the conventional weights ($\alpha = 1$) that do not take into account the statistical precision of the first-stage estimates. These global patterns remain robust but become less significant as we move towards the variance-adjusted weights ($\alpha = 0$). The Moran's I tests suggest no significant global spatial patterns in terms of MWTP for improvements in clarity, oyster abundance, and freshwater lakes. We next examine the nature of these global patterns and whether more local spatial patterns may exist that cannot be identified via the Moran's I statistic. In specific, we conduct local spatial associations analyses to test for the presence of statistically significant hot or cold spots using the Getis-Ord G_i^* statistic (Getis and Ord, 1992) and our variance-adjusted variant of the G_i^* statistic. We perform this analysis separately for each attribute, where the sampled households are the spatial units.

3.4.4 Spatial clustering analysis

The hotspot analysis results for the MWTP of each attribute are illustrated by the set of maps in the right panels of Figures 3.1-3.5. The maps show the status of each household (i.e., whether it is in a hot spot, cold spot, or demonstrates no statistically significant higher or lower MWTP values relative to other nearby households). These designations are based on the estimated G_i^* statistic for each household. The corresponding distributions of the G_i^* estimates across households are shown in Figure 3.7 for different values of α. This same information is shown in Table 3.5, which displays the number of households identified as being in a hot or cold spot.

G_i^* is assumed to be normally distributed under the null hypothesis (Getis and Ord, 1992), and so we fail to reject the null hypothesis when $-1.645 \leq G_i^* \leq 1.645$ (i.e., a

statistically insignificant result). Such instances on the maps would suggest that there is no clustering of high or low values around the corresponding household. Hot spots (black points on the figures) represent clusters of atypically high MWTP estimates, indicating a MWTP hot spot significant at the 90%, 95%, or 99% level depending on the size of the dot. Cold spots (white points) represent clusters of atypically low MWTP estimates, those with parallel negative G_i^* statistics indicating a MWTP cold spot at the same levels of significance, again varying by size of the white dot.

The hotspot analysis for *Bass* and *Clarity* reveals clusters of high MWTP values among some households in close proximity to the Bay (Figures 3.1 and 3.2, respectively). And for water clarity, we find clusters of systematically lower MWTP values among households near notable substitute waterbodies, like the Great Lakes and Finger Lakes in New York. Comparison to the variance-adjusted hotspot analyses when $\alpha = 0$ reveals similar results, but accounting for statistical precision in the underlying estimates reduces the number of households that belong to a statistically significant local cluster, especially for identified hot spots, as shown in Table 3.5.

In Figure 3.3, we find scattered hot spots of MWTP for crabs, mainly outside the watershed. There is also a concentration of cold spots within the watershed, mainly in central Virginia and Maryland. The finding that households in closest proximity to the Bay have the lowest values for improvements in crab populations, and those farthest have the highest values, is again surprising, but is in line with the interpolation exercise. This unexpected spatial pattern could be driven, at least partially, by relatively large values held by nonusers for this iconic resource. Comparing the conventional hotspot analysis ($\alpha = 1$) to our variance-adjusted hotspot analysis ($\alpha = 0$) reveals little difference, but substantially reduces the number of cold spots.

Up until this point we have found little evidence of discernible spatial patterns in oyster populations. The hotspot analysis in Figure 3.4, however, does suggest statistically significant clusters of high MWTP values for increases in oyster abundance, namely among those living closest to the Bay. There is a noticeable pattern of clustered low MWTP values, particularly around New York City and going North along the Hudson River (near the east-most border of New York state). Again, the variance-adjusted hotspot analysis ($\alpha = 0$) seems to reduce the number of hotspots, but we find slightly increased evidence of cold spots, as reported in Table 3.5.

The hotspot analyses for improvements in freshwater lakes in the broader Chesapeake Bay Watershed are displayed in the right panel of Figure 3.5. As one might expect, the conventional hotspot analysis reveals evidence of a concentration of statistically higher MWTP estimates among households living in the Watershed for improvements in freshwater lakes within the Watershed. There are also a few scattered cold spots, and most notably a concentration of lower MWTP values just outside the southwest corner of the Watershed; perhaps reflecting that there are several substitute lakes in western Virginia that are outside of the watershed. As suggested by the previous variance-adjusted hotspot analyses, we again see similar patterns in hot and cold spots, but the number of statistically significant local clusters of high MWTP estimates are substantially reduced when accounting for the statistical precision of the first-stage estimates ($\alpha = 0$).

The broader finding that the number of identified clusters are reduced after accounting for the statistical precision of the first-stage estimates is better demonstrated by the distributions of the household-specific G_i^* statistics shown in Figure 3.7. In general, we see that accounting for the statistical precision of the MWTP estimates makes one less likely to identify statistically significant clusters (i.e., a larger portion of the distribution of G_i^* is located towards zero). This finding is consistent with the hotspot maps in Figures 3.1-3.6 and makes intuitive sense.

Extreme MWTP estimates are often less precise, and so when these estimates are appropriately discounted due to this lack of precision one is less likely to falsely identify a statistically significant cluster of high or low WTP values. When performing scoping exercises like this to try and identify spatial patterns, this application demonstrates that it may be important to account for the fact that these MWTP values are estimates, and not observed values. Not taking into account the precision of the household MWTP estimates may lead researchers to falsely identify patterns of spatial heterogeneity.

The bottom panel in Figure 3.7 reveals a finding that is unique to the estimated SQ effects. Incorporating household-specific variances into the spatial weights shifts the mass of the G_i^* distributions for the MWTP estimates towards zero. However, for the G_i^* statistics corresponding to the SQ effects, we see the distribution shift more negative. In some cases, as we previously saw, this reduces the number of identified hot spots. For example, the conventional hotspot results ($\alpha = 1$) in Figure 3.7 suggest that respondents near the Finger Lakes, a notable substitute, are more susceptible to exhibiting potentially biasing behaviors against a policy option that improves water quality in the more distant Chesapeake Bay (e.g., "cold feet"). But this identified cluster of significantly high SQ effect estimates disappears once the statistical precision of those underlying estimates is accounted for.

The more unique finding is that accounting for the statistical precision behind the estimated SQ effects identifies more statistically significant cold spots. In other words, we identify more clusters of households exhibiting "warm glow" or other potentially biasing behaviors in favor of a policy option. For example, in Table 3.5 we see a 230% increase in the number of households that belong to a SQ effect cold spot when going from $\alpha = 1$ to $\alpha = 0$. This is also evident in the maps in Figure 3.6. The rightmost hotspot analysis in Figure 3.6 where $\alpha = 0$ reveals noticeably more cold spots, especially in western Maryland, and the area around New York City, the Long Island Sound, and Narragansett Bay and Cape Cod. Perhaps

respondents near these other iconic estuaries have an implicit bias or strategic response that pushes them towards an option that leads to improvements in the Chesapeake Bay. Alternatively, this may reflect preference heterogeneity in favor of estuary quality that is not otherwise captured by or correlated with other attributes in the experimental design. In any case, these effects are excluded from the policy illustration we discuss next, but the location and clustering of respondents exhibiting such behaviors are important to keep in mind when designing future stated preference studies and could be important to identify prior to specifying parametric models to estimate welfare changes. For example, the results of our spatial analyses allude to the potential importance of substitute waterbodies. Insights like this are an example of what can be gained from semi-parametric scoping exercises to examine spatial patterns.

3.4.5 Policy illustration: Chesapeake Bay Total Maximum Daily Loads

To examine the potential policy implications of accounting for spatial heterogeneity in WTP estimates, and the underlying statistical precision of those estimates, we repeat the benefit calculations carried out by Moore et al. (2018). As reported in Table 3 of Moore et al. (2018), the projected improvements from the Chesapeake Bay TMDLs are an average increase of 4.33 inches in Bay water clarity, 1.03 million striped bass, 41 million blue crab, 541 tons of oysters, and 455 freshwater lakes reaching "low algae" status. Following Holmes and Adamowicz (2003), the annual WTP for each household i is calculated as the difference of the deterministic component of the indirect utility, divided by the marginal utility of income:

$$WTP_i = \frac{\beta_i ln(\boldsymbol{x^1}) + \delta_i ln(\boldsymbol{x^1}) user_i - [\beta_i ln(\tilde{\boldsymbol{x}}) + \delta_i ln(\tilde{\boldsymbol{x}}) user_i]}{-\gamma} \tag{3.9}$$

where $\tilde{\boldsymbol{x}}$ is a vector of the baseline attribute levels (Table 3.1) and $\boldsymbol{x^1}$ are the projected policy levels (baseline levels plus the improvements). The household-specific WTP estimates are derived for all 559 households in the sample. In order to extrapolate the WTP estimates to the population of 44,353,441 households in the study area, we then create interpolated WTP

surfaces using the same procedure described in section 3.2.4.1. We next take the average of the interpolated cells within each census tract and multiply that by the number of households in that tract according to the 2010 U.S. Census. These total WTP estimates for each tract are then summed over all census tracts in the study area. The resulting total annual WTP estimates are displayed in Table 3.6. We emphasize that our policy illustration is based solely on the spatial interpolations and does not rely on the analyses of global spatial autocorrelation and local clusters.

The first two columns in Table 3.6 show the total benefit estimates taken from Moore et al. (2018). Their model 1 estimates assume homogeneity across the entire study area by applying a single average WTP estimate to all households in the population. Their model 2 estimate is based on a similar procedure, and although it does not explicitly account for spatial heterogeneity, it does account for heterogeneity regarding the use of the resource and extrapolates those values based on estimates of the proportion of the population that are users versus nonusers. The next four columns in Table 3.6 show the results of our spatially explicit extrapolation exercise, and suggest total benefit estimates for the entire study area ranging from \$6.6 to \$6.9 billion per year. These total benefit estimates are largely in line with those from Moore et al. (2018). This suggests that accounting for spatial heterogeneity may not yield substantial differences in terms of broader policy implications, at least not in this specific context and when considering the entire study area as a whole. Such spatially explicit details may be important, however, for more local policies. For example, we do find significant variation in household-level WTPs across the study area, ranging from an annual household WTP of \$23 to \$312.

Comparison of the total WTP estimates from the conventional spatial interpolation exercise ($\rho = 1$) to our variance-adjusted spatial interpolations ($\rho < 1$) suggests that total WTP estimates decrease as more weight is given to the statistical precision of the first-stage

estimates. This is consistent with our broader findings that accounting for statistical precision reduces the influence of less precisely estimated outliers that could otherwise unduly influence empirical analyses. In this particular context, however, the differences in total WTP inferred from the spatial interpolations may not be economically significant. Surprisingly, relatively small differences are also revealed when examining total WTP at more local levels, such as by state or county. In fact, even at the individual tract-level, comparing our variance-adjustment estimates when $\rho = 0.5$ to the benefits inferred from conventional interpolation techniques ($\rho = 1$), suggests that the latter leads to only a 10% difference in total tract-level WTP for the majority (90%) of the 27,117 census tracts in the study area. In short, although we find that accounting for spatial heterogeneity is important, the proposed variance-adjustment may not make much of a practical difference in this particular setting.

3.5 Conclusions

We propose a novel extension of existing semi-parametric techniques to analyze spatial patterns when the variables of interest are estimates and not observed values, as is the case in many applications to nonmarket valuation. When examining spatial welfare patterns, we account for the fact that our first-stage model will estimate some households' values less precisely than others. The methodology in this study estimates household-specific MWTP variances using Bayesian estimation techniques and incorporates that information into the spatial weights matrix used in tests for global spatial autocorrelation, spatial interpolation maps, and hotspot analyses. Similar spatial analyses have been increasingly introduced in the nonmarket valuation literature (e.g., Campbell et al., 2009; Meyerhoff, 2013; Johnston and Ramachandran, 2014; Johnston et al., 2015; Czajkowski et al., 2017), but our study is the first to formally incorporate the statistical precision of the first-stage estimates into the second-stage spatial analyses.

We demonstrate our two-step methodology using a SP study of water quality improvements in the Chesapeake Bay. Accounting for the statistical precision of the MWTP estimates generally seems to result in less statistically significant evidence of spatial patterns, as reflected by tests for global spatial autocorrelation and the hotspot analysis. This tendency increases as additional weight is given to the statistical precision of the MWTP estimates. A similar finding is found with regards to households exhibiting "cold feet," or a tendency to disproportionally favor the status quo. Overall, the analysis suggests that accounting for the statistical precision of the estimated economic phenomena being analyzed may reduce the chances of falsely identifying statistically significant spatial heterogeneity. In contrast, we also find that accounting for the statistical precision of the underlying estimates can lead to increased identification of areas where respondents disproportionally favor a policy option for reasons not explained by the choice attributes. Identifying locations where households exhibit such "warm glow" or other potentially biasing behaviors can aid in future survey design and help inform econometric model specifications to estimate welfare changes.

We estimate the total benefits projected to result from the Chesapeake Bay TMDLs to examine the importance of our extension of traditional spatial analyses from a practical standpoint. We find that in a broader regional setting, at least with our data, the difference between the benefits inferred from traditional spatial interpolation techniques versus those that accommodate for statistical precision are small. More applications of the methods discussed in this study to data valuing other environmental amenities are needed to see whether accounting for the statistical precision of the first-stage WTP estimates reveals similar findings, particularly in cases where the environmental amenities of interest are more local in nature. One might not necessarily expect as much spatial heterogeneity in preferences for a well-known iconic resource, like the Chesapeake Bay. Examination of more localized amenities,

perhaps where familiarity with the resource is more varied, may yield different findings in how accounting for statistical precision impacts the identification of spatial patterns.

Although our two-step methodology provides an intuitive path for accounting for the statistical precision of the first-stage estimates when conducting spatial analyses, and presumably allows for more accurate identification of spatial patterns, future simulation studies are needed to formally examine the potential improvements in accuracy. Such studies might entail analysis of simulated data where the researcher knows the true data generating process over space. Nonetheless, given the emphasis of these spatial analytic techniques mainly for purposes of data diagnostics and scoping (Johnston and Ramachandran, 2014; Johnston et al., 2015), we encourage researchers to implement our variance-adjustment methods when attempting to identify potential spatial patterns that may not be immediately apparent through conventional parametric models.

Table 3.1: Attribute Descriptions and Levels

Attribute	Description	Status Quo Level	Post-policy Levels
Bay Water Clarity	Number of feet below the water surface you can see	3 feet	3; 3.5; 4.5
Striped Bass Population	Number of adult striped bass in the Chesapeake Bay (millions)	24 million	24; 30; 36
Blue Crab Population	Number of adult blue crab in the Chesapeake Bay (millions)	250 million	250; 285; 328
Oyster Abundance	Tons of oysters living in the Chesapeake Bay	3,300 tons	3,300; 5,500; 10,000
Low Algae Lakes	Out of 4,200 freshwater lakes in the Chesapeake Bay Watershed, number with low algae levels	2,900 lakes	2,900; 3,300; 3,850
Annual Cost to Household	Permanent increase in the annual cost of living starting the following calendar year	$0 per year	$20; $40; $60; $180; $250; $500

Table 3.2: Posterior Distributions of Coefficient Estimates

	Mean		SD	Non-User MWTP
ln(clarity) [a]	0.505		0.548	1.47
ln(bass)	0.874		0.584	3.83*
ln(crab)	2.070	***	0.638	0.87***
ln(oyster)	0.198		0.206	0.01
ln(lake)	3.769	***	0.824	0.14***
				User MWTP
user × ln(clarity)	1.101	***	0.720	4.69***
user × ln(bass)	0.615	**	0.863	6.52***
user × ln(crab)	-0.366		0.933	0.72
user × ln(oyster)	0.355		0.320	0.02
user × ln(lake)	1.305	*	1.069	0.18**
SQC	-1.938	***	0.352	
cost	-0.009	***	0.001	

Note: All coefficient estimates modelled as random, except the coefficient on cost is treated as fixed to ensure MWTP distributions are finite. Marginal willingness to pay (MWTP) estimates expressed in 2014$. (a) Note that clarity is expressed as inches in the empirical models and subsequent MWTP estimates. All other environmental attributes are expressed in the same units originally specified in the survey, and as reported in Table 3.1 (i.e., millions of bass, millions of crabs, tons of oysters, and the number of low algae lakes). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3.3: Distribution of Coefficients and MWTP in the Population

	Mean	SD	Mean Non-User MWTP	Inner 90th Percentile MWTP	
ln(clarity)	0.4327	3.0252	1.38	-5.69	9.52
ln(bass)	0.5519	2.2868	3.84	1.63	6.92
ln(crab)	2.058	0.6847	0.87	0.74	1.03
ln(oyster)	0.2532	1.3006	0.01	-0.01	0.03
ln(lake)	3.6197	1.3837	0.14	0.08	0.24
			Mean User MWTP		
user × ln(clarity)	1.5235	1.6738	4.91	-4.55	14.34
user × ln(bass)	1.0359	1.2955	6.55	0.37	13.35
user × ln(crab)	-0.9166	1.4795	0.72	0.44	0.99
user × ln(oyster)	0.0865	0.6691	0.02	-0.08	0.26
user × ln(lake)	1.4167	2.1951	0.17	0.07	0.27

Note: Marginal willingness to pay (MWTP) estimates expressed in 2014$. MWTP for an increase in Bay water clarity is expressed in inches. MWTP for all other attributes are expressed in the same units originally specified in the survey, and as reported in Table 1 (i.e., millions of bass, millions of crabs, tons of oysters, and the number of low algae lakes).

Table 3.4: Moran's I Tests for Global Spatial Autocorrelation

	$\alpha = 1.0$		$\alpha = 0.5$		$\alpha = 0.0$	
	Moran's *I*	z-score	Moran's *I*	z-score	Moran's *I*	z-score
Clarity	-0.0047	-0.1408	-0.0056	-0.1824	-0.0058	-0.1947
Striped Bass	0.0459	2.3157**	0.0419	2.1029**	0.0355	1.7525*
Blue Crab	0.0577	2.8897***	0.0533	2.6599***	0.0476	2.3411**
Oysters	0.0305	1.5682	0.0300	1.5368	0.0294	1.4872
Lakes	0.0194	1.0276	0.0163	0.8737	0.0131	0.7116
SQ Effect	0.0726	4.0285***	0.0705	3.692***	0.0708	3.2323***

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

Table 3.5: Number of Households Identified as Being in a Hot or Cold Spot

	Numbers of Hot Spots			Numbers of Cold Spots		
	$\alpha = 1.0$	$\alpha = 0.5$	$\alpha = 0.0$	$\alpha = 1.0$	$\alpha = 0.5$	$\alpha = 0.0$
Clarity	24	21	16	20	21	20
Striped Bass	62	42	33	1	2	1
Blue Crab	34	41	46	46	35	20
Oysters	49	38	27	23	24	31
Lakes	47	35	16	21	19	23
SQ Effect	57	37	24	52	132	172

Note: The displayed counts show the number of households (out of the sample of 559) that are identified as being part of a spatial cluster of statistically higher (hot spot) or lower (cold spot) values. Hot spots are those with $G_i^* \geq 1.645$ and cold spots are those with $G_i^* \leq -1.645$.

Table 3.6: Total Annual Willingness to Pay for Improvements under Total Maximum Daily Loads (2014$, billions)

Moore et al. (2018)		Spatial Interpolation in this Study			
Model 1	Model 2	$\rho = 1.00$	$\rho = 0.75$	$\rho = 0.50$	$\rho = 0.25$
$6.813	$6.488	$6.870	$6.790	$6.711	$6.635

Figure 3.1: MWTP for Bass

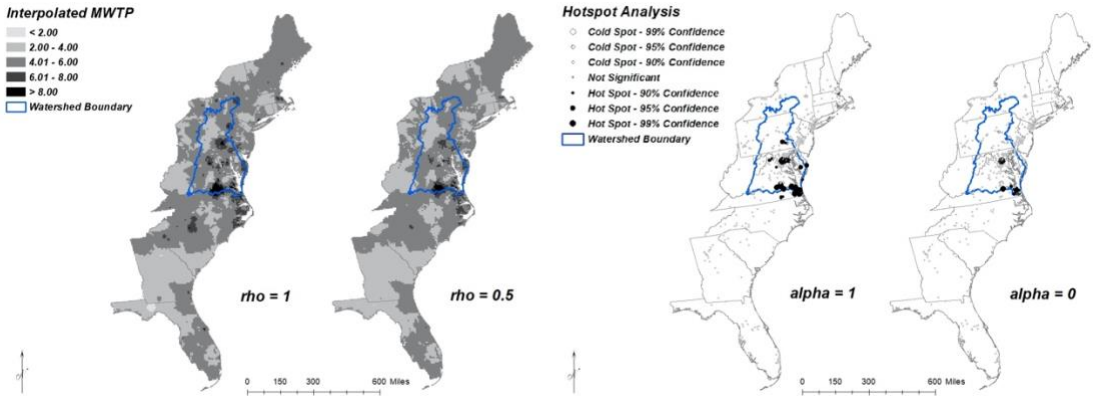


Figure 3.2: MWTP for Clarity

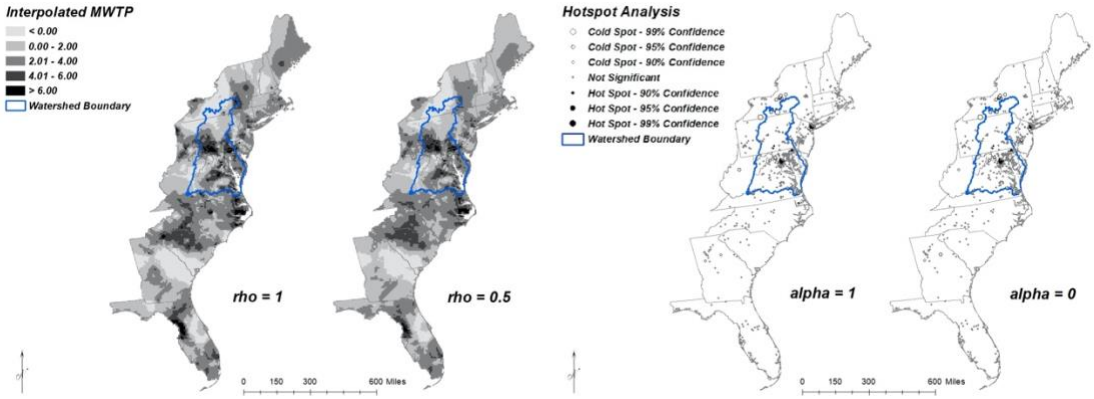


Figure 3.3: MWTP for Crabs

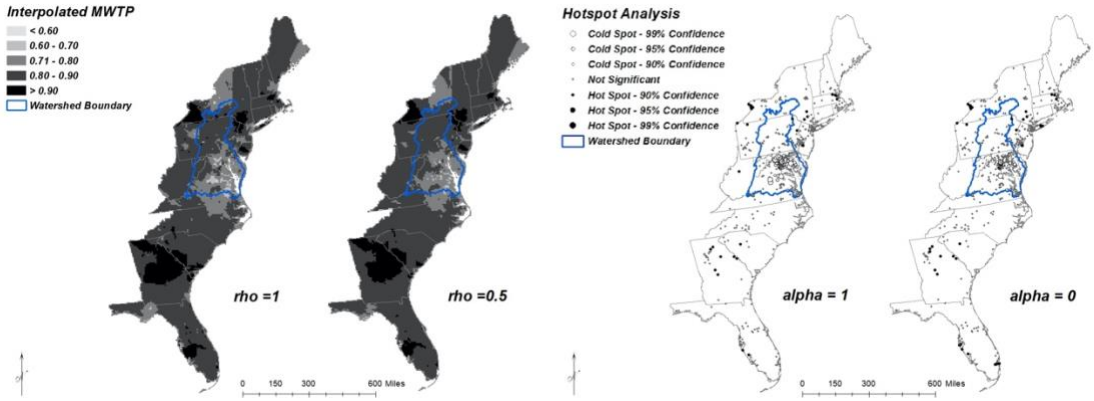


Figure 3.4: MWTP for Oysters

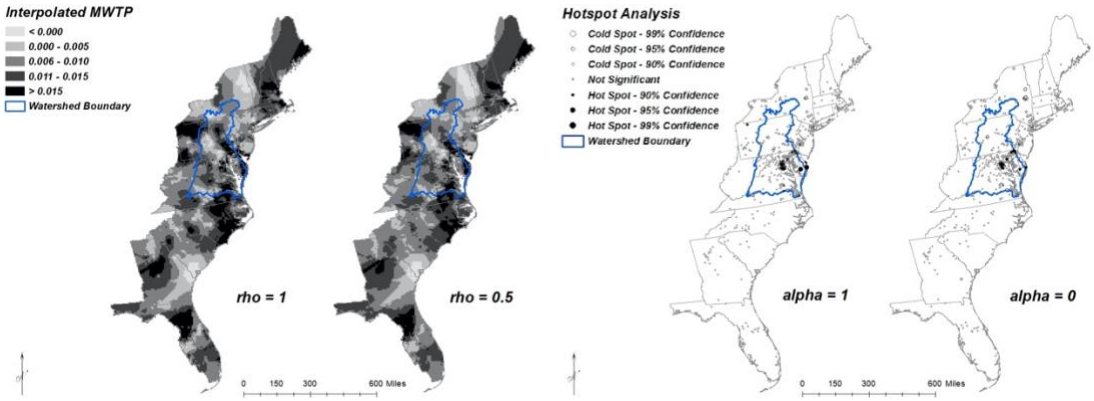


Figure 3.5: MWTP for Lakes

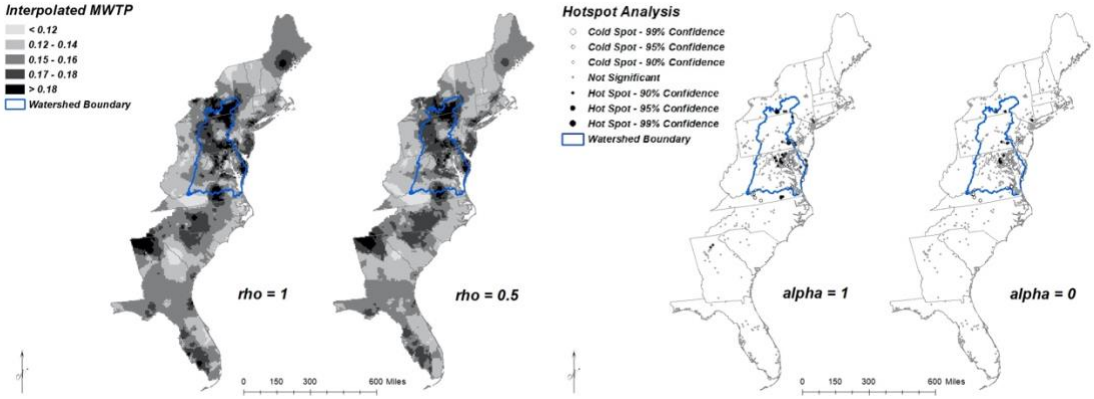


Figure 3.6: Status Quo Effects

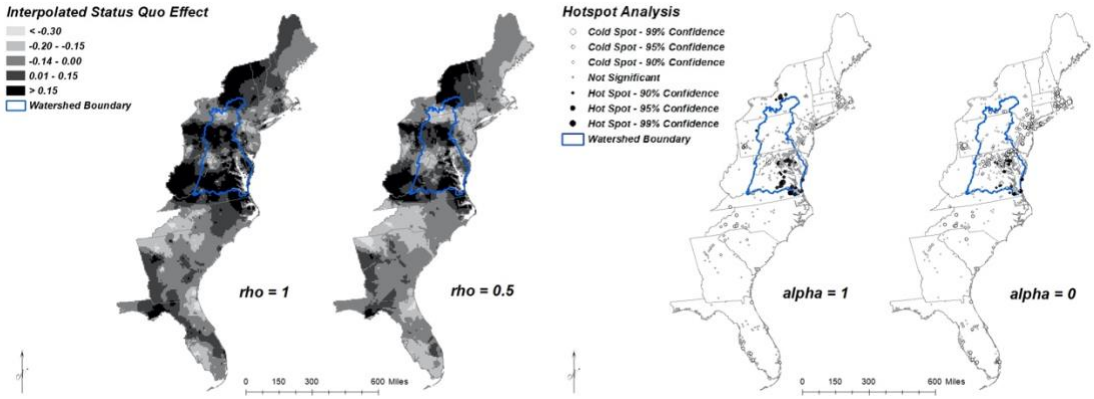
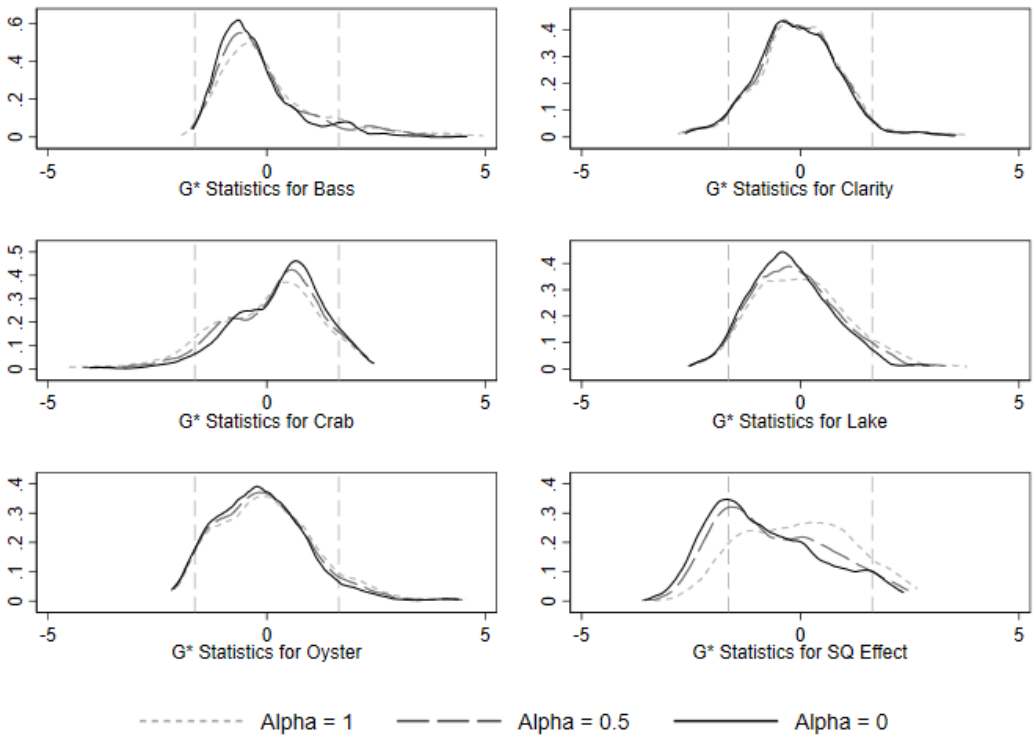


Figure 3.7: Distribution of Hotspot Analysis Getis-Ord Statistics



Note: Vertical long-dashed grey lines denote 90% confidence interval.

Appendix A: Appendix to Chapter 1

A.1 A Two-agent model with Spatial Externality in a Socially Optimal Case

Now consider two counties $k-1$ and k such that county $k-1$ is the upstream county and county k is the downstream county. Levee building in the upstream county may increase the occurrence of flooding in the downstream county. As a result, the probability of flooding occurrence in county k becomes $\Phi_{k,t}(1+\gamma L_{k-1,t})$. Instead of each county choosing its own levee building to minimize the sum of the expected value lost due to flooding and the cost of levee building on its own, the social planner chooses $L_{.,t}$ to minimize the sum of the total expected value lost due to flooding and the total cost of levee building, C_t:

$$\begin{aligned} C_t = \Phi_{k-1,t}V_{k-1,t} - \Phi_{k-1,t}V_{k-1,t}\left(\frac{L_{k-1,t}}{\mathcal{L}_{k-1}}\right)^a + \Phi_{k,t}\left(1+\gamma L_{k-1,t}\right)V_{k,t} \\ - \Phi_{k,t}(1+\gamma L_{k-1,t})\, V_{k,t}\left(\frac{L_{k,t}}{\mathcal{L}_k}\right)^a + AL_{k-1,t}^b + AL_{-k,t}^b \end{aligned} \tag{A.1}$$

A socially optimal interior solution satisfies the following two first-order constraints:

$$\frac{\partial C_t}{\partial L_{k-1,t}} = 0 \Rightarrow \frac{a\Phi_{k-1,t}V_{k-1,t}L_{k-1,t}^{a-1}}{\mathcal{L}_{k-1}^a} = bAL_{k-1,t}^{b-1} + \frac{\gamma\Phi_{k,t}V_{k,t}\left(1-L_{k,t}^a\right)}{\mathcal{L}_k^a} \tag{A.2}$$

$$\frac{\partial C_t}{\partial L_{k,t}} = 0 \Rightarrow \frac{a\Phi_{k,t}\left(1+\gamma L_{k-1,t}\right)V_{k,t}L_{k,t}^{a-1}}{\mathcal{L}_k^a} = bAL_{k,t}^{b-1} \tag{A.3}$$

Equations (A.2) and (A.3) respectively define the optimal levee heights in county $k-1$ and k where $L_{k-1,t}^{**} < L_{k-1,t}^{*} = \left(\frac{a\Phi_{k-1,t}V_{k-1,t}}{bA\mathcal{L}_{k-1}^a}\right)^{\frac{1}{b-a}}$ and $L_{k,t}^{**} = \left[\frac{a\left(\Phi_{k,t}+\gamma L_{k-1,t}^{**}\right)V_{k,t}}{bA\mathcal{L}_k^a}\right]^{\frac{1}{b-a}} < L_{k,t}^{*}$ are socially optimal levee heights and $L_{k-1,t}^{*}$ and $L_{k,t}^{*}$ are privately optimal levee heights as shown in Section 1.3. We can see that compared to the privately optimal case, both counties' optimal levee heights are lower in the socially optimal case.

A.2 A Multiple-agent Model with Spatial Externality

Now consider N ($N \geq 2$) counties such that county j is the downstream county of the other $N-1$ counties. Levee building in all the upstream counties k ($k \neq j$) may increase the occurrence of flooding in the downstream county. As a result, the probability of flooding occurrence in county j changes from Φ_{jt} to $\Phi_{jt} \prod_{k \neq j}^{N}(1+\gamma_k L_{kt})$. This assumes that if one of the upstream counties does not have levee building, its contribution to the downstream county's flooding occurrence becomes zero. Each county chooses its own levee building $L_{.,t}$ to minimize the sum of the expected value lost due to flooding and the cost of levee building, $C_{.,t}$:

$$C_{kt} = \Phi_{kt} V_{kt} - \Phi_{kt} V_{kt} \left(\frac{L_{kt}}{\mathcal{L}_k}\right)^a + A L_{kt}^b \tag{A.4}$$

$$C_{jt} = \Phi_{jt} \prod_{k \neq j}^{N} (1+\gamma_k L_{kt})\, V_{jt} - \Phi_{jt} \prod_{k \neq j}^{N} (1+\gamma_k L_{kt})\, V_{jt} \left(\frac{L_{jt}}{\mathcal{L}_j}\right)^a + A L_{jt}^b \tag{A.5}$$

Take the derivative with respect to levee height for each county and re-arrange:

$$L_{kt}^* = \left(\frac{a \Phi_{kt} V_{kt}}{b A \mathcal{L}_k^a}\right)^{\frac{1}{b-a}} \tag{A.6}$$

$$L_{jt}^* = \left[\frac{a \Phi_{jt} \prod_{k \neq j}^{N} (1+\gamma_k L_{kt})\, V_{jt}}{b A \mathcal{L}_j^a}\right]^{\frac{1}{b-a}} \tag{A.7}$$

Equation (A.6) defines the privately optimal levee height in the upstream county k. Substituting L_{kt}^* from equation (A.6) into equation (A.7), we get the privately optimal levee height in the downstream county j:

$$L_{jt}^* = \left(\frac{a \Phi_{jt} V_{jt}}{b A \mathcal{L}_j^a}\right)^{\frac{1}{b-a}} \left(\prod_{k \neq j}^{N} \left[1 + \underbrace{\gamma_k \left(\frac{a \Phi_{kt} V_{kt}}{b A \mathcal{L}_k^a}\right)^{\frac{1}{b-a}}}_{spatial\ externality}\right]\right)^{\frac{1}{b-a}} \tag{A.8}$$

Equation (A.8) shows that compared to the single-agent case, the privately optimal levee height in the downstream county is higher if there is spatial externality from any of the upstream counties' levee building.

A.3 A Two-agent Model with Spatial Externality (Cross-river Case)

Now consider N ($N \geq 2$) counties such that county j is the downstream county of the other $N-1$ counties. Levee building in all the upstream counties k ($k \neq j$) may increase the occurrence of flooding in the downstream county. Now consider two counties k and $-k$ such that these two counties are cross-river neighbors. Levee building on either side of the river may increase the occurrence of flooding on the other side. As a result, the probability of flooding occurrence in each county becomes $\Phi_{k,t}(1+\gamma L_{-k,t})$ and $\Phi_{-k,t}(1+\gamma L_{k,t})$, respectively for county k and $-k$. Each county chooses its own levee building $L_{.,t}$ to minimize the sum of the expected value lost due to flooding and the cost of levee building, $C_{.,t}$:

$$C_{k,t} = \Phi_{k,t}(1+\gamma L_{-k,t})\, V_{k,t} - \Phi_{k,t}(1+\gamma L_{-k,t})\, V_{k,t}\left(\frac{L_{k,t}}{\mathcal{L}_k}\right)^a + AL_{k,t}^b \tag{A.9}$$

$$C_{-k,t} = \Phi_{-k,t}(1+\gamma L_{k,t})\, V_{-k,t} - \Phi_{-k,t}(1+\gamma L_{k,t})\, V_{-k,t}\left(\frac{L_{-k,t}}{\mathcal{L}_{-k}}\right)^a + AL_{-k,t}^b \tag{A.10}$$

Take the derivative with respect to levee height for each county and re-arrange:

$$L_{k,t}^* = \left(\frac{a\Phi_{k,t}V_{k,t}}{bA\mathcal{L}_k^a}\right)^{\frac{1}{b-a}}\left[1+\underbrace{\gamma L_{-k,t}^*}_{spatial\ externality}\right]^{\frac{1}{b-a}} \tag{A.11}$$

$$L_{-k,t}^* = \left(\frac{a\Phi_{-k,t}V_{-k,t}}{bA\mathcal{L}_{-k}^a}\right)^{\frac{1}{b-a}}\left[1+\underbrace{\gamma L_{k,t}^*}_{spatial\ externality}\right]^{\frac{1}{b-a}} \tag{A.12}$$

Equations (A.11) and (A.12) respectively define the privately optimal levee heights in the two counties. We can observe that compared to the single-agent case, the optimal levee height in each county is higher if there is spatial externality from the cross-river county's levee building.

A.4 Additional Tables

Table A.1: List of Years When Levee Elevations Were Measured

Pre-flood Period 2		Pre-flood Period 1		Post-flood Period 1		Post-flood Period 2	
Year	Percent	Year	Percent	Year	Percent	Year	Percent
1982	2.96	1982	5.04	2012	73.02	2014	4.16
1983	10.73	1983	5.60	2013	11.37	2015	78.06
1984	7.05	1984	2.48	2014	10.81	2016	2.16
1985	18.99	1985	10.89	2015	4.80	2017	15.37
1986	10.89	1986	9.29	-	-	2018	0.24
1987	0.40	1987	0.16	-	-	-	-
1988	0.08	1989	4.72	-	-	-	-
1989	2.56	1990	7.21	-	-	-	-
1990	2.08	1991	2.48	-	-	-	-
1991	2.56	1992	0.80	-	-	-	-
1992	2.64	1993	5.68	-	-	-	-
1993	3.52	1994	5.28	-	-	-	-
1994	1.52	1995	0.88	-	-	-	-
1995	0.08	1996	5.52	-	-	-	-
1996	0.40	1997	0.32	-	-	-	-
1997	0.48	1998	12.25	-	-	-	-
1998	0.16	1999	0.80	-	-	-	-
N/A	32.99	2000	1.68	-	-	-	-
-	-	2002	0.16	-	-	-	-
-	-	2003	1.92	-	-	-	-
-	-	2009	1.92	-	-	-	-
-	-	2010	16.65	-	-	-	-
Total	100	Total	100	Total	100	Total	100

Notes: This table lists the years when levee elevations were measured using the USGS topographic maps for the two pre-flood periods (i.e., pre-flood periods 1 and 2) and the DEMs for the two post-flood periods (i.e., post-flood periods 1 and 2). N/A denotes that measurements are not available. For the levees with elevations not available in the pre-flood period 2, the majority of them have elevations measured prior to 1986 in the pre-flood period 1.

Table A.2: Impact of Flooding on Levee Elevation by Region

Variable	Logged levee elevation			
	Model I	Model II	Model III	Model IV
	Panel A: Upper Mississippi River Basin			
Flooded × Post	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
	[0.001]	[0.001]	[0.001]	[0.001]
Mean of dependent variable	6.285	6.285	6.285	6.285
Number of counties	115	115	114	114
Number of observations	1,700	1,700	1,698	1,698
	Panel B: Upper Mississippi River Basin			
Flooded × Post	0.020***	0.027***	0.027***	0.027***
	(0.007)	(0.007)	(0.004)	(0.004)
	[0.007]	[0.007]	[0.004]	[0.004]
Mean of dependent variable	3.892	3.892	3.935	3.935
Number of counties	92	92	92	92
Number of observations	798	798	770	770

Notes: Panel A and B in this table show the results from estimating equation (1.7), respectively for the sample in the upper and lower Mississippi River Basin. The dependent variable is the logged levee elevation (ft). Flooded × Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county level, are in parentheses. Standard errors, adjusted to allow for spatial autocorrelation as modeled in Conley (1999) with a distance cutoff at 50 km, are in brackets. *** $p<0.01$

Table A.3: Impact of Flooding on Levee Elevation After 1993

Variable	Logged levee elevation			
	Model I	Model II	Model III	Model IV
Flooded × Post	0.004***	0.004***	0.008**	0.008**
	(0.001)	(0.001)	(0.003)	(0.003)
	[0.001]	[0.001]	[0.003]	[0.003]
Mean of dependent variable	5.285	5.285	5.337	5.337
Number of counties	133	133	132	132
Number of observations	1,270	1,270	1,242	1,242
Levee controls		✓		✓
County controls			✓	✓
County fixed effects	✓	✓	✓	✓
State-by-year fixed effects	✓	✓	✓	✓

Notes: This table shows the results from estimating equation (1.7) using the data sample that excludes any levees whose elevations were measured prior to 1993. The dependent variable is the logged levee elevation (ft). Flooded °ø Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county level, are in parentheses. Standard errors, adjusted to allow for spatial autocorrelation as modeled in Conley (1999) with a distance cutoff at 50 km, are in brackets. *** p<0.01; ** p<0.05

Table A.4: Robustness Checks on Impact of Flooding on Levee Elevation

Variable	Logged levee elevation			
	Model I	Model II	Model III	Model IV
Flooded × Post	0.007***	0.007***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)
Mean of dependent variable	5.520	5.520	5.549	5.549
Number of counties	207	207	206	206
Number of observations	2,498	2,498	2,498	2,498
Levee controls		✓		✓
County controls			✓	✓
County fixed effects	✓	✓	✓	✓
State-by-year fixed effects	✓	✓	✓	✓

Notes: This table shows the results from estimating equation (1.7). The dependent variable is the logged levee elevation (ft). Flooded °ø Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county by year level, are in parentheses for all models. *** p<0.01

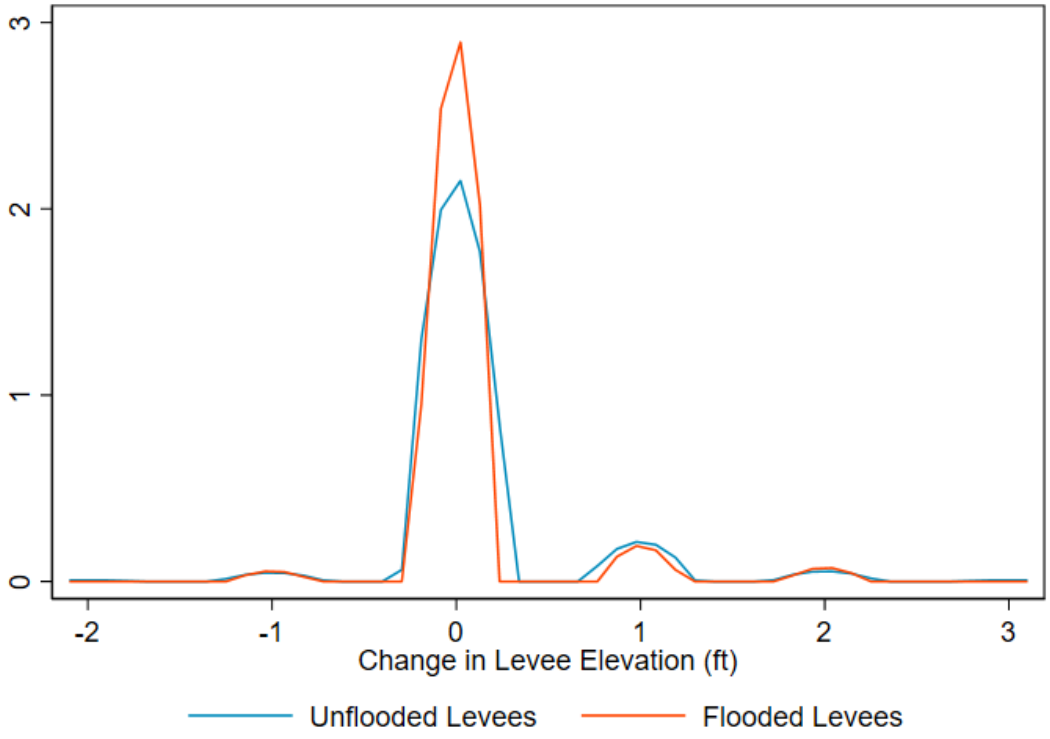
Table A.5: Impact of Flooding on Levee Elevation with Spillover Effects

Variable	Logged levee elevation			
	Model I	Model II	Model III	Model IV
Flooded × Post	0.005***	0.005***	0.008**	0.008**
	(0.003)	(0.003)	(0.004)	(0.004)
log(upstream levee elevation)	0.733***	0.733***	0.727***	0.727***
	(0.180)	(0.180)	(0.206)	(0.206)
log(cross-river levee elevation)	-0.149	-0.149	0.014	0.014
	(0.346)	(0.346)	(0.264)	(0.264)
log(upstream-cross-river levee elevation)	-0.187	-0.187	0.190	0.190
	(0.374)	(0.374)	(0.347)	(0.347)
Mean of dependent variable	5.520	5.520	5.549	5.549
Number of counties	207	207	206	206
Number of observations	2,498	2,498	2,498	2,498
Levee controls		✓	✓	✓
County controls				✓
County fixed effects	✓	✓	✓	✓
State-by-year fixed effects	✓	✓	✓	✓

Notes: This table shows the results from estimating equation (1.8) with cross-river and upstream-cross-river neighbors. The dependent variable is the logged levee elevation (ft). Flooded × Post is equal to 1 if the levee was flooded in 2011 and the levee elevation was measured after 2011. log(upstream levee elevation), log(cross-river levee elevation), and log(upstream-cross-river levee elevation) are respectively the logged average elevations of levee(s) in the county's immediate upstream, cross-river, and upstream-cross-river county. Levee controls include levee length, the number of levee segments, and whether the levee is located in a 100-year floodplain (500-year floodplain as baseline). County controls include logged farm acreage, logged farmland value, logged total agricultural sales, logged median housing value, and flood counts. Standard errors, clustered at the county level, are in parentheses. *** p<0.01; ** p<0.05

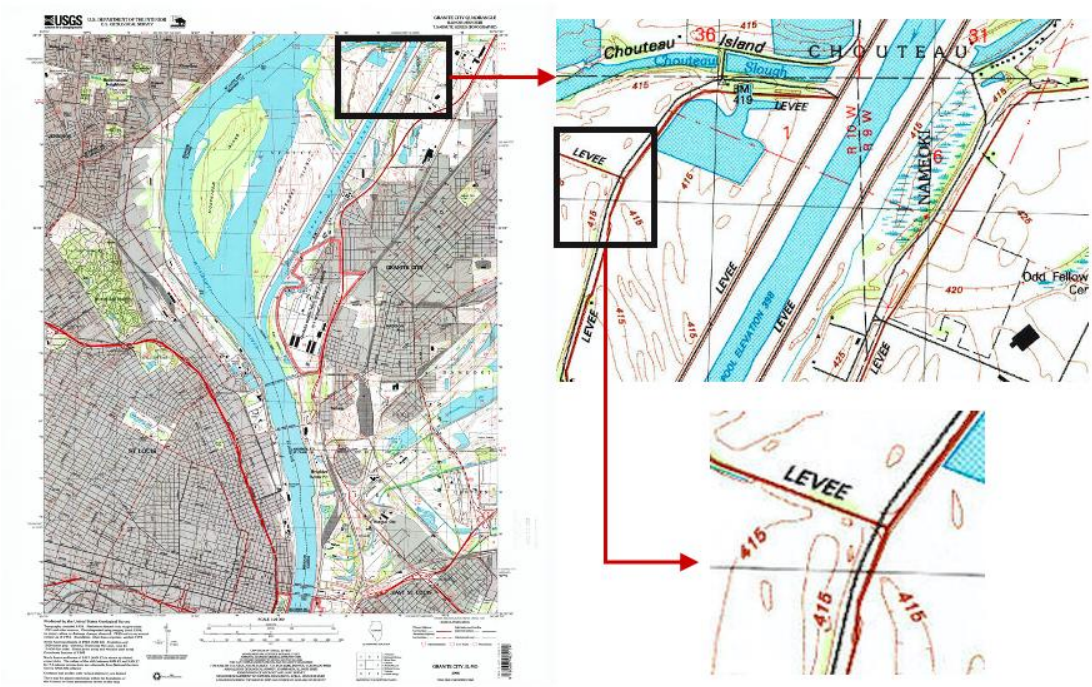
A.5 Additional Figures

Figure A.1: Kernel Density Plot of the Change in Levee Elevations Pre-flood



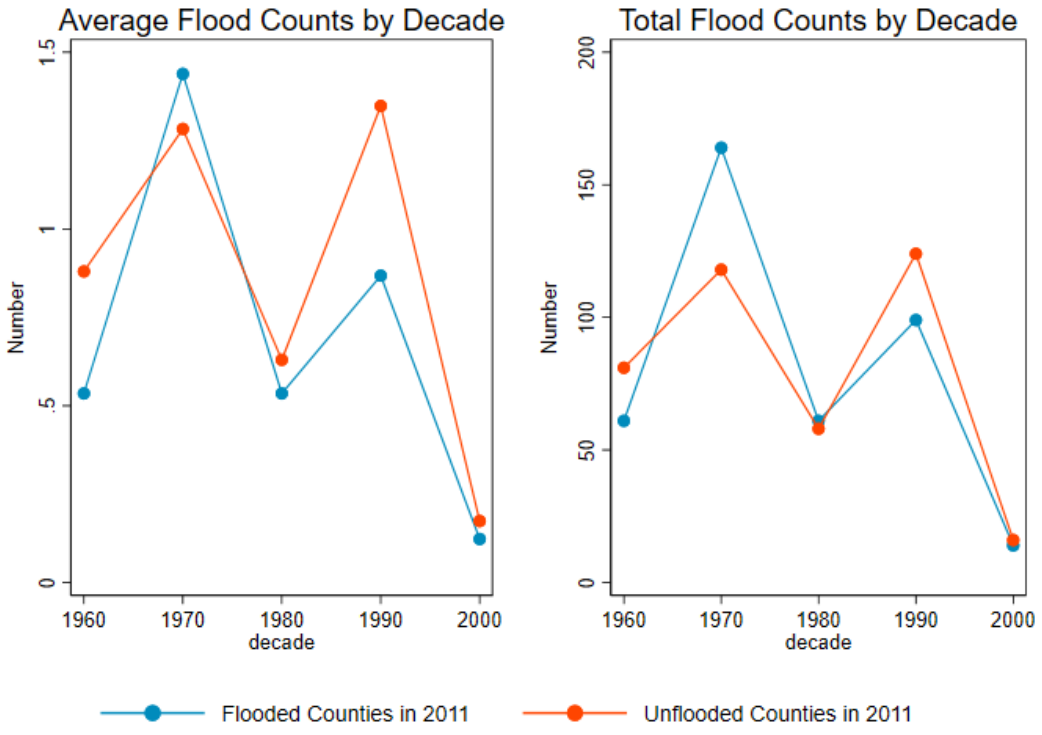
Notes: This figure plots Epanechnikov kernel densities of the change in levee elevations between the two pre-flood periods for levees that were flooded and unflooded in 2011. Only levees that have both pre-flood measurements are used (N=837, with N=522 for flooded levees and N=315 for unflooded levees), which consist about 67% of the entire data sample. See Table A.1 for more details on the list of years when levee elevations were measured for the two pre-flood periods. For the change in levee elevations between the two pre-flood periods, the mean (S.D.) of the unflooded distribution is 0.114 (0.027), and the mean (S.D.) of the flooded distribution is 0.088 (0.018). A two-sample t-test with equal variances shows that the difference in mean change in levee elevation between flooded and unflooded levees is statistically insignificant.

Figure A.2: Example of Topographic Maps and Levee Elevation Measurement



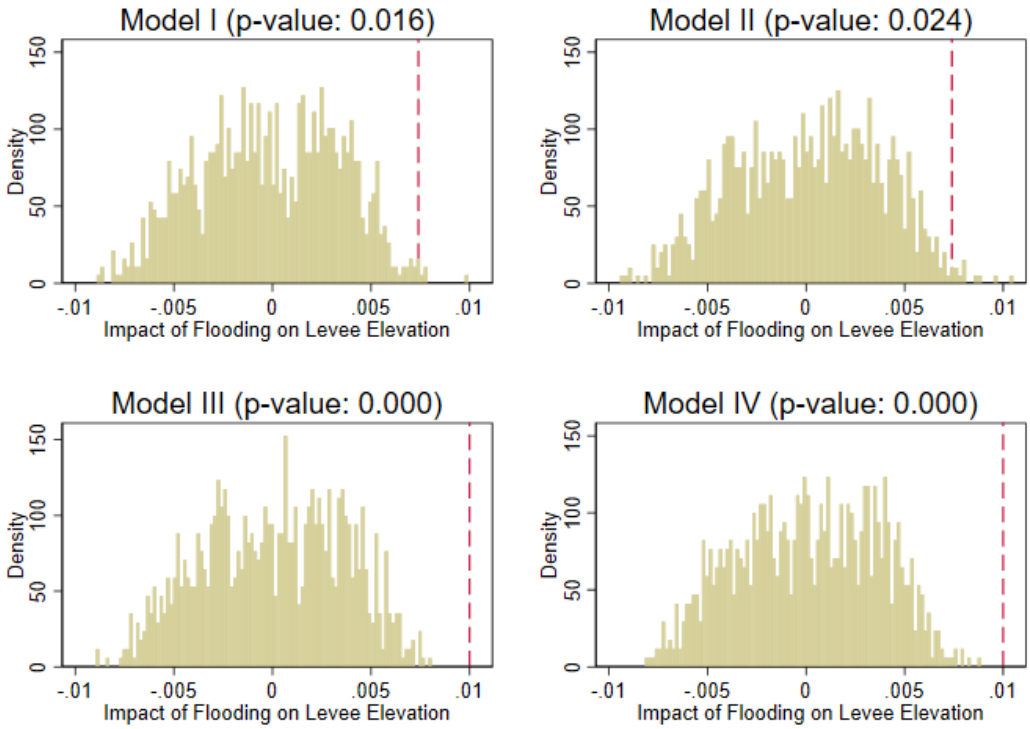
Notes: This topographic map shows the levees with elevations, using the Granite City Quadrangle Illinois-Missouri 7.5-minute Series at the scale of 1:24,000 (the finest scale that is available) in 1998 as an example. For the particular levee highlighted in the map, it shows that the elevation was 415 ft in 1998.

Figure A.3: Flood Counts during 1960–2000s



Notes: This figure plots the average and total number of flood counts at county level in each decade during the period 1960–2010s for counties that were flooded or unflooded during the Great Mississippi Flood of 2011. Similar pre-trending in terms of flood counts between flooded and unflooded counties in 2011 can be observed. A two-sample t-test with equal variances shows that the difference in mean flood counts between flooded and unflooded counties is statistically insignificant in all decades except for the decade of 1990.

Figure A.4: Randomization Inference Test



Note: This figure displays the results of a randomization inference test for the four model specifications used to estimate the impact of flooding on levee elevation in Table 1.3. For each model specification, I randomly re-assign flooding to exactly 114 counties 1,000 times, re-estimate each model specification, and save the coefficient of interest $\widehat{\beta_3}$. These permutated coefficients are in yellow histograms. The dashed vertical red lines are the actual estimated coefficients in Table 1.3. The p-value is the fraction of permutated coefficients that are larger than the actual coefficient. I can reject the null of no impact of flooding on levee elevation with a $p<0.05$ for all the four model specifications.

Appendix B: Appendix to Chapter 2

Table B.1: Probability of Responding to the Survey: Probit Model

Variable	Dependent variable: Respond (Yes=1)	
	Coefficient	S.E.
Land use land cover		
% in hay and pasture	-0.0001	(0.0009)
% in fruits and vegetables	0.0018	(0.0041)
% in forest	0.0007	(0.0010)
% in wetland	0.0059***	(0.0017)
% in other land uses	0.0006	(0.0007)
Land capability class		
% in class 3	-0.0006	(0.0014)
% in class 4	-0.0016	(0.0019)
% in class 5-8	-0.0016	(0.0017)
Drainage class		
% in poorly drained	-0.0005	(0.0012)
% in excessively drained	-0.0005	(0.0025)
Farmland class		
% in prime farmland	-0.0020	(0.0017)
% in other farmland	-0.0031**	(0.0014)
Hydrologic class		
% in high infiltration rate	-0.0004	(0.0009)
% in moderate infiltration rate	-0.0003	(0.0007)
Parcel size (acre)	-0.0002	(0.0002)
Soil rental rate ($/acre)	0.0007	(0.0007)
Average slope (% grade)	0.0034	(0.0069)
Average elevation (ft)	-0.0001	(0.0004)
Crop productivity index	-0.3335	(0.2179)
Distance to the Bay (mile)	-0.0006	(0.0010)
Crop reporting district		
Lower shore	-0.3678***	(0.1141)
Upper shore	-0.1315	(0.0862)
South	-0.3057***	(0.0957)
West	-0.3972**	(0.1569)
Constant	-0.4629**	(0.2167)
Number of observations	8,923	

Note: This table presents the probit regression results on landowners' probability of responding to the survey, where the dependent variable equals one if a landowner responded to our survey and explanatory variables include a list of parcel-specific characteristics. For land use land cover, "% in crops" is the baseline. For land capability class, "% in class 1-2" is the baseline. For drainage class, "% in well drained" is the baseline. For farmland class, "% in not prime farmland" is the baseline. For hydrologic class, "% in low infiltration rate" is the baseline. For crop reporting district, "central" is the baseline. Standard errors are in the parenthesis. *** p<0.01; ** p<0.05

Table B.2: Probability of Facing Stated Preference Questions: Probit Model

Variable	Dependent variable: Faced stated preference questions (yes=1)	
	Coefficient	S.E.
Corn (yes=1)	0.185	(0.114)
Soybeans (yes=1)	0.150	(0.114)
Small grains (yes=1)	-0.087	(0.124)
Vegetables (yes=1)	0.103	(0.145)
Hay and pasture (yes=1)	0.105	(0.114)
Beef cattle and calves (yes=1)	-0.103	(0.134)
Milk cows (yes=1)	0.041	(0.250)
Hogs and pigs (yes=1)	0.203	(0.151)
Sheep, lambs, and goats (yes=1)	-0.293	(0.307)
Horses and mules (yes=1)	-0.127	(0.168)
Poultry (yes=1)	0.095	(0.126)
Gross farm sales revenue		
$2,500-$4,999	0.003	(0.167)
$5,000-$9,999	0.347**	(0.159)
$10,000-$24,999	0.112	(0.168)
$25,000-$99,999	0.412**	(0.187)
$100,000-$499,999	-0.478*	(0.286)
$500,000-$999,999	-0.011	(0.409)
$1,000,000 or more	-0.300	(0.596)
I prefer not to say	0.281*	(0.169)
Rent_out	-0.079	(0.111)
Farm_income	-0.165	(0.239)
Senior	-0.006	(0.098)
College	-0.187*	(0.108)
Risk_averse	0.027	(0.109)
Enrollee	-3.577***	(0.252)
Self_funder	-3.422***	(0.230)
SRR	0.001	(0.001)
Attitude_property	0.083	(0.098)
Attitude_tax	0.022	(0.134)
Crop insurance (yes=1)	0.296*	(0.153)
Livestock insurance (yes=1)	-0.093	(0.322)
Constant	2.234***	(0.276)
Observations	1,335	

Note: This table presents the probit regression results on landowners' probability of facing DCE questions, where the dependent variable equals one if the landowner faced DCE questions and explanatory variables include a list of parcel and landowner characteristics. For gross farm sales revenue, "less than $2,500" is the baseline. Missing observations are not included in the regression. *** p<0.01; ** p<0.05; * p<0.1

Table B.3: Summary Statistics Comparison between Landowners Who Faced Discrete Choice Experiment (DCE) Questions and Those Who did not

Variable	Landowners who faced DCE questions	Landowners who did not face DCE questions
Corn (yes=1)	0.400	0.361
Soybeans (yes=1)	0.397	0.366
Small grains (yes=1)	0.230	0.234
Vegetables (yes=1)	0.102	0.129
Hay and pasture (yes=1)	0.509	0.563
Beef cattle and calves (yes=1)	0.172	0.181
Milk cows (yes=1)	0.040	0.032
Hogs and pigs (yes=1)	0.147	0.136
Sheep, lambs, and goats (yes=1)	0.017	0.033
Horses and mules (yes=1)	0.087	0.116
Poultry (yes=1)	0.177	0.218
Gross farm sales revenue		
Less than $2,500	0.329	0.387
$2,500-$4,999	0.112	0.117
$5,000-$9,999	0.146	0.109
$10,000-$24,999	0.102	0.113
$25,000-$99,999	0.096	0.084
$100,000-$499,999	0.055	0.051
$500,000-$999,999	0.007	0.015
$1,000,000 or more	0.007	0.010
I prefer not to say	0.145	0.114
Rent_out	0.504	0.488
Farm_income	0.154	0.139
Senior	0.570	0.606
College	0.606	0.700
Risk_averse	0.286	0.251
Enrollee	0.058	0.194
Self_funder	0.264	0.792
SRR	0.047	0.050
Attitude_property	0.605	0.536
Attitude_tax	0.190	0.149
Crop insurance (yes=1)	0.142	0.121
Livestock insurance (yes=1)	0.029	0.021

Note: This table presents the summary statistics comparison between landowners who faced DCE questions and those who did not in our survey.

Table B.4: Summary Statistics Comparison between Landowners Who Faced Discrete Choice Experiment (DCE) Questions and All Maryland Farmers (2017 Census of Agriculture)

Variable	Landowners who faced DCE questions	2017 Census of Agriculture
Corn (yes=1) [a]	0.400	0.238
Soybeans (yes=1)	0.397	0.202
Small grains (yes=1) [b]	0.230	0.141
Vegetables (yes=1)	0.102	0.077
Hay and pasture (yes=1)	0.509	0.372
Beef cattle and calves (yes=1)	0.172	0.200
Milk cows (yes=1)	0.040	0.041
Hogs and pigs (yes=1)	0.147	0.086
Sheep, lambs, and goats (yes=1)	0.017	0.143
Horses and mules (yes=1)	0.087	0.318
Poultry (yes=1)	0.177	0.228
Gross farm sales revenue [c]		
Less than $2,500	0.385	0.395
$2,500 to $4,999	0.132	0.098
$5,000 to $9,999	0.170	0.089
$10,000 to $24,999	0.120	0.111
$25,000 to $99,999	0.112	0.066
$100,000 to $499,999	0.064	0.148
$500,000 or more	0.018	0.094
Average age (years)	65	57
Crop insurance (yes=1)	0.142	0.128

Note: This table presents the summary statistics comparison between landowners who faced DCE questions and all Maryland farmers (2017 Census of Agriculture).

[a]: Sum of corn for grain and silage or greenchop in Census of Agriculture. Overlap not accounted for.

[b]: Sum of wheat, oats, barley, and sorghum in Census of Agriculture. Overlap not accounted for.

[c]: The "I prefer not to say" category in our survey is dropped in the calculation.

Appendix C: Appendix to Chapter 3

C.1 Description of Bayesian Estimation Routine

We can appeal to Bayes' Rule to define the relationship between the conditional distribution $\eta(\lambda|y,z,v)$ and the distribution over the population $g(\lambda|v)$. First consider the probability for the mixed logit:

$$P(y_i|z_i,v) = \int P(y_i|z_i,\lambda)g(\lambda|v)d\lambda \tag{C.1}$$

which gives the probability of household i's set of responses, given the data and the parameters of the population level parameter distribution. Using Bayes Rule, we express $\eta(\lambda|y,z,v)$ as $\frac{P(y_i|z_i,\lambda)g(\lambda|v)}{P(y_i|z_i,v)}$ and since the denominator is constant with respect to λ, $\eta(\lambda|y,z,v)$ is proportional to the numerator which provides a useful interpretation of $\eta(\cdot)$. The density of λ in the subpopulation that chose y_i when faced with z_i is proportional to the density of λ in the entire population, given by $g(\cdot)$., multiplied by the probability that someone would choose y_i given the data and the set of parameters λ.

Estimating the mixed logit model via hierarchical Bayes' requires iteratively drawing from the distributions for the mean and variance of $\boldsymbol{\lambda}$ and the household-level parameters while always conditioning on the most recent draw of the other parameters. We refer the reader to Train (2003, pages 302-308) for a complete description of the algorithm but for our purposes, consider a simulation that begins with starting values for the mean vector of $\boldsymbol{\lambda}$, a covariance matrix W, and a vector of household-level parameters for each respondent λ_i. The first step of the algorithm draws a realization of the mean vector $\bar{\lambda}$ conditional on λ_i and W which is distributed $N(\sum_i \frac{\lambda_i}{N}, W)$. The second step draws the covariance matrix W from an inverse Wishart distribution with $M+N$ degrees of freedom and scale matrix $\frac{MI+N\bar{S}}{M+N}$, where I is a M-

dimensional identity matrix and $\bar{S} = \frac{\sum_i (\lambda_i - \bar{\lambda})(\lambda_i - \bar{\lambda})'}{N}$. The third and final step of the algorithm draws household-level parameter vectors from a density proportional to $\prod_t \frac{e^{\lambda_i z_i y_{it} t}}{\sum_i e^{\lambda_i z_{it}}} \phi(\lambda_i | \bar{\lambda}, W)$ which requires a MH algorithm. After a burn-in period, the draws will converge to the joint posterior distribution of the model parameters.

C.2 Parametric Estimation of Distance Decay

As a preliminary step to examine whether preferences for environmental improvements vary with distance to the resource, researchers often include a measure of the distance of respondent i to the resource (e.g., Pate and Loomis, 1997; Hanley et al., 2003; Bateman et al., 2006). We denote this distance measure as $f(d_i)$, and examine various functional forms for such a global distance gradient using the SP data from Moore et al. (2018). The model in equation (3.1) of section 3.2.1 in the main text is thus augmented as follows:

$$P_i\left(j \mid \boldsymbol{x}_q, cost_q, SQ_q, d_q\right) = \frac{exp\left\{ln\left(\boldsymbol{x}_j\right)\boldsymbol{\beta} + \left[ln\left(\boldsymbol{x}_j\right) \times f(d_i)\right]\theta + \gamma cost_j + \varphi SQ_j\right\}}{\sum_q exp\left\{ln\left(\boldsymbol{x}_q\right)\boldsymbol{\beta} + \left[ln\left(\boldsymbol{x}_q\right) \times f(d_i)\right]\theta + \gamma cost_q + \varphi SQ_q\right\}} \tag{C.2}$$

In these preliminary regression models, equation (C.2) is estimated as a conditional logit model. The added parameter vector θ will reflect any spatial heterogeneity in respondent's preferences for improvements in the environmental commodity over space. If estimates of θ are statistically significant, then it would suggest that preferences, and hence subsequent MWTP calculations, vary with distance. If we are unable to find a specification of $f(d_i)$ that yields statistically significant estimates of θ, then it is reasonable to conclude that MWTP does not change in a smooth, parametric way over space. Such a finding motivates our more flexible semi-parametric examination of spatial heterogeneity in MWTP.

The results of a series of conditional logit models following equation (C.2) are estimated using different functional form assumptions for $f(d_i)$, including linear distance, the natural log of distance, inverse distance, a quadratic specification, and various stepwise functions (e.g., within versus outside a 50-kilometer buffer, within versus outside of the watershed, and based on the three geographic strata in the experimental design of the survey application). The results are presented in Table C.1 below. The individual coefficient results are difficult to interpret given the numerous interaction terms, and in any case, are not of primary interest here.

The results of interest pertain to the statistical significance of the coefficients, or lack thereof, in the lower panel of Table C.1. These estimates correspond to θ in equation (C.2). As can be seen, the estimates are often statistically insignificant, both individually and jointly. This demonstrates that, at least under the lens of this conventional parametric distance decay paradigm, there is little evidence of spatial heterogeneity in MWTP. Researchers and practitioners may often stop here and assume spatial homogeneity in the household benefits going forward. Such a conclusion, however, may be premature. Statistically significant spatial heterogeneity can still be identified using the two-step semi-parametric approaches we discuss in the main paper.

Table C.1: Conditional Logit Models with Parametric Distance Gradient

VARIABLES	Linear distance (1)	ln(distance) (2)	Inverse distance (3)	Quadratic distance (4)		Stepwise: 50km (5)	Stepwise: In watershed (6)	Stepwise: Geographic Strata (7)	
ln(clarity)	2.336e-01	1.973e-01	2.589e-01	1.746e-01		3.183e-01	3.367e-01	7.411e-03	
	(0.353)	(0.934)	(0.259)	(0.466)		(0.287)	(0.336)	(0.430)	
ln(bass)	-5.943e-01	-1.388e+00	-2.395e-01	-8.126e-01*		-2.248e-02	-1.869e-02	-6.930e-01	
	(0.371)	(0.982)	(0.276)	(0.486)		(0.310)	(0.363)	(0.447)	
ln(crab)	-1.859e-01	-5.876e-01	3.363e-01	-2.027e-01		4.288e-01	7.597e-01	-5.320e-01	
	(0.519)	(1.386)	(0.387)	(0.691)		(0.428)	(0.480)	(0.636)	
ln(oyster)	-1.131e-01	-7.380e-02	-5.586e-02	-7.431e-02		-6.705e-02	-1.439e-03	-1.870e-02	
	(0.127)	(0.335)	(0.095)	(0.166)		(0.107)	(0.123)	(0.150)	
ln(lake)	1.160e+00**	4.660e-01	1.573e+00***	1.359e+00**		1.883e+00***	1.788e+00***	8.586e-01	
	(0.492)	(1.423)	(0.358)	(0.662)		(0.386)	(0.451)	(0.613)	
SQC	-8.971e-01***	-1.293e+00***	-7.081e-01***	-1.051e+00***		-5.598e-01***	-5.090e-01***	-1.063e+00***	
	(0.178)	(0.432)	(0.137)	(0.227)		(0.151)	(0.172)	(0.215)	
cost	-5.008e-03***	-5.009e-03***	-5.045e-03***	-5.005e-03***		-5.035e-03***	-5.004e-03***	-5.064e-03***	
	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	
				distance	**distance^2**			**watershed states**	**other east coast states**
ln(clarity) × f(dist)	1.591e-05	1.022e-02	-1.376e-02	4.576e-04	-4.662e-07	-3.273e-01	-2.313e-01	6.531e-01	1.747e-02
	(0.001)	(0.178)	(0.474)	(0.003)	(0.000)	(0.668)	(0.529)	(0.628)	(0.639)
ln(bass)× f(dist)	1.257e-03	2.352e-01	3.654e-01	3.060e-03	-1.689e-06	-1.010e+00	-5.013e-01	3.860e-01	1.242e+00*
	(0.001)	(0.191)	(0.577)	(0.003)	(0.000)	(0.662)	(0.553)	(0.626)	(0.727)
ln(crab) × f(dist)	1.862e-03	1.920e-01	7.844e-01	1.956e-03	3.062e-08	-2.114e-01	-9.659e-01	6.992e-01	2.407e+00**
	(0.001)	(0.269)	(0.589)	(0.004)	(0.000)	(0.950)	(0.794)	(0.875)	(1.005)
ln(oyster) × f(dist)	2.262e-04	4.815e-03	2.075e-01	-1.094e-04	3.101e-07	9.369e-02	-1.228e-01	-2.197e-01	2.024e-01
	(0.000)	(0.065)	(0.161)	(0.001)	(0.000)	(0.224)	(0.190)	(0.223)	(0.230)
ln(lake) × f(dist)	1.551e-03	2.283e-01	3.689e-01	-1.544e-04	1.592e-06	-1.455e+00	-5.109e-01	6.058e-01	2.075e+00**
	(0.001)	(0.273)	(0.147)	(0.004)	(0.000)	(0.977)	(0.739)	(0.841)	(0.938)
Joint significance	$\chi2(4) = 4.64$	$\chi2(4) = 1.94$	$\chi2(4) = 6.30$	$\chi2(8) = 5.51$		$\chi2(4) = 2.56$	$\chi2(4) = 2.78$	$\chi2(8) = 12.40$	
	p = 0.3267	p = 0.7464	p = 0.1779	p = 0.7014		p = 0.6334	p = 0.5958	p = 0.1341	
SQC × f(dist)	6.697e-04	1.189e-01	6.904e-02	1.896e-03	-1.130e-06	-7.387e-01**	-4.840e-01*	5.166e-01*	6.248e-01*
	(0.000)	(0.084)	(0.100)	(0.001)	(0.000)	(0.304)	(0.260)	(0.296)	(0.330)
Observations	4,719	4,719	4,719	4,719		4,719	4,719	4,719	
LL	-1567.4058	-1569.1278	-1568.1169	-1564.6207		-1566.0111	-1568.7936	-1555.4703	

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Bibliography

Allen, D.W., Lueck, D., 1995. Risk preferences and the economics of contracts. *American Economic Review* 85(2): 447–451.

Alix-Garcia, J., Wolff, H., 2014. Payment for ecosystem services from forests. *Annual Review of Resource Economics* 6(1): 361–380.

Anderson, M.L., Auffhammer, M., 2014. Pounds that kill: The external costs of vehicle weight. *Review of Economic Studies* 81: 535–571.

Anselin, L.,1995. Local indicators of spatial association—LISA. *Geographical Analysis* 27(2): 93–115.

Anselin, L., Gallo, J.L., 2006. Interpolation of air quality measures in hedonic house price models: Spatial aspects. *Spatial Economic Analysis* 1(1): 31–52.

Arnold, J.L., 1988. The evolution of the 1936 Flood Control Act. Office of History, U.S. Army Corps of Engineers.

Badura, T., Ferrini, S., Burton, M., Binner, A., Bateman, I.J., 2020. Using individualised choice maps to capture the spatial dimensions of value within choice experiments. *Environmental and Resource Economics* 75: 297–322.

Banzhaf, H.S., Chupp, B.A., 2012. Fiscal federalism and interjurisdictional externalities: New results and an application to US Air pollution. *Journal of Public Economics* 96: 449–464.

Barry, J.M., 1998. Rising tide: The Great Mississippi Flood of 1927 and how it changed America. New York: Simon & Schuster.

Bateman, I.J., Day, B.H., Georgiou, S., Lake, I., 2006. The aggregation of environmental benefit values: Welfare measures, distance decay and total WTP. *Ecological Economics* 60(2): 450–460.

Belasen, A.R., Polachek, S.W., 2009. How disasters affect local labor markets? The effects of hurricanes in Florida. *Journal of Human Resources* 44(1): 251–276.

Beltran, A., Maddison, D., Elliott, R., 2018. Is flood risk capitalised into property values? *Ecological Economics* 146: 668–685.

Beltran, A., Maddison, D., Elliott, R., 2019. The impact of flooding on property prices: A repeat-sales approach. *Journal of Environmental Economics and Management* 95: 62–86.

Bin, O., Landry, C.E., 2013. Changes in implicit flood risk premiums: Empirical evidence from the housing market. *Journal of Environmental Economics and Management* 65(3): 361–376.

Blundell, R., Ham, J., Meghir, C., 1987. Unemployment and female labour supply. *The Economic Journal* 97: 44–64.

Bond, C.A., Cullen, K.G., Larson, D.M., 2009. Joint estimation of discount rates and willingness to pay for public goods. *Ecological Economics* 68: 2751–2759.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2010. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods* 1(2): 97–111.

Boustan, L.P., Kahn, M.E., Rhode, P.W., Yanguas, M.L., 2020. The effect of natural disasters on economic activity in US counties: A century of data. *Journal of Urban Economics*, forthcoming.

Brick, K., Visser, M., 2015. Risk preferences, technology adoption and insurance uptake: A framed experiment. *Journal of Economic Behavior & Organization* 118: 383–396.

Brown, J.P., Maniloff, P., Manning, D.T., 2020. Spatially variable taxation and resource extraction: The impact of state oil taxes on drilling in the US. *Journal of Environmental Economics and Management* 103: 102354.

Budish, E., Cramton, P., Shim, J., 2015. The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics* 130(4): 1547–1621.

Budziński, W., Campbell, D., Czajkowski, M., Demšar, U., Hanley, N., 2017. Using geographically weighted choice models to account for the spatial heterogeneity of preferences. *Journal of Agricultural Economics* 69(3): 606–626.

Budziński, W., Czajkowski, M., 2021. Accounting for spatial heterogeneity of preferences in discrete choice models. *Central European Journal of Economic Modelling and Econometrics* 13: 1–24.

Burlig, F., Stevens, A.W., 2019. Reap what your friends sow: Social networks and technology adoption. Working Paper.

Burton, M., Rigby, D., 2009. Hurdle and latent class approaches to serial non-participation in choice models. *Environmental and Resource Economics* 42: 211–226.

Cai, H., Chen, Y., Gong, Q., 2015. Polluting thy neighbor: Unintended consequences of China's pollution reduction mandates. *Journal of Environmental Economics and Management* 76: 86–104.

Canales, E., Bergtold, J.S., Williams, J., Peterson, J., 2015. Estimating farmers' risk attitudes and risk premiums for the adoption of conservation practices under different contractual arrangements: A stated choice experiment. Selected Paper prepared for presentation at the 2015 AAEA and WAEA Annual Meeting.

Camillo, C.A., 2012. Divine providence: The 2011 flood in the Mississippi River and Tributaries Project. U.S. Army Corps of Engineers, 142.

Campbell, D., Hutchison, W.G., Scarpa, R., 2009. Using choice experiments to explore the spatial distribution of willingness to pay for rural landscape improvements. *Environment and Planning A* 41: 97–111.

Cattaneo, M.D., Titiunik, R., Vazquez-Bare, and G., 2017. Comparing inference approaches for RD designs: A reexamination of the effect of head start on child mortality. *Journal of Policy Analysis and Management* 36(3): 643–681.

Chen, Z., Swallow, S.K., Yue, I.T., 2020. Non-participation and heterogeneity in stated: A double hurdle latent class approach for climate change adaptation plans and ecosystem services. *Environmental and Resource Economics* 77: 35–67.

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

Cohen, J., Dupas, P., 2010. Free distribution or cost-sharing? Evidence from a randomized Malaria prevention experiment. *Quarterly Journal of Economics* 125: 1–45.

Conley, T., 1999.GMMestimation with cross sectional dependence. *Journal of Econometrics* 92(1): 1–45.

Costello, C., QuÅLerou, N., Tomini, A., 2017. Private eradication of mobile public bads. *European Economic Review* 94: 23–44.

Cowdrey, A.E., 1977. Land's end: A history of the New Orleans District, U.S. Army Corps of Engineers, and its lifelong battle with the lower Mississippi and other rivers wending their way to the sea. U.S. Army Corps of Engineers, New Orleans District.

Cunningham, S., Bennear, L.S., Smith, M.D., 2016. Spillovers in regional fisheries management: Do catch shares cause leakage? *Land Economic* 92(2): 344–362.

Czajkowski, M., Budziński W., Campbell, D., Giergiczny, M., Hanley, N., 2017. Spatial heterogeneity of willingness to pay for forest management. *Environmental and Resource Economics* 68: 705–727.

Danley, B., Sandorf, E.D., Campbell, D., 2021. Putting your best fish forward: Investigating distance decay and relative preferences for fish conservation. *Journal of Environmental Economics and Management* 108: 102475.

Demsetz, H., 1967. Toward a theory of property rights. *American Economic Review: Papers and Proceedings* 57: 347–359.

Dundas, S.J., Lewis, D.J., 2020. Estimating option values and spillover damages for coastal protection: Evidence from Oregon's Planning Goal 18. *Journal of the Association of Environmental and Resource Economists* 7(3): 519–554.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. *Quarterly Journal of Economics* 133(4): 1645–1692.

Falk, A., Becker, A., Dohmen, T., Huffman, D., Sunde, U., 2016. The preference survey module: A validated instrument for measuring risk, time, and social preferences. IZA DP No. 9674.

Federal Emergency Management Agency (FEMA), 2016. Factsheet: What is a levee? https://www.fema.gov/media-library/assets/documents/22951.

Fell, H, Kaffine, D.T., 2014. Can decentralized planning really achieve first-best in the presence of environmental spillovers? *Journal of Environmental Economics and Management* 68: 46–53.

Fischer, S., Wollni, M., 2018. The role of farmers' trust, risk and time preferences for contract choices: Experimental evidence from the Ghanaian pineapple sector. *Food Policy* 81: 67–81.

Flor, A., Pinter, N., Remo, J.W.F., 2011. The ups and downs of levees: GPS–based change detection, Middle Mississippi River, USA. *Geology* 39(1): 55–58.

Foelske, L., van Riper, G.J., 2020. Assessing spatial preference heterogeneity in a mixed-use landscape. *Applied Geography* 125: 102355.

Galloway, G.E., 1995. Learning from the Mississippi Flood of 1993: Impacts, management issues, and areas for research. U.S.- Italy Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods Perugia (Italy).

Getis, A., 2007. Reflections on spatial autocorrelation. *Regional Science and Urban Economics* 37(4): 491–496.

Getis, A., 2010. Spatial autocorrelation. In: Fischer, M.M., Getis, A. (Eds.). Handbook of Applied Spatial Analysis. Springer, Heidelberg.

Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24(3): 189–206.

Gibson, M., Mullins, J., 2020. Climate risk and beliefs in New York floodplains. *Journal of the Association of Environmental and Resource Economists* 7(6): 1069–1111.

Glenk, K., Johnston, R.J., Meyerhoff, J., Sagebiel, J., 2020. Spatial dimensions of stated preference valuation in environmental and resource economics: Methods, trends and challenges. *Environmental and Resource Economics* 75: 215–242.

Grammatikopoulou, I., Artell, J., Hjerppe, T., Pouta, E., 2020. A mire of discount rates: Delaying conservation payment schedules in a choice experiment. *Environmental and Resource Economics* 77: 615–639.

Greene, W.H., 2003. Econometric Analysis. New Jersey: Prentice-Hall.

Ham, J.C., 1982. Estimation of a labour supply model with censoring due to unemployment and underemployment. *Review of Economic Studies* 49(3): 335–354.

Hanemann, W.M., 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics* 66(3): 332–341.

Hanley, N., Schalpfer, F., Spurgeon J., 2003. Aggregating the benefits of environmental improvements: Distance-decay functions for use and non-use values. *Journal of Environmental Management* 68: 297–304.

Harrison, R.W., 1951. Levee districts and levee building in Mississippi: A study of State and local efforts to control Mississippi River floods. Stoneville Miss. Delta Council.

Heine, R.A., Pinter, N., 2012. Levee effects upon flood levels: An empirical assessment. *Hydrological Processes* 26: 3225–3240.

Hennighausen, H., Suter, J.F., 2020. Flood risk perception in the housing market and the impact of a major flood event. *Land Economics* 96(3): 366–383.

Hersher, 2018. Levees make Mississippi River floods worse, but we keep building them. NPR, 21 May 2018.

He., S., 2017. Randomization inference with Stata: A guide and software. *The Stata Journal* 17(3): 630–651.

Hoffer, A., Humphreys, B.R., Lacombe, D.J., Ruseski, J.E., 2015. Trends in NCAA Athletic Spending: Arms Race or Rising Tide? *Journal of Sports Economics* 16(6): 576–596.

Hole, A.R. Kolstad, J.R., 2012. Mixed logit estimation of willingness to pay distributions: A comparison of models in preference and WTP space using data from a health-related choice experiment. *Empirical Economics* 42(2): 445–469.

Holmes T.P., Adamowicz, W.L., 2003. Attribute-based methods. In: Champ P.A., Boyle K.J., Brown T.C. (eds) A Primer on Nonmarket Valuation. The Economics of Non-Market Goods and Resources, vol 3. Springer, Dordrecht.

Hornbeck, R., Naidu, S., 2014. When the levee breaks: Black migration and economic development in the American South. *American Economic Review* 104(3): 963–990.

Howard, G., Whitehead, J.C., Hochard, J., 2021. Estimating discount rates using referendum-style choice experiments: An analysis of multiple methodologies. *Journal of Environmental Economics and Management* 105: 102399.

Humphreys, B.G., 1914. Floods and levees of the Mississippi River. Washington, D.C., The Mississippi River Association.

Interagency Floodplain Management Review Committee (IFMRC), 1994. Sharing the challenge: Floodplain management into the 21st century. Report to the Administration Floodplain Management Task Force, U.S. Government Printing Office, Washington D.C., 191 pp.

Iyer, P., Bozzola, M., Hirsch, S., Meraner, M., and Finger, R., 2020. Measuring farmer risk preferences in Europe: A systematic review. *Journal of Agricultural Economics* 71(1): 3–26.

Jacobson, R.B., Linder, G., Binter, C., 2015. The role of floodplain restoration in mitigating flood river, Lower Missouri River, USA. In: Hudson, P.F., Middelkoop, H. (Eds.), Geomorphic Approaches to Integrated Floodplain Management of Lowland Fluvial System in North America and Europe. Springer, New York.

Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R., et al., 2017. Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists* 4(2): 319–405.

Johnston, R.J., Jarvis, D., Wallmo, K., Lew, D.K., 2015. Multiscale spatial pattern in nonuse willingness to pay: Applications to threatened and endangered marine Species. *Land Economics* 91(4): 739–761.

Johnston, R.J., Ndebele, T., Newburn, D.A., 2022. Modeling transaction costs in household adoption of landscape conservation practices. American Journal of Agricultural Economics. Forthcoming.

Johnston, R.J., Ramachandran, M., 2014. Modeling spatial patchiness and hot spots in stated preference willingness to pay. *Environmental and Resource Economics* 59: 363–387.

Jørgensen, S.L., Olsen, S.B., Ladenburg, J., et al., 2013. Spatially induced disparities in users' and non-users' WTP for water quality improvements—testing the effect of multiple substitutes and distance decay. *Ecological Economics* 92: 58–66.

Kim, S.I., Haab, T.C., 2009. Temporal insensitivity of willingness to pay and implied discount rates. *Resource and Energy Economics* 31: 89–102.

Klein, C., Zellmer, S., 2007. Mississippi River stories: Lessons from a century of unnatural disasters. *SMU Law Review* 60: 1471–1538.

Kousky, C., 2010. Learning from extreme events: Risk perceptions after the flood. *Land Economics* 86(3): 395–422.

Kuwayama, Y., Brozovic, N., 2013. The regulation of a spatially heterogeneous externality: Tradable groundwater permits to protect streams. *Journal of Environmental Economics and Management* 66: 364–382.

Lakdawalla, D., Zanjani, G., 2005. Insurance, self-protection, and the economics of terrorism. *Journal of Public Economics* 89: 1891–1905.

Lawrence, J., Quade, D., Becker, J., 2014. Integrating the effects of flood experience on risk perception with responses to changing climate risk. *Natural Hazards* 74: 1773–1794.

Layton, D.F., Brown, G., 2000. Heterogeneous preferences regarding global climate change. *Review of Economics and Statistics* 82(4): 616–624.

Lew, D.K., 2018. Discounting future payments in stated preference choice experiments. *Resource and Energy Economics* 54: 150–164.

Lipscomb, M., Mobarak, A.M., 2017. Decentralization and pollution spillovers: Evidence from the re-drawing of county borders in Brazil. *Review of Economic Studies* 84(1): 464–502.

Lloyd-Smith, P., and Adamowicz, W., 2018. Can stated measures of willingness-to-accept be valid? Evidence from laboratory experiments. *Journal of Environmental Economics and Management* 91: 133–149.

Lloyd-Smith, P., Adamowicz, W., Entem, A., Fenichel, E.P., Rad, M.R., 2021. The decade after tomorrow: Estimation of discount rates from realistic temporal decisions over long time horizons. *Journal of Economic Behavior and Organization* 183: 158–174.

Maddala, G.S., 1983. Limited-dependent and qualitative variables in economics. New York: Cambridge University Press.

Maher, K., 2019. Mississippi River neighbors battle over levee heights. The Wall Street Journal, 4 February 2019.

Manley, J. Mathias, J., 2017. CREP: cattle receiving enhanced pastures? Investigating landowner response to federal incentives. *Land Economics* 93(1): 59–73.

McCann, L., Claassen, R., 2016. Farmer transaction costs of participating in federal conservation programs: Magnitudes and determinants. *Land Economics* 92(2): 256–272.

Meyerhoff, J., 2013. Do turbines in the vicinity of respondents' residences influence choices among programmes for future wind power generation? *Journal of Choice Modelling* 7: 58–71.

McCoy, S.J., Zhao, X., 2018. A city under water: A geospatial analysis of storm damage, changing risk perceptions, and investment in residential housing. *Journal of Association of Environmental and Resource Economists* 5(2): 301–330.

Michler, J.D., Wu, S.Y., 2020. Relational contracts in agriculture: Theory and evidence. *Annual Review of Resource Economics* 12: 18.1–18.17.

Missirian, A., 2020. Yes, in your backyard: Forced technological adoption and spatial externalities. Job Market Paper.

Moore, C., Guignet, D., Maguire, K., Simon, N., Dockins, C., 2018. Valuing ecological improvements in the Chesapeake Bay and the importance of ancillary benefits. *Journal of Benefit-Cost Analysis* 9(1): 1–26.

Myers, K., Parsons, G., Train, K., 2017. Inadequate response to frequency of payments in contingent valuation of environmental goods. In: Mcfadden, D.,Train, K. (Eds.),

Chapter 3 in Contingent Valuation of Environmental Goods: A Comprehensive Critique. Edward Elgar, Northampton, MA, pp. 43–57.

National Research Council, 2000. Risk analysis and uncertainty in flood damage reduction studies. Washington, DC: The National Academies Press.

Nelson, T.A., Boots, B., 2008. Detecting spatially explicit hot spots in landscape-scale ecology. *Ecography* 31(5): 556–566.

Olsen, S.B., Jensen, C.U., Panduro, T.E., 2020. Modelling strategies for discontinuous distance decay in willingness to pay for ecosystem services. *Environmental and Resource Economics* 75: 351–386.

Ord, J.K., Getis, A., 1995. Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis* 27(4): 286–306.

Palm-Forster, L.H., Swinton, S.M., Lupi, F., Shupp, R.S., 2016. Too burdensome to bid: Transaction costs and pay-for-performance conservation. *American Journal of Agricultural Economics* 98(5): 1314–1333.

Pate, J., Loomis J.B., 1997. The effect of distance on willingness to pay values: A case study of wetlands and salmon in California. *Ecological Economics* 20: 199–207.

Peterson, J.M., Smith, C.M., Leatherman, J.C., Hendricks, N.P., Fox, J.A., 2014. Transaction costs in payment for environmental service contracts. *American Journal of Agricultural Economics* 97(1): 219–238.

Pinter, N., Huthoff, F., Dierauer, J., Remo, J.W.F., Damptz, A., 2016. Modeling residual flood risk behind levees, Upper Mississippi River, USA. *Environmental Science & Policy* 58: 131–140

Pitlick, J., 1997. A regional perspective of the hydrology of the 1993 Mississippi River Basin floods. *Annals of the Association of American Geographers* 87(1): 135–151.

Ramsey, S.M., Bergtold, J.S., Canales, E., Williams, J.R., 2019. Effects of farmers' yield-risk perceptions on conservation practice adoption in Kansas. *Journal of Agricultural and Resource Economics* 44(2): 380–403.

Remo, J.W.F., Carlson, M., Pinter, N., 2012. Hydraulic and flood-loss modeling of levee, floodplain, and river management strategies, Middle Mississippi River, USA. *Natural Hazards* 61: 551–575.

Remo, J.W.F., Ickes, B.S., Ryherd, J.K., Guida, R.J., Therrell, M.D., 2018. Assessing the impacts of dams and levees on the hydrologic record of the Middle and Lower Mississippi River, USA. *Geomorphology* 313: 88–100.

Revelt, D., Train, K., 2000. Customer-specific taste parameters and mixed logit: Households' choice of electricity supplier. Working paper No. E00-274, Department of Economics University of California Berkeley.

Ribaudo, M., Shortle, J., 2019. Reflections on 40 years of applied economics research on agriculture and water quality. *Agricultural and Resource Economics Review* 48(3): 519–530.

Rode, J., Weber, A., 2016. Does localized imitation drive technology adoption? A case study on rooftop photovoltaic systems in Germany. *Journal of Environmental Economics and Management* 78: 38–48.

Rolfe, J., Windle, J., 2012. Distance decay functions for iconic assets: Assessing national values to protect the health of the great barrier reef in Australia. *Environmental and Resource Economics* 53: 347–365.

Scarpa, R., Rose, J. M., 2008. Design efficiency for non-market valuation with choice modelling: How to measure it, what to report and why. *Australian Journal of Agricultural and Resource Economics* 52(3): 253–282.

Scarpa, R., Thiene, M., Train, K., 2008. Utility in willingness to pay space: A tool to address confounding random scale effects in destination choice to the Alps. *American Journal of Agricultural Economics* 90(4): 994–1010.

Schaafsma, M., Brouwer, R., Gilbert, A., van den Bergh, J., Wagtendonk, A., 2013. Estimation of distance-decay functions to account for substitution and spatial heterogeneity in stated preference research. *Land Economics* 89: 514–537.

Shafran, H., 2005. Transboundary spillovers and decentralization of environmental policies. *Journal of Environmental Economics and Management* 50: 82–101.

Sigman, A.P., 2005. Risk externalities and the problem of wildfire risk. *Journal of Urban Economics* 64: 488–495.

Stiglitz, J., 1974. Incentives and risk sharing in sharecropping. *Review of Economic Studies* 41(2): 219–255.

Strobl, E., 2011. The economic growth impact of hurricanes: Evidence from U.S. coastal counties. *Review of Economics and Statistics* 93(2): 575–589.

Struyk, R.J., 1971. Flood risk and agricultural land value: A test. *Water Resources Research* 7(4): 789–797.

Suter, J.F., Poe, G.L., Bills, N.L., 2008. Do landowners respond to land retirement incentives? Evidence from the Conservation Reserve Enhancement Program. *Land Economics* 84(1): 17–30.

Tanner, A., Arvai, J., 2018. Perceptions of risk and vulnerability following exposure to a major natural disaster: The Calgary Flood of 2013. *Risk Analysis* 38(3): 548–561.

Tarlock, A.D., 2012. United States flood control policy: The incomplete transition from the illusion of total protection to risk management. *Duke Environmental Law Policy Forum* 23: 151–183.

Tobin, G.A., 1995. The levee love affair: A stormy relationship. *Water Resources Bulletin* 31(3): 359–367.

Toledo-Gallegos, V.M., Long, J., Campbell, D., Börger, T., Hanley, N., 2021. Spatial clustering of willingness to pay for ecosystem services. *Journal of Agricultural Economics* 72: 673–697.

Train, K.E., 2003. Discrete choice analysis with simulation. Cambridge, Cambridge, UK.

U.S. Army Corps of Engineers (USACE), 1999. Digest of water resources policies and authorities. Washington, D.C.: U. S. Army Corps of Engineers.

U.S. Army Corps of Engineers (USACE), 2011. Great Flood of ’11. Our Mississippi. Rock Island, IL: U.S. Army Corps of Engineers.

U.S. Army Corps of Engineers (USACE), 2014. Mississippi River and tributaries project: Levee system evaluation report for the National Flood Insurance Program. Mississippi Valley Division: U. S. Army Corps of Engineers.

Vasquez-Lavín F., Carrasco, M., Barrientos, M., Gelcich, S., Olivia, R.D.P., 2021. Estimating discount rates for environmental goods: Are people’s responses inadequate to frequency of payments? *Journal of Environmental Economics and Management* 107: 102446.

Vasquez-Lavín, F., Oliva, R.D.P., Hernández, J.I., Gelcich, S., Carrasco, M. and Quiroga, M., 2019. Exploring dual discount rates for ecosystem services: Evidence from a marine protected area network. *Resource and Energy Economics* 55: 63–80.

Villanueva, A.J., Glenk, K., Rodriguez-Entrena, M., 2017. Protest responses and willingness to accept: Ecosystem services providers’ preferences towards incentive-based schemes. *Journal of Agricultural Economics* 68(3): 801–821.

Viscusi, W.K., Huber, J., Bell, J., 2008. Estimating discount rates for environmental quality from utility-based choice experiments. *Journal of Risk Uncertainty* 37(2-3): 199–220.

Von Haefen, R.H., Massey, D.M., Adamowicz, W.L., 2005. Serial nonparticipation in repeated discrete choice models. *American Journal of Agricultural Economics* 87(4): 1061–1076.

Vousdoukas, M.I., Mentaschi, L., Hinkel, J.,Ward, P.J., Mongelli, I., Ciscar, J., Feyen, L., 2020. Economic motivation for raising coastal flood defenses in Europe. *Nature Communications* 11: 2119.

Wang, H, Qiu, F., 2017. Investigation of the dynamics of agricultural land at the urban fringe: A comparison of two peri-urban areas in Canada. *The Canadian Geographer* 61(3): 457–470.

Wang, H., Swallow, B.M., 2016. Optimizing expenditures for agricultural land conservation: Spatially-explicit estimation of benefits, budgets, costs and targets. *Land Use Policy* 59: 272–283.

Weng, W., Morrison, M.D., Boyle, K.J., Boxall, P.C., Rose, J., 2021. Effects of the number of alternatives in public good discrete choice experiments. *Ecological Economics* 182: 106904.

West, G.H., Snell, H., Kovacs, K.F., Nayga, R.M., 2021. Flexible estimation of groundwater service values and time preferences. *Journal of the Association of Environmental and Resource Economists* 8(4): 825–861.

White, L., 2004. The “arms race” on American roads: The effect of sport utility vehicles and pickup trucks on traffic safety. *Journal of Law and Economics* 47(2): 333–355.

Wunder, S., Engel, S., Pagiola, S., 2008. Taking stock: A comparative analysis of payments for environmental services programs in developed and developing countries. *Ecological Economics* 65(4): 834–852.

Yang, M., Bhatta, R.A., Chou, S., H., C., 2017. The impact of prenatal exposure to power plant emissions on birth weight: Evidence from a Pennsylvania power plant located upwind of New Jersey. *Journal of Policy Analysis and Management* 36(3): 557–583.

Yen, B.C., 1995. Hydraulics and effectiveness of levees for flood control. U.S.-Italy Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods Perugia (Italy), November 1995.