2024

Essays on Impacts of Natural Disasters

Fang-Yu Yeh
University of Kentucky, fye224@uky.edu
Author ORCID Identifier:
https://orcid.org/0009-0009-9683-2605
Digital Object Identifier: https://doi.org/10.13023/etd.2024.312b

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Fang-Yu Yeh, Student

Dr. William Hoyt, Major Professor

Dr. Carlos Lamarche, Director of Graduate Studies

Essays on Impacts of Natural Disasters

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Arts and Sciences
at the University of Kentucky

By
Fang-Yu Yeh
Lexington, Kentucky

Directors: Dr. William Hoyt, Bryan Endowed Chair of Public Finance
Dr. Lala Ma, Associate Professor of Economics
Lexington, Kentucky
2024

https://orcid.org/0009-0009-9683-2605

ABSTRACT OF DISSERTATION

Essays on Impacts of Natural Disasters

This dissertation consists of three chapters that study how natural disasters affect households and individuals. The first chapter examines how Kentucky's housing market responds to changes in flood risk information. The second chapter looks at disproportionate drinking water non-compliance post-disaster. The third chapter studies the effect of flooding events and the national school meals program in Kentucky on education outcomes.

The first chapter examines Kentucky's housing market response to changes in the Federal Emergency Management Agency's (FEMA) flood maps and how the responses differ depending on whether an area has been flooded recently. I use Zillow's ZTRAX property transaction data and current and historical floodplain maps to estimate a hedonic property value model and to recover the price impact of residential properties that have experienced a change in their flood zone status. Importantly, I also allow for the price effect of a flood map change to depend on whether a property was recently affected by a flood event. I find that when properties are switched into flood zones in recently flooded areas, sale prices decrease by 5.2% on average. In contrast, prices increase by 4.7%, on average, when houses are mapped out of a flood zone. Understanding how housing markets respond to flooding events and flood risk can help regulators evaluate the effectiveness of programs aimed to adapt to increasing flood risk, such as disaster assistance programs and the National Flood Insurance Program, and can provide guidance on ways to improve these programs. This research provides evidence for policymakers to provide detailed and personalized information on flood risk to better serve the housing and insurance markets.

The second chapter analyzes drinking water quality at the intersection of race and socioeconomic status and whether the clean-up time is longer in more disadvantaged communities after a flooding event. We match Safe Drinking Water Act (SDWA) violations with county-level demographic and economic information from the U.S. Census. We find that larger minority groups and higher poverty rates are associated with extended non-compliance periods, and post-flooding clean-up times are longer for communities with higher poverty rates. The environmental justice literature has focused on the inequalities of racial and socioeconomic status concerning the application of environmental regulations and the inequitable recovery processes for vulnerable communities. These may help target under-performing systems that might benefit from assistance in achieving consistent compliance. Our results also suggest that attention to the distributional impact of regulatory actions should be incorporated into post-disaster recovery prioritization decisions.

The third chapter studies the effect of flooding events and the national school meals program in Kentucky on education outcomes. Literature has provided evidence of the importance of food security in the disaster recovery phase for children's academic performance, and natural disasters pose significant challenges to the education system in affected regions. With the Kentucky Department of Education's (KDE) school-level academic achievement scores and attendance records, I identify the student cohorts affected by flooding events. I find that affected students with free/reduced-price meals experienced smaller decreased test scores. Understanding what mechanisms the food assistance programs can affect education outcomes can shed light on potential pathways for implementing and improving programs that mitigate the adverse effects of natural disasters on education.

KEYWORDS: natural disaster, flooding, housing market, water quality, academic performance

Fang-Yu Yeh

August 7, 2024

Essays on Impacts of Natural Disasters

By
Fang-Yu Yeh

Dr. William Hoyt and Dr. Lala Ma
Directors of Dissertation

Dr. Carlos Lamarche
Director of Graduate Studies

August 7, 2024
Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Map Updates and Flood Events on Kentucky's Housing Market

1.1 Introduction

Flood events are the most common and costly natural disasters in the U.S., affecting millions of individuals each year. According to the National Centers for Environmental Information, the U.S. has witnessed over $67.8 billion in flood damages since 2010 (Smith, 2020). In Kentucky, flooding is the state's most frequent and costly natural disaster[1]. Kentucky's varying degrees of topography play a role in the state's vulnerability to flooding. Figure 1.1 shows the ratio of total property damage by floods to median home value in the last 2 decades. At the end of July 2022, several counties in Eastern Kentucky were hit by severe flash floods resulting from a week-long heavy rain. The "1-in-1000 year"[2] flood event claimed more than 30 lives and destroyed hundreds of homes, bridges, and roads in the area. Unfortunately, most residents in this area do not have flood insurance because they are not in a floodplain and 6 of those counties do not have an updated flood maps since 2009. The residents are not updated with the information and therefore underestimate their flood risk.

While climate events and their impact on real estate assets are not unprecedented, the increasing prominence of extreme weather in the past few decades has become more apparent. Climate change, population growth, and changes in land use have exposed more people to flood hazards. Flood damages are expected to increase annually, where the financial consequences are being borne by homeowners and the

[1] Estimates are based on the Storm Events Database from NOAA/National Centers for Environmental Information (NCEI). Digital data are available at `http://www.ncdc.noaa.gov/stormevents/ftp.jsp`

[2] According to the United States Geological Survey (USGS), the term "1,000-year flood" means that a flood of that magnitude (or greater) has a 1 in 1,000 chance of occurring in any given year. In terms of probability, the 1,000-year flood has a 0.1% chance of happening in any given year.

government. This creates difficulties for the Federal Emergency Management Agency (FEMA) to accurately quantify a property's risk as flood risks change over time, and thus presents a challenge to all those engaged in the housing market, whether as potential homeowners making housing decisions, as insurers setting actuarially fair insurance rates, or as policymakers choosing appropriate land-use regulations and flood preparation plans.

Economists have been examining the impact of negative shocks from natural disasters on risk attitudes and perceptions. The Bayesian learning model states that individuals update their prior risk beliefs in response to new information. In the case of flooding, there is publicly available information on the likelihood of the event (flood risk maps) and historical flood records for people to adjust their risk beliefs. For example, individuals who recently experienced a flood event will likely increase their perception of risk and they tend to exhibit higher levels of risk aversion afterward (Tversky and Kahneman, 1973). I focus on the housing market's responses to flood risk as a means to adapt to climate change. The housing literature relevant to flood risk suggests that the correction of flood risk will affect home values because home buyers will account for changes in insurance (if changes in flood risks are reflected in insurance payments) and expected damages.

This paper uses flood map updates and flood events in Kentucky to investigate how the housing market responds to flood risks associated with different information: observed flooding and the flood hazard maps put forth by FEMA. This study asks two questions: How does the housing market respond to a change in floodplain status? Second, does the housing market response differ depending on how the update in flood risk information occurs (e.g., due to flood-related events as opposed to an update in flood maps)? Using Zillow's ZTRAX property transaction data and FEMA's floodplain maps, I find that the properties that are switched into the floodplain in an area that has experienced a large flood within a year saw a decrease in price and the property price increases when it is removed from a floodplain in an area without flooding within a year. I also find that the housing markets' responses to flood risk vary by neighborhood characteristics.

The study contributes to the hedonic literature discussing the impacts of flooding events on properties inside and outside the floodplain. Previous studies show that located within a flood zone lowers the property value more after a major flood event (Bin and Polasky, 2004; Kousky, 2010; Bin and Landry, 2013). However, the immediate post-flood discount for properties inside the floodplain diminishes with time (Atreya et al., 2013; Beltrán et al., 2019). Literature also shows that the negative risk salience effect for high-risk properties that were not actually inundated (Bakkensen et al., 2019; Hennighausen and Suter, 2020; Yi and Choi, 2020). To my knowledge, this paper provides the first evidence of the effects of the impact of multiple large regional floods on housing market within one inland state over time.

Additionally, the paper contributes to the recent literature that uses changes in flood risk mapping. Empirical work has shown that the sale prices of previously flood-free properties being assigned into flood zones decrease (Hino and Burke, 2020) but properties previously located in flood zones that become flood-free see no significant impact on sale prices (Shr and Zipp, 2019). Gibson and Mullins (2020) show that after Hurricane Sandy, properties in New York City included in the new floodplain experienced a large price discount compared to those who were not in the new floodplain. Furthermore, home buyers are more responsive to the actual occurrence of a flood event than to the release of flood maps to the public (Rajapaksa et al., 2016). My paper will extend on the context by studying the effects of both changes in the flood maps and flood events on the housing market, comparing how changes in flood risk increase and decrease housing prices between recently flooded areas and areas that have not experienced flooding lately.

The paper proceeds as follows. The first section introduces the institutional background. The second section provides the theory for the research. The third section shows details on the data and variables of interest. The fourth section presents the research design and identification strategies. The fifth section presents the results and assesses robustness. Finally, the sixth section concludes and discusses the limitations of the study.

1.2 Institutional Background

In 1968, Congress passed the National Flood Insurance Act, tasking the Federal Emergency Management Administration (FEMA) with creating and facilitating the NFIP. The NFIP's stated purpose is twofold: 1) to provide access to federally subsidized flood insurance and distribute the cost of flooding and 2) to reduce the nation's flood risk through the implementation of floodplain management standards. To accomplish these goals, the NFIP requires communities to work collaboratively to employ flood risk mitigation strategies and develop Flood Insurance Risk Maps (FIRM). FIRMs delineate Special Flood Hazard Areas (SFHA), which are areas that have a 1% or greater risk of flooding every year. In communities that participate in the NFIP, homeowners of properties in the SFHA are required to purchase flood insurance as a condition of receiving federally backed mortgages or federally regulated mortgages. Moderate to low-risk areas are marked as 500-year floodplains, where the properties have 0.2% flooding risk every year but are not required to purchase flood insurance. The old flood insurance premium rating, which had not fundamentally changed since the 1970s, evaluated the structures' flood risk by the floodzone status, the elevation of the structure relative to the Base Flood Elevation (BFE) in each risk zone, and the occupancy type. Taking effect on all NFIP policyholders in April, 2022, a new premium rating system, Risk Rating 2.0, calculates flood insurance premiums by incorporating a broader range of measures such as the distance to water, the type and size of nearest bodies of water, flood frequency and the elevation of the property relative to the flooding source. According to FEMA, Risk Rating 2.0 will reflect more types of flood risk in the premium rates and provide rates that are easier for policyholders to understand.

Given that the severity of flooding is expected to increase over time, the National Flood Insurance Program Reform Act of 1994 mandated FEMA to review the FIRMs every five years. However, there is no consistent timetable for when a particular community will have its maps revised and updated. Generally, flood maps may require updating when there have been significant new building developments in or near the flood zone, changes to flood protection systems, or environmental changes in the community. Because of the variability in how and when a FIRM is updated, one community may have had its map last updated in 2018 while a neighboring community had its last revised in 2005. Figure 1.2 shows the years in which each county had its last map updates. Of 120 counties in Kentucky, 63 counties have not updated their flood maps since 2011.

1.3 Theory

I apply a Bayesian learning model (Viscusi, 1991) to formulate an individual's subjective perceptions of flood risk as a Bayesian learning process (Gayer et al., 2000). Individuals are assumed to update their prior probability assessment of flood risk based on both new information provided by FEMA and flood events. The updated subjective probability of flooding (π) is a function of the risk from the new information, which includes flood map updates (m) and flooding events (e), and the individual's prior risk belief (k) that is based on the previously assigned flood zone and past experience with flood events:

$$\pi(k, m, e) = \frac{\phi_0 k + \kappa_0 m + \psi_0 e}{\phi_0 + \kappa_0 + \psi_0} \tag{1.1}$$

where ϕ_0, κ_0 and ψ_0 are the information parameters which measure information content associated with, respectively, the prior risk assessment, flood map updates, and flooding events. Denote the weight of each information source on an individual's risk belief as $\phi = \frac{\phi_0}{\phi_0+\kappa_0+\psi_0}$, $\kappa = \frac{\kappa_0}{\phi_0+\kappa_0+\psi_0}$ and $\psi = \frac{\psi_0}{\phi_0+\kappa_0+\psi_0}$. The risk-perception function can be re-written as

$$\pi(k, m, e) = \phi k + \kappa m + \psi e \tag{1.2}$$

The new information may serve as good news or as bad news, therefore, m and e may be less or greater than k. If a house is moved outside of the floodplain, $m < k$, then the individual would lower their risk beliefs. If a house is re-zoned to be in a floodplain, $m > k$, then the individual would increase their risk belief. If new flood maps do not provide any new information to the individuals, the risk belief would remain the same. If a flooding event caused damages to the house or the surrounding areas, we would expect that $e > k$ and the individual would increase their risk belief.

I extend the models of MacDonald et al. (1987); Hallstrom and Smith (2005); Bin and Landry (2013); Kousky (2010); Shr and Zipp (2019) by accounting for the information of flood map updates and recent flood events. The decision is modeled using a state-dependent expected utility framework, where there are two states: flooding (F) and no flooding (NF). Let U_F denote the utility in the flooding state and U_{NF} as the utility when there is no flood. Three assumptions are necessary to establish the household decision making problem: (1) For any given level of income, households prefer being safe, i.e. $U_{NF} > U_F$ (2) Within each state of the world, households are risk-neutral or risk-averse (utility function is quasi-concave); (3) Marginal utility of income is higher when there is no risk. Household utility in each state of the world is defined as:

$$U = U(X, Z) \tag{1.3}$$

where X is a numeraire good and the price is set equal to 1; Z is the set of neighborhood and structural characteristics of the home. The function $P()$ maps housing characteristics, neighborhood attributes, and individuals' risk perceptions to a price:

$$P = P(Z, \pi(k, m, e)) \tag{1.4}$$

Given total income Y, individuals choose the level of X and Z to maximize their utility subject to their budget constraint:

$$\begin{aligned} &Max\ EU = \pi(k, m, e)U(X_F, Z) + [1 - \pi(k, m, e)]U(X_{NF}, Z) \\ &s.t.\ Y = X + P(Z, \pi(k, m, e)) \end{aligned} \tag{1.5}$$

The numeraire X can be expressed as $X_{NF} = Y - P(Z, \pi(k, m, e)) - I(k, m)$ when no flood occurs and as $X_F = Y - P(Z, \pi(k, m, e)) - I(k, m) - L + C$ in the case of a flood. I is the flood insurance premium payment[3]; L is the loss during a flood event; and C is the insurance coverage for a flood event.

The expected utility can be rewritten as

$$\begin{aligned} EU = &\pi(k, m, e)U(Y - P(Z, \pi(k, m, e)) - I(k, m) - L + C, Z) \\ &+ [1 - \pi(k, m, e)]U(Y - P(Z, \pi(k, m, e)) - I(k, m), Z) \end{aligned} \tag{1.6}$$

Under the assumptions that the housing market is a perfectly competitive and the consumers are rational, have identical preferences, perfect information, and perfect mobility,[4] we can solve for the partial derivative of the hedonic function with respect to the risk belief, which gives the marginal implicit price of the risk, or the risk discount.

$$\frac{\partial P}{\partial \pi} = \frac{U_F - U_{NF}}{\pi \frac{\partial U_F}{\partial X} + (1 - \pi) \frac{\partial U_{NF}}{\partial X}} < 0 \tag{1.7}$$

Using the chain rule, we can solve for the partial derivatives of the hedonic function with respect to each information source. The marginal implicit prices of the risk

[3]The NFIP rating methods during the sample period used basic characteristics to classify properties based on flood risks, which are evaluated by the flood zone, occupancy type and the elevation of the structure.

[4]Property prices are higher in an area with better amenities because households would want to move into the areas and drive up the prices. Perfect mobility of households between different locations ensures that the property prices reflect the benefits of amenities.

prior to a new event, the risk associated with the new flood zone information, the risk associated with a new flooding event are, respectively,

$$\frac{\partial P}{\partial k} = \frac{\partial \pi}{\partial k}\frac{\partial P}{\partial \pi} - \frac{\partial I}{\partial k}, \qquad \frac{\partial P}{\partial m} = \frac{\partial \pi}{\partial m}\frac{\partial P}{\partial \pi} - \frac{\partial I}{\partial m}, \qquad \frac{\partial P}{\partial e} = \frac{\partial \pi}{\partial e}\frac{\partial P}{\partial \pi} \tag{1.8}$$

The Bayesian model suggests that people will increase or decrease their willingness to pay for risk reduction after the release of the new map by FEMA or a flooding event. The impact of the new information enters the hedonic price analysis by a comparison of the marginal price of the risk before ($\frac{\partial P}{\partial k}$) and after ($\frac{\partial P}{\partial \pi}$) such events. I discuss the impact of new information on the implicit marginal price of flood risks in three cases.

Case 1: Properties mapped into floodplain

Mapping a property into a floodplain increases the individual's risk belief and requires the individual to purchase flood insurance ($\frac{\partial I}{\partial m} > 0$). The effect of the new information would have a negative impact on housing prices and we would expect an increase in willingness to pay for risk reduction ($\frac{\partial P}{\partial \pi} > \frac{\partial P}{\partial k}$).

Case 2: Properties mapped out of floodplain

For properties mapped outside of a floodplain, the change indicates that flood risk is lower than previously perceived. The individual's risk belief will decrease and they would no longer be required to purchase flood insurance ($\frac{\partial I}{\partial m} < 0$). The effect of the new floodplain status would have a positive impact on housing prices and we would expect a decrease in willingness to pay for risk reduction ($\frac{\partial P}{\partial \pi} < \frac{\partial P}{\partial k}$).

Case 3: Properties that experience a flooding event

Previous studies have confirmed that following a flood event, there is a significant negative effect on the value of properties at risk ($\frac{\partial P}{\partial e} < 0$). Individuals will increase their risk belief and the new information would cause an increase in willingness to pay for risk reduction ($\frac{\partial P}{\partial \pi} > \frac{\partial P}{\partial k}$). Unlike Cases 1 and 2, the price of flood insurance is not experience-rated. The flood insurance rates set by FEMA are at nearly identical rates before and after each flooding event. Therefore, the change in the implicit marginal price is purely associated with the change in the individuals' subjective assessment of flood risk.

1.4 Data

Floodplain maps Using geographic information system (GIS), I match all properties to their flood zone. Current floodplain maps (officially "Digital Flood Insurance Rate Maps") are downloaded as state-level National Flood Hazard Layer (NFHL) from FEMA's Flood Map Service Center. For historical floodplain maps, I obtained Q3 Flood Data from FEMA[5], the first digitization of floodplain maps. They were initially produced in 1996 and updated through 1998. For the counties that have had two updates since 1998 (one update between Q3 and the current flood maps), I acquired the second flood maps from the county offices in Kentucky or from FEMA historical raster files. Each property was overlaid on both the current and historical flood maps and assigned one of two conditions for each time period: in a Special Flood Hazard Area (SFHA, equivalent to the 1% floodplain) or outside the floodplain. Figure 1.3 shows the 70 studied counties in this paper.

Real estate data Property sales and characteristics data are sourced from Zillow's ZTRAX database. I matched each recorded sales event in the transaction table to property attribute information in the assessor table. I include the records in Kentucky from January 2005 to October 2021. The dataset contains the transaction date, sale price, the properties' location, structural characteristics, and residential type (i.e., Single family, condominium, mobile home). Housing prices are converted to 2010 Q1 dollars using the All-Transactions House Price Index for Kentucky (KYSTHPI) from the U.S. Federal Housing Finance Agency. I remove outliers with prices below $10,000 or with prices above 100 million dollars. I constructed geographic variables for each property: the distance to the nearest waterbody using the National

[5]Additional information on Q3 Flood Data is available here: `https://hazards.fema.gov/femaportal/usercare/guidesAndDocs/Documents/flood_map_svc.htm`

Hydrography Dataset (NHD) from the United States Geological Survey (USGS)[6] and the distances to the boundaries of current and historical floodplains.

Table 1.1 provides summary statistics for property and neighborhood characteristics. The average sale price is \$40,000 lower for the houses inside a floodplain compared to properties located outside the floodplain. Properties inside the floodplain have larger lot sizes but smaller square footage. Using the tract-level American Community Survey (ACS) 2015-2019 5-year estimates, the houses inside floodplains are more likely to be in neighborhoods that have lower median income and median home value. Among these transactions, 6,539 (1.6%) are always in the floodplain, 1,974 (0.5%) switched into an SFHA, 3,870 (1%) were mapped out of an SFHA and 400,018 (96.9%) are always outside the floodplain. Table 1.2 reports summary statistics for switchers and non-switchers inside and outside the floodplain.

Flood event data For large regional floods, I use Presidential Disaster Declaration (PDD) Floods events and NFIP redacted claims as data sources. The PDD system is a formalized process to request and receive federal assistance following large natural disasters. PDD Summaries from FEMA provides information on all approved federal disaster declaration requests, including data on the disaster type, disaster event start and end dates, and affected counties.[7] NFIP redacted claims data[8] provides claim transactions on property type, date of loss, flood zone, and the amount paid on claims. I match the date of loss and the location of each property to the incident period of PDD floods to determine if the flood damage is caused by a large regional flood. Since PDD floods are determined at the county level, not all communities within a county are affected by the flood. I construct a variable to identify which communities in PDD counties are "hit" by each flood. I consider a

[6]Additional information on the NHD is available here: `https://www.usgs.gov/national-hydrography/national-hydrography-dataset`

[7]Additional information on the PDD data is available here: `https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2`

[8]Additional information on the NFIP redacted claims data are available here: `https://www.fema.gov/openfema-data-page/fima-nfip-redacted-claims-v1`

community to be hit if there are at least \$100,000 in building claims linked to the PDD floods within the county subdivision.

1.5 Empirical Framework

If home buyers update their risk accordingly with new flood maps or flood events, then the price discount should fully capture the risk information. To test the hypothesis, I employ a difference-in-differences (DID) specification with 4 different approaches. Model 1 estimates the effect of floodplain status and flood events on properties that are initially outside the floodplain:

$$ln(P_{ict}) = \beta_1 SwitchIn_{it} + \beta_2 SwitchIn_{it} * Event_{ct} + \beta_3 Event_{ct} + \alpha Z_i + \kappa_{ct} + \epsilon_{ict} \quad (1.9)$$

Model 2 estimates the effect of floodplain status and flood events on properties that are initially inside the floodplain:

$$ln(P_{ict}) = \beta_4 SwitchOut_{it} + \beta_5 SwitchOut_{it} * Event_{ct} + \beta_6 Event_{ct} + \alpha Z_i + \kappa_{ct} + \epsilon_{ict} \quad (1.10)$$

$ln(P_{i,c,t})$ is log sale price of property i in county c at year t. $SwitchIn_i$ is a dummy variable equal to one if property i is sold and was mapped in SFHA after the flood map updates but outside SFHA before the update. Similarly, $SwitchOut_i$ is a dummy variable equal to one if property i is sold and was mapped out SFHA after the update but was inside SFHA before the update. $Event_t$ is a dummy variable equal to one if a sale occurred after the area has experienced a flood within one or two calendar years and zero if not. Z_i denotes property specific characteristics. κ_{ct} is a fixed effect for each county-quarter, which controls for local market trends. β_1 represents the effect of switching from non-SFHA to SFHA, compared to the properties that are never in the SFHA, and β_4 represents the effect of switching from SFHA to non-SFHA, compared to the properties that are always in the SFHA. β_3 represents the effect of flood events on non-SFHA properties and β_6 represents the effect of flood events on

SFHA properties. β_2 isolates the unique effect of the flood event on the properties that are mapped into the floodplain after the map updates and β_5 is the effect of a flood event on the properties that are mapped outside the floodplain after the map updates.

The hypothesis is that being switched into the floodplain will have a larger risk discount for the properties in areas that have experienced a large regional flood within one year of sale compared to areas that have relatively lower flood risk. Individuals may increase their flood risk belief more as they witness both updates and events: $\beta_2 < \beta_1 < 0$. On the other hand, being switched outside the floodplain will have a larger positive impact for the properties that are in relatively lower flood risk areas. Individuals lower their flood risk belief when they are not required to purchase flood insurance and the area did not experience a PDD flood recently: $\beta_4 > \beta_5$.

Model 3 combines (9) and (10) and estimates the effect of changes in floodplain status and flood events on all properties, comparing to the properties that do not have have change in floodplain status:

$$ln(p_{ict}) = \delta_1 SwitchIn_{it} + \delta_2 SwitchOut_{it} + \delta_3 SwitchIn_{it} * Event_{ct} + \delta_4 SwitchOut_{it} * Event_{ct} \\ + \delta_5 Event_{ct} + \alpha Z_i + \kappa_{ct} + \epsilon_{ict} \quad (1.11)$$

For the properties in areas that did not experience a large regional flood within 1 or 2 years of sale, δ_1 is the effect of switching into floodplain, δ_2 is the effect of switching outside the floodplain, comparing to the houses that did not have a change in flood zone status. Similarly, for the properties in communities that had experienced a PDD flood within 1 or 2 years, δ_3 is the effect of switching into a floodplain, δ_4 is the effect of switching outside the floodplain. This model allows us to look at the effect of updating flood maps in areas that have/have not experienced flooding recently.

1.6 Empirical Results

Before showing the price effects of both flood map updates and events, I first present the effects of being inside a floodplain. Table 1.3 presents these results. All specifications include county-by-quarter fixed effects and block fixed effects. Column 1 shows that house prices inside the flood zone are 4.77% lower than those outside. Column 2 shows that the flood risk discount increases to 7.34% if the area has experienced a PDD flood within 1 year. The increase in discount reflects updated expectations of future flooding and costs related to inundation, such as damage. The estimated flood discount is consistent with previous studies that find marginal impacts of flood risk ranging from 1.1% to 28.7%.

Figure 1.4 plots the flood zone effects by distance to the flood zone boundary. For properties located in communities with no flooding within 1 year of sale, prices decrease 4% just inside the flood zone comparing to the ones just outside the flood zone boundary. For properties located in communities with no flooding within 1 year of sale, prices decrease 8% just inside the flood zone compared to the ones just outside the flood zone boundary. The hypothesis is that for the houses that gets switched in/out around the flood zone boundary, we would see a larger price difference for the areas that had been flooded recently.

1.6.1 Main Specification Results

Table 1.4 presents the main estimates corresponding to equations (9), (10), and (11). Model 1 shows the effect of mapping into a SFHA is statistically insignificant for communities that did not experience a PDD flood within 1 or 2 years. For the houses that switched into a flood zone in areas that have experienced a PDD flood within 1 year, the housing prices are 6.53% lower than those of houses that stayed outside the floodplain in the communities where there was no PDD flood within 1 year. This is equivalent to an average decrease of $11,152.02 ($170,781.4*0.0653) in adjusted sale prices. Model 2 shows that the estimated effects of switching out from the floodplain are statistically significant and range from 3.96% for communities that experienced a large flood within 1 year to 4.09% for communities that experienced a large flood within 2 years. This is equivalent to an average increase of $5,379.51 ($135,846.1*0.0396) to $5,556.11 ($135,846.1*0.0409). Model 3 shows that compared to the houses that did not have a change in flood zone status, houses that are removed from the floodplain experience price increases of 5.17% but the effect of being mapped in is not significant. All models show insignificant effects of switching into a flood zone in areas without a major flood and switching out from a floodplain in areas that experienced one within 1 year. These findings suggest that home buyers in those areas do not internalize the potential increase/decrease in flood risk solely with the information provided by FEMA's updated maps. The responses to the changes in flood risks based on the area's flooding history and the updated flood zones.

Table 1.5 re-estimates the main specifications from equations (9), (10), and (11) with property fixed effects. By comparing the same property over time, it is possible to control for unobserved, time-invariant characteristics that are correlated with flood risk and contribute to price. The drawback of the repeat sales model is that it assumes that there are no structural changes such as physical improvements in the property between sales. The results from Model 3 show that, comparing to the estimates with block fixed effects, the estimates are larger for the effect of switching out and the effect of switching in in recently flooded areas. The comparison indicates that the repeat sales model controls for, at least partly, the omitted variable biases stemming from using coarser fixed effects.

1.6.2 Heterogeneous Impacts of Floodings on Properties

Due to the absence of individual home buyer information, I use neighborhood characteristics such as median income and median home value by census tract level to examine if the effects of map updates and flooding events vary by socio-economic status. Table 1.6 reports the estimates for different neighborhood categories. The effect of switching into flood zones in a flood-prone area is largely driven by the properties in higher income tracts, while properties in lower income communities see significant price increases when mapped out of floodplains in a non flood-prone area. This suggests that communities with lower socio-economic status may have lower salience of flood risk in regards to whether the property is located in a higher risk area. From the summary statistics, houses inside floodplain are likely to be in lower income/home value neighborhoods, home buyers in low socio-economic status communities are more likely to be attracted to properties that are removed from a floodplain in a flood-free area. For higher income/home value neighborhoods, the houses that are now at higher flood risk in areas that have been recently flooded are less desirable and the home buyers may take future flood damage and cost into consideration.

Following Gibson and Mullins (2020), I examine whether the changes in flood risk belief by property values are partially responsible for the observed price changes. They hypothesize that the properties with structural values that are below the flood insurance coverage cap ($250,000) would have smaller effects of switching flood zone status because there is little to no uninsured value and premiums increase slowly in structural value. For the houses above the cap, one would expect larger effects of switching as the prices increase. Figure 1.5 plots the effects in $75,000 property value bins. The effects of switching out are significant for properties below the cap, which suggests that higher home value buyers do not recognize the reduction in flood risk. However, the interaction effect for switching in and flood events is significant but in the opposite direction. Due to data limitations, I cannot observe the elevation of the building and whether there are structural improvements for reducing flood risk on the property. Home buyers may decrease their risk belief for higher value houses in flood-prone areas as they are more likely to be elevated above the base flood elevation or have flood mitigation on the property.

1.6.3 Potential Threats to Identification

A key identifying assumption in a Difference-and-differences model is that treatment and control groups have common counterfactual trends, which means that in the absence of the treatment, the treatment and the control groups would have changed in the same way during the post-treatment period. I test this assumption using event study models. Results are shown in Figure 1.6. I use 3 months as a time unit, where period 0 represents the 3 months before the flood map updates. The pre-updates period exhibits no significant differential trends. Switching into the flood zone in areas without flooding recently and switching out in areas with flooding have small and insignificant effects on properties before and after flood map updates. Beginning in the 3 quarters after the map updates, the price of the properties that are switched into the flood zone in flood-prone fall by 1%. After 1 year of the updates, properties that are switched out from floodplain in non flood-prone areas increase by 1% in price.

A potential threat to identification is the perfect information assumption of buyers. If buyers are not aware of the flood zone status of properties, the flood risk discount could be only a lower bound. Passed in July 2000, Kentucky Revised Statutes §324.360 requires sellers of single-family residential properties to make certain disclosures to potential buyers. This law included a Seller's disclosure of conditions form and questions regarding the property's flooding history and the flood zone classification. Therefore, I consider the impact of information asymmetry between sellers and buyers to be minimal in the cases of properties in Kentucky and that buyers are well-informed about the flood risk.

A second potential threat is that houses in different flood zones are systematically different and that these differences are time-varying and unobserved by the researcher. If so, using properties from all over the state to construct a counterfactual price path could introduce bias. Therefore, I restrict the sample to the properties within 300 meters of the floodplain boundary and report the estimates in Table 1.7. The results are similar to the main results in Table 1.4 with slightly larger effects.

A third potential threat is the collinearity between flood map updates and flooding events. If new flood maps were released soon after a flooding event in response to the concern of outdated flood risk information (i.e., the dummy variables of switch in and switch out are correlated with the event dummy variable.), then the estimated standard errors would be larger therefore reducing the statistical significance. Table 1.8 shows the results of restricting the sample to the counties that had new flood map updates after 6 months of a flooding event. The results are similar to the main results in Table 1.4 so I consider the collinearity issue to be limited.

One worry with the main estimation might be that there is a change in which kinds of houses end up being sold within a year or two of a flood. If we see significant increase or decrease in the number of houses inside the flood zone comparing to the houses outside floodplain being sold, the negative selection could affect the housing market and the housing price. Figure 1.7 shows the number of housing sold by flood zone status before and after a flooding event. The volume sold for all types of flood zone status follow the same trend and I do not see a significant increase or decrease of houses being sold within one year of flood.

Due to data availability, a limitation of the study is that it may under/overestimate the risk belief since I do not observe the structural characteristics such as elevation level and flood mitigation. Homeowners and communities can submit Letter of Map Change (LOMC) and Letter of Map Amendment (LOMA) to remove properties from the floodplain if they believe that the property was incorrectly identified. Identifying the precise flood zone status for properties is of interest for future work. Another future avenue of investigation would involve distinguishing inundated structures and ”near misses”, defined as structures not directly flooded but located inside the floodplain. Previous literature (Bakkensen et al., 2019) has shown that home buyers perceive inundated properties as being riskier and near misses as relatively less risky. Given these considerations, recovering flooding history and monetary damages may help to better explain the full range of behavioral responses to flood events.

1.7 Discussion and Conclusion

This study uses a hedonic pricing framework to investigate how the housing market in Kentucky reacts to information from flood map updates provided by FEMA and from flooding events. The paper contributes to the literature by comparing the changes in flood risk belief associated with both flooding history and changes in flood zone status. The results show that housing values decrease by 6% when a property is assigned to a flood zone where the area has experienced a large flood within a year and that housing values increase by 4% when a property is removed from a flood zone where the area has not experienced a large flood recently.

However, the effects are not symmetric. Housing prices do not rebound when removed from a recently flooded area and do not drop when assigned into a flood zone in an area with no flooding within 1 or 2 years. This indicates that the mapping of properties into floodplains is generally not internalized by residents in areas that have not experienced flood events recently, even when facing mandatory flood insurance costs. Similarly, the removal of properties from floodplains in the areas that witnessed flooding recently does not reduce home buyers’ flood risk beliefs. These results also provide evidence of heterogeneous responses to flood risk information within different communities and different property values.

The findings imply that individuals’ responses to changes in flood risk are based on both recent flooding and flood maps provided by FEMA. The findings suggest some potential improvements to the National Flood Insurance Program. First, FEMA's floodplain maps should provide more detailed and personalized information on flood risk to better serve the housing and insurance markets. FEMA's new flood insurance premium rating system, Risk Rating 2.0, incorporates a wider measurement to calculate each property's individual risk. The additional information should take other relevant factors such as previous flooding events into account. Secondly, FEMA and local governments can increase education and outreach efforts about flood risk and the importance of flood insurance in order to reduce the asymmetric responses by home buyers from different socio-economic statuses. The awareness of differences in the responses/behaviors of home buyers on flood risk is also important for banks and other financial institutes in order to implement appropriate mortgage plans.

1.8 Tables

Table 1.1: Summary Statistics by flood zone status: SFHA and non-SFHA

	Inside Floodplain		Outside Floodplain	
	Mean	S.D.	Mean	S.D.
Housing Attributes				
Price	126,531	183,193	170,699	400,699
House age when sold	45.91	24.72	39.06	30.69
Bedrooms	1.484	1.570	2.039	1.559
Bathrooms	1.563	0.836	1.886	0.924
Lot size (sqft)	142,152	1.007e+06	114,134	825,556
Square footage	1,482	713.5	1,713	1,548
Dummy for single family	0.941	0.236	0.936	0.246
Dummy for condo	0.0275	0.164	0.0348	0.183
Dummy for mobile home	0.0235	0.151	0.0123	0.110
Dummy for townhouse	0.00840	0.0913	0.0174	0.131
Distance to nearest waterbody (meters)	84.40	100.9	252.9	343.0
Neighborhood Characteristics				
Median income	58,270	29,531	67,849	28,289
Median home value	157,554	99,932	188,179	92,431
Fraction of white	81.20	16.27	81.04	16.99
Fraction in poverty	15.39	9.007	12.36	10.41
Population	4,691	1,749	5,105	2,073
Observations	8,567		408,589	

Notes. Table provides the mean attributes of houses and the neighborhoods inside the floodplain to ones outside the floodplain. Tract-level neighborhood attributes are from American Community Survey (ACS) 2015-2019 5-year estimate.

Table 1.2: Summary Statistics by flood zone status: switching and non-switching

	Switch In		Switch Out		Always In		Never In	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Housing Attributes								
Price	125,837	216,553	151,290	164,609	126,733	172,309	170,887	402,304
House age when sold	38.53	27.87	34.35	22.36	48.05	23.29	39.10	30.76
Bedrooms	2.055	1.574	1.805	1.620	1.317	1.529	2.041	1.558
Bathrooms	1.553	0.860	1.819	0.839	1.566	0.830	1.886	0.925
Lot size (sqft)	257,955	1.556e+06	184,593	973,907	108,500	775,438	113,452	823,961
Square footage	1,536	652.3	1,591	841.4	1,466	729.6	1,715	1,553
Dummy for single family	0.884	0.320	0.894	0.308	0.957	0.203	0.936	0.245
Dummy for condo	0.0575	0.233	0.0722	0.259	0.0188	0.136	0.0344	0.182
Dummy for mobile home	0.0560	0.230	0.0135	0.116	0.0140	0.118	0.0123	0.110
Dummy for townhouse	0.00259	0.0509	0.0204	0.141	0.0101	0.1000	0.0174	0.131
Distance to nearest waterbody (meters)	71.31	78.89	90.66	92.02	88.20	106.2	254.5	344.1
Neighborhood Characteristics								
Median income	54,806	23,647	60,194	23,212	59,277	30,962	67,923	28,324
Median home value	143,903	73,850	169,987	76,235	161,522	105,994	188,356	92,556
Fraction of white	87.03	11.64	80.11	12.81	79.51	17.02	81.05	17.02
Fraction in poverty	16.55	9.709	13.78	9.536	15.05	8.765	12.34	10.42
Population	4,761	1,807	5,192	1,867	4,671	1,732	5,104	2,075
Observations	1,929		6,638		6,591		401,901	

Notes. Table provides the mean attributes of houses and the neighborhoods inside the floodplain to ones outside the floodplain. Tract-level neighborhood attributes are from American Community Survey (ACS) 2015-2019 5-year estimate.

Table 1.3: Effect of being in the floodzone on housing price

	(1)	(2) Within 1year	(3) Within 2years
SFHA	-0.0477***	-0.0442***	-0.0419***
	(0.00879)	(0.00980)	(0.0106)
Event		-0.00998	-0.00999
		(0.00932)	(0.00978)
SFHA*Event		-0.0734***	-0.0694***
		(0.0144)	(0.0126)
ln(Lot size)	0.0434***	0.0433***	0.0434***
	(0.00975)	(0.00975)	(0.00975)
ln(Squared Footage)	0.498***	0.499***	0.499***
	(0.0184)	(0.0184)	(0.0184)
squared House age when sold	-1.09e-06	-1.09e-06	-1.09e-06
	(1.02e-06)	(1.02e-06)	(1.02e-06)
Bedrooms	0.00914***	0.00913***	0.00914***
	(0.00314)	(0.00314)	(0.00314)
Bathrooms	0.0897***	0.0897***	0.0897***
	(0.00716)	(0.00716)	(0.00716)
Dummy for single family	0.0898***	0.0898***	0.0897***
	(0.0261)	(0.0261)	(0.0261)
Dummy for condo	-0.0820**	-0.0820**	-0.0820**
	(0.0339)	(0.0339)	(0.0339)
Dummy for mobile home	-0.364***	-0.364***	-0.364***
	(0.0369)	(0.0369)	(0.0369)
ln(Distance to nearest waterbody)	-0.000562	-0.000563	-0.000557
	(0.00233)	(0.00233)	(0.00233)
Constant	7.405***	7.407***	7.408***
	(0.148)	(0.148)	(0.148)
Observations	410,325	410,325	410,325
R-squared	0.692	0.692	0.692
County by quarter FE	Yes	Yes	Yes
Block FE	Yes	Yes	Yes

Notes. Dependent variable is log sale price. Standard errors are stated in parentheses and are clustered at county subdivision level. Price, lot size, squared footage, and distance to nearest waterbody were transformed with natural logs. House age was transformed by squaring the variable. *** p<0.01, ** p<0.05, * p<0.1

Table 1.4: Effect of map updates and flood events within 1 year or 2 years on housing price

	Model 1		Model 2		Model 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	Within 1year	Within 2years	Within 1 year	Within 2 years	Within 1 year	Within 2 years
SwithIn	-0.0136	-0.0169			-0.0142	-0.0158
	(0.0220)	(0.0231)			(0.0207)	(0.0221)
SwithIn*Event	-0.0653**	-0.0356			-0.0464	-0.0189
	(0.0276)	(0.0311)			(0.0287)	(0.0317)
SwitchOut			0.0396**	0.0409**	0.0517***	0.0588***
			(0.0157)	(0.0194)	(0.0139)	(0.0131)
SwitchOut*Event			-0.0103	-0.0155	0.00776	-0.0165
			(0.0252)	(0.0178)	(0.0184)	(0.0140)
Event	-0.00947	-0.00881	-0.0402*	-0.0486***	-0.0104	-0.0101
	(0.00928)	(0.00990)	(0.0230)	(0.0160)	(0.00936)	(0.00989)
ln(Lot size)	0.0426***	0.0426***	0.0349***	0.0350***	0.0434***	0.0434***
	(0.00983)	(0.00983)	(0.0105)	(0.0104)	(0.00975)	(0.00975)
ln(Squared Footage)	0.500***	0.500***	0.440***	0.440***	0.499***	0.499***
	(0.0184)	(0.0184)	(0.0498)	(0.0498)	(0.0184)	(0.0184)
squared House age when sold	-1.05e-06	-1.05e-06	-2.48e-05***	-2.48e-05***	-1.09e-06	-1.09e-06
	(9.87e-07)	(9.87e-07)	(7.18e-06)	(7.21e-06)	(1.02e-06)	(1.02e-06)
Bedrooms	0.00910***	0.00911***	0.00990*	0.00994*	0.00914***	0.00915***
	(0.00323)	(0.00323)	(0.00561)	(0.00568)	(0.00314)	(0.00314)
Bathrooms	0.0893***	0.0893***	0.0628***	0.0629***	0.0897***	0.0897***
	(0.00726)	(0.00726)	(0.0170)	(0.0168)	(0.00716)	(0.00716)
Dummy for single family	0.0867***	0.0867***	0.221**	0.222**	0.0896***	0.0896***
	(0.0264)	(0.0265)	(0.0949)	(0.0960)	(0.0261)	(0.0261)
Dummy for condo	-0.0848**	-0.0848**	0.141	0.143	-0.0822**	-0.0822**
	(0.0338)	(0.0338)	(0.148)	(0.148)	(0.0339)	(0.0339)
Dummy for mobile home	-0.367***	-0.367***	-0.0559	-0.0484	-0.364***	-0.364***
	(0.0373)	(0.0374)	(0.188)	(0.188)	(0.0369)	(0.0369)
ln(Distance to nearest waterbody)	-0.000593	-0.000589	0.00701	0.00729	-0.000616	-0.000609
	(0.00235)	(0.00235)	(0.0143)	(0.0143)	(0.00234)	(0.00233)
Constant	7.410***	7.411***	7.713***	7.719***	7.406***	7.407***
	(0.147)	(0.147)	(0.368)	(0.368)	(0.148)	(0.148)
Observations	400,042	400,042	9,284	9,284	410,325	410,325
R-squared	0.691	0.691	0.763	0.763	0.692	0.692
County by quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Dependent variable is log sale price. Standard errors are stated in parentheses and are clustered at county subdivision level. Price, lot size, squared footage, and distance to nearest waterbody were transformed with natural logs. House age was transformed by squaring the variable. *** p<0.01, ** p<0.05, * p<0.1

Table 1.5: Effect of map updates and flood events within 1 year or 2 years on housing price: Repeated sales

	Model 1		Model 2		Model 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	Within 1year	Within 2years	Within 1 year	Within 2 years	Within 1 year	Within 2 years
SwithIn	0.155	0.149			0.0318	0.0382
	(0.127)	(0.125)			(0.0454)	(0.0451)
SwithIn*Event	0.0427	0.0722			-0.0947**	-0.0577
	(0.126)	(0.129)			(0.0374)	(0.0426)
SwitchOut			0.180*	0.180*	0.123*	0.127*
			(0.107)	(0.107)	(0.0646)	(0.0649)
SwitchOut*Event			0.105	0.114	-0.0134	-0.0258
			(0.100)	(0.105)	(0.0255)	(0.0241)
Event	-0.0120	-0.0137	-0.0541	-0.0671**	-0.0131	-0.0153
	(0.0103)	(0.0110)	(0.0343)	(0.0299)	(0.0105)	(0.0110)
squared House age when sold	6.76e-05**	6.66e-05**	9.72e-05*	8.97e-05*	6.73e-05**	6.62e-05**
	(3.05e-05)	(3.06e-05)	(4.91e-05)	(4.97e-05)	(3.04e-05)	(3.05e-05)
Constant	11.59***	11.60***	11.23***	11.26***	11.59***	11.59***
	(0.0695)	(0.0701)	(0.115)	(0.114)	(0.0691)	(0.0698)
Observations	279,928	279,928	5,695	5,695	286,618	286,618
R-squared	0.811	0.811	0.866	0.866	0.813	0.813
County by quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Property FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Sample limits to properties sold more than once during the studied period. Dependent variable is log sale price. Standard errors are stated in parentheses and are clustered at county subdivision level. Price, lot size, squared footage, and distance to nearest waterbody were transformed with natural logs. House age was transformed by squaring the variable. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.6: Hetergeneous effect by neighborhood characteristics

	Lower income tracts			Higher income tracts		
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
SwithIn	-0.0535		-0.0344	0.00667		0.00799
	(0.0394)		(0.0277)	(0.0268)		(0.0252)
SwithIn*Event	-0.0535		-0.000208	-0.0725**		-0.116*
	(0.0557)		(0.0326)	(0.0317)		(0.0663)
SwitchOut		0.0818***	0.0648***		0.0251	0.0396
		(0.0212)	(0.0128)		(0.0223)	(0.0245)
SwitchOut*Event		-0.0110	-0.0260		-0.00702	0.0556*
		(0.0677)	(0.0188)		(0.0303)	(0.0302)
Event	0.0202	-0.0364	0.00248	-0.0137**	-0.0372	-0.0105*
	(0.0142)	(0.0350)	(0.00992)	(0.00630)	(0.0279)	(0.00564)
Observations	94,924	3,092	160,576	304,932	6,099	249,653
R-squared	0.620	0.727	0.628	0.630	0.772	0.608
County by quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes	Yes	Yes
	Lower home value tracts			Higher home value tracts		
	(7)	(8)	(9)	(10)	(11)	(12)
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
SwithIn	-0.0162		-0.0339	-0.0108		0.00375
	(0.0414)		(0.0319)	(0.0235)		(0.0262)
SwithIn*Event	0.00839		-0.0376	-0.0855***		-0.0659*
	(0.0586)		(0.0640)	(0.0278)		(0.0339)
SwitchOut		0.0472***	0.0645***		0.0488**	0.0377*
		(0.0158)	(0.0147)		(0.0228)	(0.0208)
SwitchOut*Event		-0.0221	-0.0125		0.00963	0.0264
		(0.0249)	(0.0220)		(0.0312)	(0.0277)
Event	0.0227	-0.0406***	0.00714	-0.0126	-0.0445	-0.0105*
	(0.0181)	(0.0146)	(0.0107)	(0.00759)	(0.0337)	(0.00608)
Observations	85,341	3,849	132,301	314,600	5,280	277,962
R-squared	0.559	0.582	0.559	0.618	0.775	0.607
County by quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Samples are divided by tract-level median income and median home value in Kentucky from American Community Survey (ACS) 2015-2019 5-year estimate. Dependent variable is log sale price. Standard errors are stated in parentheses and are clustered at county subdivision level. Control variables include lot size, squared footage, house age, number of bedrooms, number of bathrooms, residential type, and distance to nearest waterbody. Price, lot size, squared footage, and distance to nearest waterbody were transformed with natural logs. House age was transformed by squaring the variable. *** p<0.01, ** p<0.05, * p<0.1

Table 1.7: Robustness: Exclude properties outside 300m of flood zone boundary

	Model 1		Model 2		Model 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	Within 1year	Within 2years	Within 1 year	Within 2 years	Within 1 year	Within 2 years
SwithIn	-0.0249	-0.0265			-0.0235	-0.0240
	(0.0214)	(0.0232)			(0.0202)	(0.0222)
SwithIn*Event	-0.0726**	-0.0482*			-0.0445	-0.0227
	(0.0279)	(0.0283)			(0.0270)	(0.0288)
SwitchOut			0.0451***	0.0473**	0.0625***	0.0693***
			(0.0166)	(0.0204)	(0.0142)	(0.0137)
SwitchOut*Event			-0.00569	-0.00751	-0.000866	-0.0200*
			(0.0252)	(0.0173)	(0.0176)	(0.0117)
Event	-0.00509	-0.00452	-0.0313	-0.00751	-0.00791	-0.00774
	(0.0121)	(0.0101)	(0.0205)	(0.0173)	(0.0120)	(0.0102)
Observations	127,616	127,616	8,874	8,874	137,454	137,454
R-squared	0.694	0.694	0.763	0.763	0.696	0.696
County by quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Sample limits to properties inside 300m of flood zone boundary. Dependent variable is log sale price. Standard errors are stated in parentheses and are clustered at county subdivision level. Control variables include lot size, squared footage, house age, number of bedrooms, number of bathrooms, residential type, and distance to nearest waterbody. Price, lot size, squared footage, and distance to nearest waterbody were transformed with natural logs. House age was transformed by squaring the variable. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.8: Robustness: Exclude counties that had flood maps updated within 6 months after a flooding event

	Model 1		Model 2		Model 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	Within 1year	Within 2year	Within 1year	Within 2year	Within 1year	Within 2year
SwithIn	-0.0179	-0.0216			-0.0164	-0.0186
	(0.0242)	(0.0252)			(0.0228)	(0.0242)
SwithIn*Event	-0.0685**	-0.0327			-0.0455	-0.0153
	(0.0295)	(0.0345)			(0.0310)	(0.0344)
SwitchOut			0.0444***	0.0474**	0.0497***	0.0588***
			(0.0157)	(0.0199)	(0.0146)	(0.0138)
SwitchOut*Event			-0.0169	-0.0171	0.00416	-0.0256
			(0.0272)	(0.0197)	(0.0187)	(0.0160)
Event	-0.00738	-0.00296	-0.0418*	-0.0480***	-0.00839	-0.00454
	(0.0106)	(0.0107)	(0.0239)	(0.0166)	(0.0107)	(0.0108)
Observations	351,225	351,225	8,527	8,527	360,711	360,711
R-squared	0.681	0.681	0.753	0.753	0.681	0.681
County by quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Block FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Dependent variable is log sale price. Standard errors are stated in parentheses and are clustered at county subdivision level. Control variables include lot size, squared footage, house age, number of bedrooms, number of bathrooms, residential type, and distance to nearest waterbody. Price, lot size, squared footage, and distance to nearest waterbody were transformed with natural logs. House age was transformed by squaring the variable. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

1.9 Figures

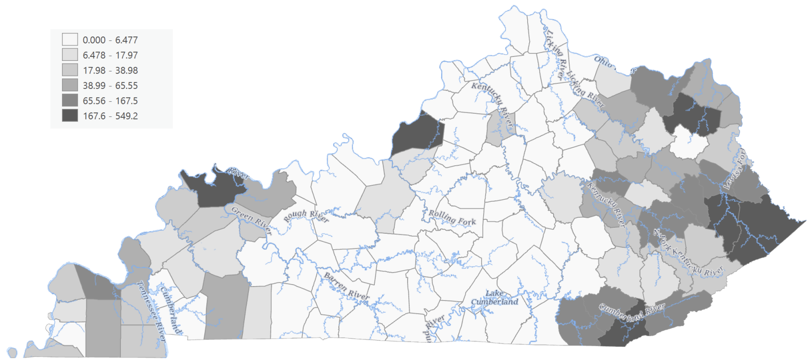

Figure 1.1: Ratio of total property damage by floods to from 2000 to 2021 to median home value

Notes. Property damage data from Storm Data developed by National Weather Service. Estimated median home value from American Community Survey (ACS) 2015-2019 5-year estimate.

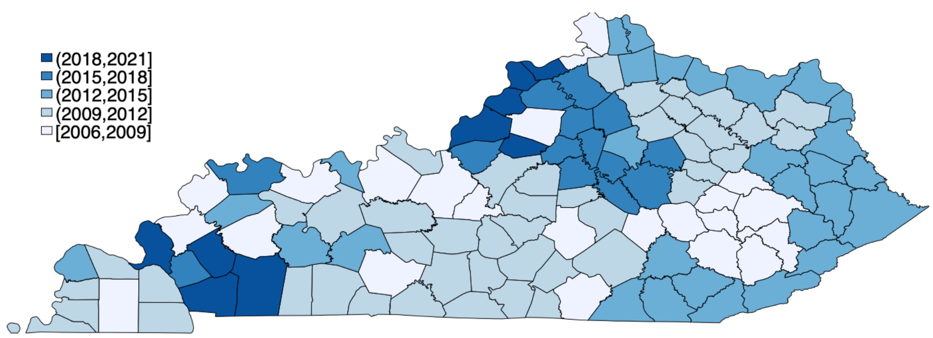

Figure 1.2: Most recent flood maps by year and by county

Notes. The figure shows county-level current flood maps' effective dates by year. Flood maps' effective dates are from FEMA flood map service center: `https://msc.fema.gov/portal/home`

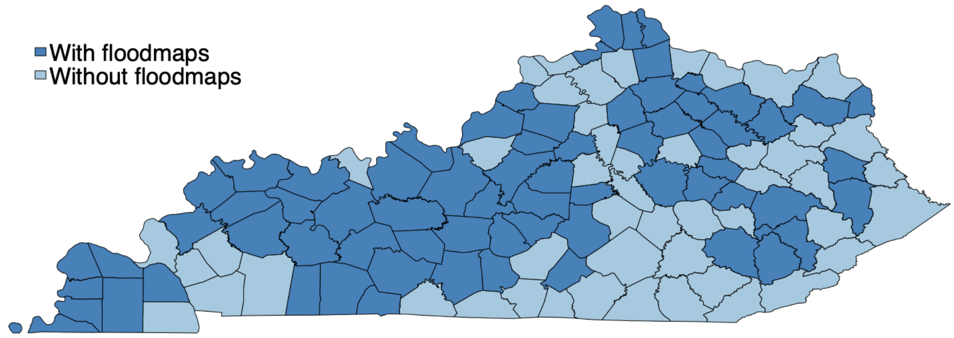

Figure 1.3: Studied counties

Notes. The figure shows the counties included in the studied sample. Current floodplain maps are from FEMA's Flood Map Service Center. Historical floodplain maps are from the county offices or the historical raster files from FEMA.

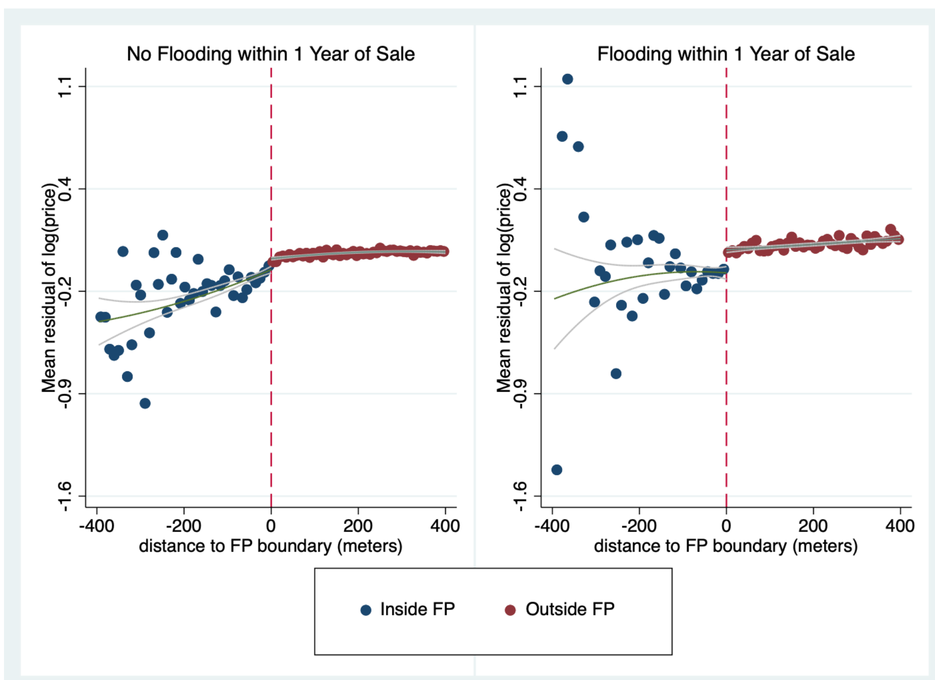

Figure 1.4: Flood zone effects by distance to the flood zone boundary

Notes. Log sale prices are regressed on a set of control variables and the coefficients are the average of log prices that belong to different bins by distance to the boundary. All averages are normalized to the bin inside floodplain closest to the boundary.

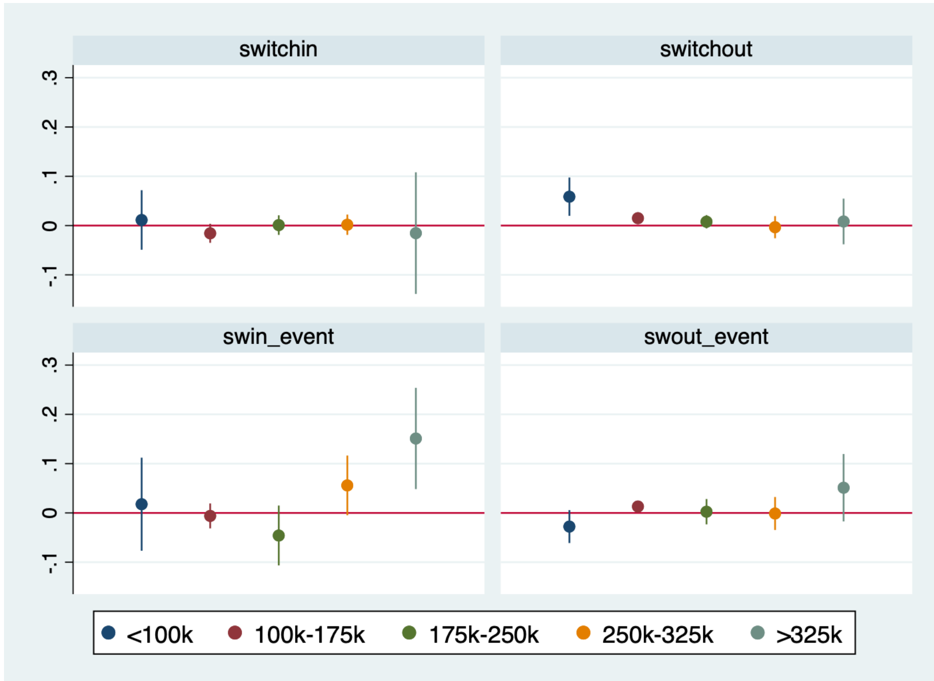

Figure 1.5: Heterogeneous effects by property value

Notes. Observations are divided into bins based on the sale prices.

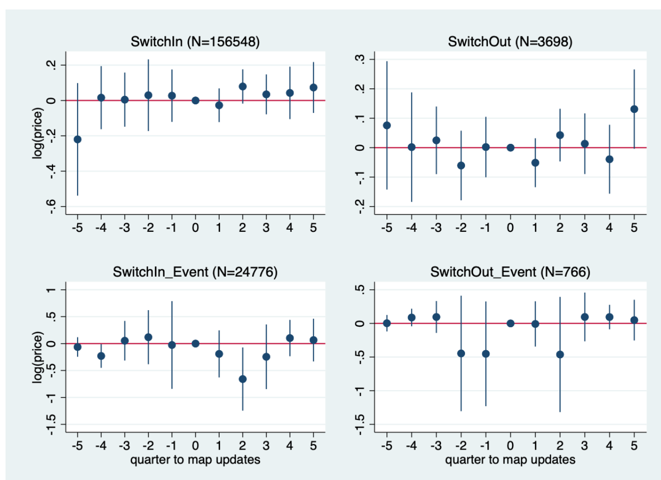

Figure 1.6: Event study

Notes. Each period is 3 months. Period 0 represents 3 months before the flood map updates. Control groups: 1) properties stayed outside and without a flood within 1 year, 2) properties stayed outside and with a flooding recently, 3) properties stayed inside and without a flooding, and 4) properties stayed inside and with a flooding.

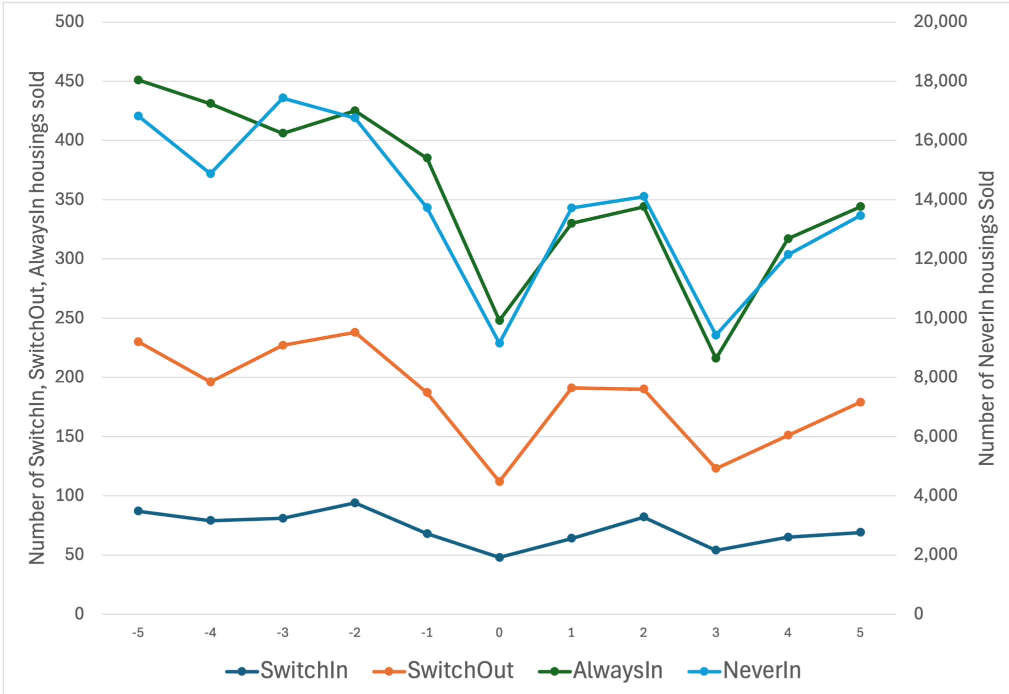

Figure 1.7: Change in volumes sold by floodzone status

Notes. Each period is 3 months. Period 0 represents 3 months before a flooding event.

Chapter 2 H 2 Oh No! Drinking Water Noncompliance and Environmental Justice

2.1 Introduction

Access to clean drinking water is essential for human health. If drinking water contains unsafe levels of microbes or chemical contaminants, it can cause gastrointestinal illnesses, nervous system damage, reproductive effects, and increased risk of cancer (Smith et al., 2000; Chen et al., 2017; Ward et al., 2018). In 1974, Congress passed the Safe Drinking Water Act (SDWA) to regulate the public drinking water supply. Authorized by the SDWA, the Environmental Protection Agency (EPA) set national health-based standards to protect against naturally occurring and man-made contaminants that may be found in drinking water. These standards include maximum contaminant levels (MCLs) for various substances, treatment techniques, and monitoring requirements to safeguard public health. Despite the intent of the SDWA to provide clean drinking water for everyone, some communities throughout the country have disproportionately experienced compromised water systems. High-impact agriculture can lead to runoff containing fertilizers and pesticides, contaminating water sources with nitrates and other harmful chemicals. Industrial pollution from factories and manufacturing plants can introduce a range of toxic substances into the water supply, including heavy metals and volatile organic compounds. Additionally, failing infrastructure, such as aging pipes and water treatment facilities, exacerbates the problem by allowing contaminants to enter the drinking water system.

Recent incidents such as the water crises in Flint, Michigan, and Jackson, Mississippi, have brought to light significant disparities in water quality and safety. These events raise the question of whether they are outliers or part of a predictable pattern of environmental injustice. Research into environmental justice, such as the Drinking

Water Disparities Framework developed by (Balazs and Ray, 2014), suggests that these cases may indeed be predictable. The framework explains that disadvantaged communities often lack the funding and expertise to comply with water quality regulations, leading to disparities in exposure to contaminants. In Flint and Jackson, predominantly low-income and minority communities suffered from systemic failures and prolonged exposure to harmful contaminants, reflecting broader trends identified in studies linking socioeconomic and racial factors to poor water quality.

The impact of these disparities is profound and multifaceted. In Flint, for example, residents were exposed to lead-contaminated water for over a year before decisive action was taken. This exposure has had lasting effects on the health of the community, particularly among children, who are more vulnerable to the neurotoxic effects of lead. Similarly, in Jackson, the predominantly Black community faced prolonged periods without access to safe drinking water, leading to increased health risks. These crises highlight how environmental injustices compound the challenges faced by already marginalized communities, exacerbating health disparities and undermining trust in public institutions. Other studies have shown that communities with higher proportions of hispanic, and communities with lower rates of home ownership are associated with higher nitrate and arsenic level (Balazs et al., 2011, 2012; Schaider et al., 2019). And predominantly Black communities are more likely to be exposed to total coliform bacteria, which could lead to acute gastrointestinal illness (Stillo and MacDonald Gibson, 2017). Other studies have found that community water systems that serve higher minority populations committed more health-based violations and are more prone to repeated violations (McDonald and Jones, 2018; Switzer and Teodoro, 2018; Allaire et al., 2018).

Understanding the intersection of environmental justice research on water quality with the increased frequency of flooding is crucial, as natural disasters often disproportionately impact marginalized communities (Cutter, 2012; Johnson, 2008; Ryder,

2017) and evidence indicates that flooding and extreme rainfalls impact water quality, increasing microbial and chemical loads in surface water (Mishra et al., 2021; Ten Veldhuis et al., 2010; Andrade et al., 2018; Yard et al., 2014), which could exacerbate existing inequalities in access to clean and safe water. Those marginalized communities often reside in areas with poorer infrastructure, which is less capable of withstanding natural disasters, resulting in more significant damage and longer recovery times. Post-disaster recovery time and effort are coincided with social, political, and economic characteristic of the population (Finch et al., 2010; Cutter et al., 2012; Blaikie et al., 2014). Limited access to financial resources, emergency services, and healthcare makes it harder to prepare for, respond to, and recover from disasters.

This paper uses SDWA water system violations and county-level socio-demographic data to investigate whether socioeconomic disparities exist in the duration of exposure to contaminants in drinking water. Additionally, we examine associations between post-flooding non=compliance period and county-level socio-demographic factors. To our knowledge, this is the first paper that looks at non-compliance period of violation on the national scale and provides new insights into the literature on environmental justice associated with drinking water. We find that larger minority groups and higher poverty rates are associated with extended noncompliance periods. This study also expands on the literature on inequitable recovery processes for vulnerable communities, and we show that post-flooding clean-up times are longer for communities with higher poverty rates. These may help target under-performing systems that might benefit from assistance in achieving consistent compliance. Our results also suggest that attention to the distributional impact of regulatory actions should be incorporated into post-disaster recovery prioritization decisions.

The paper proceeds as follows. The first section shows details on the data and variables of interest. The second section presents the research design and identification strategies. The third section presents the results, assesses robustness and discusses the limitation of the study. Finally, the forth section concludes this paper.

2.2 Data

Drinking water data We obtained drinking water system violation reports from EPA's Safe Drinking Water Information System (SDWIS). The dataset contains each public water system (PWS) ID, system type (community, non-community), number of people served, source water type (groundwater or surface water), region served by the system (city, county), violation type, the date of the beginning and the end of a monitoring period in which a PWS was in violation of a primary drinking water regulation, containment name, whether it's health based[1], whether it's a major violation[2],

compliance period begin and end date. EPA classifies water systems according to the number of people they serve: very small (≤500 people), small (501–3,300), medium (3,301–10,000), large (10,001–100,000), and very large (>100,000).

Demographic data County-level demographic data are collected from American Community Survey (ACS) 5-year estimate from 2010 to 2020. Key independent variables are the percentage of white, Black, and Hispanic, the percentage of the population below the poverty level, education level, the percentage of home ownership, median income, median home value, and percentage of the population in urban areas.

Table 3.1 provides a description of the water quality data. Each observation is a water quality violation sometime between 2010 and 2020. Non-compliance period

[1] These violations fall into three categories: 1) exceedances of the maximum contaminant levels (MCLs) which specify the highest allowable contaminant concentrations in drinking water, 2) exceedances of the maximum residual disinfectant levels (MRDLs), which specify the highest concentrations of disinfectants allowed in drinking water, and 3) treatment technique requirements, which specify certain processes intended to reduce the level of a contaminant. `https://echo.epa.gov/help/sdwa-faqs`

[2] There are two types of monitoring violations: 1) A major violation occurs when no tests were taken and/or no test results were submitted to the Department, 2) A minor violation occurs when some, but not all, of the required samples are collected and/or submitted. `https://www.epa.gov/compliance/stateterritorynavajo-nation-annual-public-water-systems-compliance-report`

measures the beginning to the end of a monitoring period in which a public water system was in violation of a primary drinking water regulation. The average non-compliance period is higher for larger water systems. For the smallest systems (<500 people served), the average non-compliance period is 290 days. For the largest system (>100,000 served), the average non-compliance period is 498 days. A large minority of small water systems are surface water systems, while the majority of larger systems are surface water systems. Across all sizes, most violations are major violations, but fewer than half of violations are health-based violations. Larger water systems serve more non-white counties. In the smallest water systems, 77% of the served county population is white, while in the larger systems, only 60% of the population is white. Poverty and education vary slightly across water system sizes. 13 to 16% of the population is below the poverty line, and 85 to 87% of the population has at least a high school diploma. Home-ownership is more common for individuals in smaller systems than for populations served by larger systems. Populations served by large water systems have higher median income, near $60,000, while populations served by smaller water systems have lower income, around $50,000 to $55,000. Median home value is substantially larger for larger water systems. Smaller systems are more likely to be located in rural counties, and almost all Very Large systems are in counties with more than 50 percent of the population living in urban areas. Figure 2.1 shows the average non-compliance period and percentage of non-white by county.

Table 2.2 reports correlations between non-compliance period and a number of key independent variables. Column (1) presents correlations for the full sample, columns (2) and (3) restrict the sample to major violations and health violations, and columns 4-8 restrict the sample to various population sizes. Non-compliance period is uncorrelated with the percentage of the white population. non-compliance period is negatively correlated with the black portion of the population, indicating that a county with a higher black population will have a shorter non-compliance period.

This relationship holds for most cuts of the data. The Hispanic portion of the county is positively correlated with non-compliance period. The portion of the population below the poverty line and the portion with a high school diploma are generally uncorrelated with non-compliance period, but the correlations are sometimes positive and sometimes negative, depending on the sample. Home-ownership is negatively correlated with non-compliance period. Median income and median house value are positively correlated in the full sample, but in the restricted samples, the correlation with non-compliance period is sometimes positive.

Flood event data For large regional floods, we use Presidential Disaster Declaration (PDD) events as data sources. The PDD system is a formalized process to request and receive federal assistance following large natural disasters. Public Assistance (PA) is FEMA's largest grant program, which provides funds to assist communities in responding to and recovering from major declared disasters or emergencies. PDD Public Assistance Funded Project Summaries[3] provides information on all approved federal disaster declaration requests, including data on the disaster type, disaster event start and end dates, affected counties, PA grant funding, and number of funded projects. We include disaster categories such as severe storms, flooding, and hurricanes, and we match the compliance period beginning date and PWS location to the disaster period and the location of PDD events to determine if the violation reports can be linked to the disasters. Figure 2.2 shows the average non-compliance period post-flooding and the number of violation reports within 1 month of flooding. Table 2.3 summarizes the violation reports linked to a PDD event and those that are not. We exclude the counties that did not experience any PDD event during the studied period. 2% of the violation reports in this sample occur within 1 month of a PDD event, and non-compliance period is shorter for the PWS that experienced PDD events recently. Public assistance's emergency work includes setting up temporary

[3]Additional information on the PDD data is available here: `https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2`

water treatment facilities and implementing measures to protect water sources from contamination, which explains the shorter contamination time and fewer violation reports categorized as Major.

2.3 Empirical Framework

We are interested in the relationship between demographics of population served by PWS and the non-compliance period, which can be specified in the following equation:

$$NoncompliancePeriod_{ict} = \beta_0 + \beta_1 SES_{ct} + \beta_2 X_{ict} + \alpha_t + \delta_c + \epsilon_{ict} \tag{2.1}$$

where $NoncompliancePeriod_{ict}$ is the non-compliance period of each violation report i in county c at year t. Independent variables include social-economic statuses of county c at year t, and number of controls X_{ict}: primary water source, violation type, whether it is health based, contaminant names and system sizes. α_t is year fixed effects, δ_c is county fixed effects which capture the effects of unobserved time-invariant local factors that affect each PWS, and ϵ_{ict} is the idiosyncratic error term that changes across time for each county. The coefficient we are interested in is β_1, which captures the differential clean-up time by the county's socio-economic status.

To understand the effect of large regional flooding on the differed contamination time intersects with the SES indicators, we modified the above equation into a difference-in-differences (DID) specification:

$$NoncompliancePeriod_{ict} = \gamma_0 + \gamma_1 PDD_{ct} + \gamma_2 PDD_{ct} * SES_{ct} + \gamma_3 SES_{ct} + \theta X_{ict} + \alpha_t + \delta_c + \epsilon_{ict} \tag{2.2}$$

where the outcome variable $NoncompliancePeriod_{ict}$ and the independent variable are same as before and $PDD_{ct} = 1$ if violation report begin within 30 days of a PDD in county c at year t. γ_1 represents the effect of large regional flooding on all PWS and γ_2 is the effect of the flooding on clean-up time associated with different levels of socio-economic status.

2.4 Empirical Results

2.4.1 Descriptive Results

Table 3.3 presents the result from equation 2.1. Our preferred model (6) includes both year and county-fixed effects. All estimates are statistically significant. A one-unit increase in the percentage of Black residents was associated with an 8.94-day increase in the non-compliance period. Both the percentage of Hispanic residents and poverty are associated with more extended non-compliance periods. By contrast, a one-unit increase in the percent of homeowner residents was associated with a 1.27-day decrease in the non-compliance period. However, education level and median home value are inversely associated with longer non-compliance periods. For the control variables, we see that surface water-sourced PWS have shorter non-compliance periods than groundwater-sourced PWS. The omitted group for system size is Very Small water systems. Compared to the smallest systems that serve less than 500 people, Small and Medium systems (serving 501 to 10,000 people) have shorter non-compliance periods (25.94 and 12.45 days). In contrast, Large and Very Large systems (serving a population of more than 10,001) are associated with 60.02 and 241 days longer in the non-compliance period. Although the coefficients for Large and Very Large seem big, those water systems are only 4.5% of all water systems, while Very Small accounts for 71% of all water systems. Previous EPA reports ((EPA), 2013) have noted the problem of varying non-compliance of different system sizes: larger systems have a greater capacity to maintain compliance than small systems and can return to compliance more quickly than small systems. Our results show the same story; if there are violations, small PWSs may need more technical capabilities to correct the underlying problems.

Table 2.5 looks at the non-compliance period by different types of violation reports. SDWIS violations generally fall into two categories: monitoring and reporting violations and health-based violations. Column (1) shows that the water systems in communities with higher rates of Black, Hispanic, and more population in poverty have longer non-compliance periods where they failed to follow established monitoring and reporting schedules. For communities with a one-unit increase in the percentage of home-ownership, the non-compliance period decreased by 3.67 days. Since most violation reports are categorized as Major Violation, the results from Table 3.3 are mainly driven by failing to monitor and report. Although these violations are not directly related to health, these violations can conceal underlying severe problems such as contamination. Without proper monitoring and reporting, it is impossible to determine whether the health-based standards have been met.

Column (2) looks at health-based violation reports that the water system either failed to comply with mandated treatment techniques or violated any Maximum Contaminant Levels (MCLs). The results show that a one-unit increase in the percent of Black residents is associated with a 1.54-day decrease in the non-compliance period, and one percentage increase in the median house value decreases the non-compliance period by 10.48 days. Other variables show statistically insignificant coefficients.

Column (3) focuses on the violation reports with water contaminated with arsenic, coliform, lead and copper, and nitrate. Those contaminants are a few of the most frequently found in violation of health-based standards in U.S. drinking water. Numerous studies indicate that exposure to chemicals in drinking water poses significant health risks to the population, particularly children, the elderly, and the immunocompromised (Council et al., 1999; Hopenhayn, 2006; Gruber et al., 2014; Ward et al., 2018; Brown and Margolis, 2012). We find that the percentage of the Black and Hispanic population in poverty is positively associated with more extended contamination time, and communities with higher median house value and median income are associated with shorter contamination time. The results expand the existing research on disparities in drinking water contaminants, particularly in relation to the length of exposure.

Columns (4) and (5) divide the sample into urban and rural counties. Many literature shows that the disparities in access to drinking water between rural and urban areas are noticeable (Strosnider, 2017; Nogueira et al., 2003). Smaller, rural water systems tend to have more violations than PWSs in larger and urban areas. Violations usually are related to poor-quality water sources and a lack of resources required to meet new regulations or maintain infrastructure (Patel et al., 2020). Delpla et al. (2015) suggests that small rural water systems serving areas with lower income, lower education levels, and higher unemployment rates are more likely to have lead levels in drinking water and less likely to have advanced water treatment to the systems. Our results show that in urban counties, a higher percentage of Hispanic residents and the population living below poverty are associated with a longer non-compliance period, whereas in rural counties, only the percentage of home-ownership is significant and negatively related to the non-compliance period.

2.4.2 Impact of Flooding on Non-Compliance Period

Table 2.6 presents the result from equation 2.2. The sample excludes the counties that did not experience any PDD event during the studied period. Column (1) shows that, for the violation reports that were reported within 1 month of a PDD event, the non-compliance period of the violation reports is negatively associated with the percentage of Black, Hispanic population and median home value; a higher percentage of the population below poverty is associated with longer non-Compliance Period after a PDD event. However, median income is related to a longer non-compliance period immediately after a PDD event. Column (2) limits the sample to the violation reports categorized as Major. The post-disaster non-compliance period is longer for populations with higher poverty rates and median income and is shorter for higher median house value areas. Column (3) only includes the violation report with arsenic, coliform, lead and copper contaminants, and nitrate. The coefficients are statistically insignificant except for the percentage of the Hispanic population. Columns (4) and (5) compare the post-disaster non-compliance period of urban and rural areas. In urban areas that have experienced a PDD event within 1 month, contamination time is shorter in areas with a lower percentage of Blacks and Hispanics but longer for a higher percentage of the population below poverty and counties with higher median income. For rural areas, the coefficients on race are statistically insignificant, but we find that higher median home value is associated with shorter post-disaster contamination time. This suggests that after a PDD event that affects the water quality, a higher poverty level results in longer non-compliance periods in both urban and rural counties, but race factors in only urban areas, while median house value has more effect in rural areas.

2.4.3 Public Assistance Program

In the aftermath of natural disasters, immediate recovery efforts focus on assisting individual victims. Communities often need to repair critical infrastructure systems that support essential services for the affected population. Among the most important of these are water infrastructure systems: drinking water treatment facilities and distribution systems that ensure the supply of safe and healthy potable water. To help address such emergencies, Congress has authorized programs over the years that can provide emergency assistance to repair and restore drinking water, wastewater, and related water infrastructure systems and facilities. Public Assistance (PA) is FEMA's largest grant program, providing funds to assist communities in responding to and recovering from major disasters or emergencies declared by the President. The program funds emergency assistance to save lives and protect property and permanent work to restore community infrastructure affected by a federally declared incident. Work that is eligible for PA funding is classified as either a) Emergency Work, which includes debris removal and emergency protective measures performed to eliminate or reduce immediate threats to public health and safety, including restoration of drinking water, or b) Permanent Work Category F (Public Utilities), which includes work to restore an eligible damaged facility to its pre-disaster design, repairing water treatment facilities and distribution systems.

Previous studies on FEMA's post-disaster relief have shown that the response to natural disasters is inequitable, and those disparities often lead to exacerbated wealth inequality (Howell and Elliott, 2018, 2019; Billings et al., 2022). Some suggest that socially vulnerable communities and communities with more Black, Hispanic, or Native American residents (Domingue and Emrich, 2019; Drakes et al., 2021) are likely to receive lower levels of assistance. This disparity could leave those more vulnerable communities worse off than before the disaster and unable to repair critical infrastructure, such as the water treatment and distribution facilities. If the amount of assistance funding and recovery projects are correlated with the social-economics status, the differences in post-disaster non-compliance period could be driven by the different level of disaster assistance funds. Table 2.7 shows that in all the studied counties that had experienced a PDD event, the amount of PA funding a county received is largely influenced by the median income and population size, with 1 percentage increase in median income associating with 1.5% increase in funding and 1% increase in population associating with 0.7% increase in funding. Column (2) shows that the number of projects in those counties are positively related to the percentage of population below poverty, median age, median income and population size, but negatively related to the percentage of home-ownership, median home value and if the county is classified as rural. We see that the amount of funding and projects are correlated with the demographics of the affected areas, the public assistance program efforts related to restore the safe drinking water post disaster could bias our main estimation.

Table 2.8 presents the estimation from equation 2.2 with two additional controls: the PA grant funding and number of funded projects. We see that the coefficients are similar to the main results in Table 2.6 with the effects of race, poverty slightly smaller on post-disaster non-compliance period and higher median home value results in a faster clean-up time. The amount of grant funding have mostly insignificant effect except for in rural areas, whereas the number of grant projects is negatively related the the non-compliance period.

2.4.4 Data Limitation

Our analyses of associations between the non-compliance period and socio-demographic factors are limited by the spatial resolution of available data at the county level. Public Water Systems (PWS) often serve communities at smaller geographic scales than the county, such as towns or neighborhoods, but there is no nationwide data available at the PWS level. If some PWS serve communities with socio-economic statuses that differ significantly from the aggregated county-level demographic characteristics, the results could be biased. This limitation can obscure disparities that might exist at more localized levels, where socio-economic conditions vary widely within a single county. Additionally, county-level analysis excludes drinking water systems that do not have information about the counties they serve, such as those catering to some Indigenous communities. It also overlooks populations that obtain their water from private wells or non-community water systems, further limiting the comprehensiveness of the study. Future research on the non-compliance period of drinking water could greatly benefit from more granular demographic data, which would allow for a more precise understanding of the relationship between water quality and socio-demographic factors. Despite these limitations, our analysis represents a valuable step toward assessing national drinking water disparities and emphasizes the need for improved data collection practices to better inform public health interventions and policy decisions.

2.5 Discussion and Conclusion

This paper provides compelling evidence of disparities in drinking water non-compliance periods at the national level, highlighting significant environmental justice concerns. Our results indicate that communities with a higher percentage of minority populations and higher percentages of populations living below the poverty line experience longer non-compliance periods, ranging from 3.9 to 8.9 days. Conversely, communities with higher median incomes and home-ownership rates tend to have shorter non-compliance periods. When examining violations related to failure to monitor and report, we find that both race and poverty are strongly associated with extended non-compliance periods. For violations involving hazardous chemicals such as arsenic, coliform, lead, copper, and nitrate exceeding the Maximum Contaminant Level (MCL), communities with higher median home values and median incomes have significantly shorter periods of contamination. Additionally, in rural areas, the rate of home-ownership appears to be a more significant factor than race and poverty indicators in determining non-compliance periods. These findings highlight the critical need to address current disparities to improve water quality and mitigate the future impact of contaminant violations, which disproportionately affect disadvantaged communities. Proactively addressing these issues through targeted policies and interventions can ensure more equitable access to safe drinking water for all communities.

This paper also assesses the environmental justice implications of disasters on water quality disparities, emphasizing the importance of understanding the relationship between post-disaster drinking water contamination and social inequality. This understanding is crucial for guiding policy decisions that prioritize assistance to communities disproportionately impacted by both the disaster and poor water quality. Our estimations reveal that in counties recently affected by a PDD event, higher poverty rates are associated with longer non-compliance periods. However, our analysis did not find a statistically significant relationship between socio-economic status and the duration of contamination by key pollutants post-disaster. Despite this, understanding the duration of exposure to contaminants and non-compliance in monitoring and reporting, especially among vulnerable populations, is essential for developing effective post-disaster recovery strategies. These strategies must focus on swiftly restoring safe drinking water in the most affected communities. By addressing these disparities, we can ensure that recovery efforts are equitable and that the health and safety of all community members are protected in the aftermath of disasters. This research underlines the need for targeted interventions and resources to support vulnerable populations, ensuring they receive the necessary aid to recover from both immediate and long-term impacts on water quality.

Overall, the results suggest critical concerns exist over disadvantaged communities bearing a disproportionate burden of water quality impairment and longer non-compliance periods. An in-depth understanding of the most affected population groups is critical to water resource management and planning decisions. While federal and state post-disaster assistance might be available to struggling communities, targeting assistance can be challenging because of the limited information on populations disproportionately burdened by water quality disparities. Our findings address the knowledge gap for the optimal distribution of state water resources to support water quality improvement, especially within under-performing communities. The findings discussed in this paper can advance the understanding of vulnerable communities and the call for environmental justice associated with drinking water quality. Reducing water quality violations period can lead to improved health outcomes and less disparity in water service. Post-disaster scenarios often reveal and exacerbate existing inequalities and regulatory measures that do not consider these disparities can unintentionally leave the most vulnerable populations without essential resources. By assessing how different communities are affected by regulations on water quality, infrastructure repair, and resource allocation, policymakers should identify areas where marginalized groups might be disproportionately impacted. Prioritizing the needs of these vulnerable groups in disaster recovery plans ensures that safe drinking water remains accessible to everyone, especially those historically underserved or at higher risk. This approach not only promotes health equity but also reduces long-term social and economic disparities, enhancing overall community resilience. By ensuring that all individuals, regardless of socio-economic status, have reliable access to clean water, policymakers can foster more inclusive and sustainable communities that are better prepared to withstand future environmental challenges.

2.6 Tables

Table 2.1: Summary statistics by water system sizes 2010-2020

	Very Small ($\leq$500)		Small (501–3,300)		Medium (3,301–10,000)		Large (10,001–100,000)		Very Large (>100,000)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Violation Report										
Non-Compliance Period	290.429	435.721	271.35	419.768	295.402	417.231	369.283	461.431	497.529	472.564
Surface water =1	0.045	0.207	0.145	0.352	0.241	0.428	0.272	0.445	0.732	0.443
Major vioaltion =1	0.64	0.48	0.704	0.457	0.689	0.463	0.684	0.465	0.784	0.411
Health based =1	0.109	0.311	0.146	0.353	0.155	0.362	0.134	0.341	0.073	0.26
Obeservations	1,989,747		511,876		168,423		142,371		16,989	
County-level demographic										
% White	76.71	18.77	72.74	20.84	72.98	18.58	69.30	18.84	60.29	17.92
% Black	5.59	8.56	12.64	18.30	11.96	15.39	8.31	10.95	9.73	9.50
% Hispanic	12.33	15.11	9.68	13.98	10.09	13.69	15.58	16.85	22.83	17.88
% Below Poverty	13.47	4.80	16.08	6.72	15.75	5.95	13.43	5.47	13.05	3.88
% with HS diploma	87.23	5.86	84.44	7.26	84.52	6.78	87.26	5.97	87.33	4.92
% Homeowner	71.23	7.36	72.14	7.37	71.45	6.96	69.32	7.24	64.23	7.06
% Male	49.80	1.55	49.66	1.77	49.53	1.84	49.36	1.20	49.16	0.89
Median age	40.41	5.24	39.69	5.00	39.15	4.53	38.22	4.96	37.02	3.67
Median income	54,116.60	14,193.87	48,306.03	14,913.92	49,290.34	14,982.90	57,798.05	16,507.31	59,569.77	13,981.14
Median home value	180,241.44	86,955.51	146,582.89	84,604.90	153,737.22	93,025.38	208,008.89	103,705.92	227,650.02	100,048.39
Population	291,854.86	695,269.65	206,663.23	626,501.09	249,164.49	621,631.45	508,343.79	896,741.71	1,608,512.20	1,728,878.86
Urban = 1	0.571	0.495	0.429	0.495	0.525	0.499	0.835	0.372	0.997	0.054
Obeservations	628,013		161,466		53,579		37,305		2,821	

Notes. Table provides the mean attributes of PWS by system sizes and the population characteristics it serves. County-level demographic data are from American Community Survey (ACS) 5-year estimate from 2010 to 2020. Counties with more than 50 percent of the population living in urban areas are classified urban.

Table 2.2: Correlations between Non-Compliance Period and key independent variables

	All	Major Violations	Health Violations	Very Small	Small	Medium	Large	Very Large
% White	-0.001	0.0017	-0.0996	-0.0139	0.0635	0.0662	-0.0876	0.0374
% Black	-0.0482	-0.0813	0.0015	0.0088	-0.1117	-0.1808	-0.1345	-0.2273
% Hispanic	0.0376	0.0655	0.105	0.0168	0.0381	0.1052	0.1614	0.0781
% Below Poverty	-0.0314	-0.0375	0.1014	0.0093	-0.1141	-0.1642	0.0572	-0.08
% with HS Diploma	0.0298	0.011	-0.1156	-0.0088	0.1145	0.1188	0.0167	0.0219
% Homeowner	-0.0273	-0.0465	-0.052	-0.0137	-0.0314	-0.0906	-0.0465	-0.025
Median Income	0.0246	0.0088	-0.0685	-0.0071	0.1117	0.1273	-0.0827	-0.0187
Median House Value	0.024	0.0284	-0.0936	-0.0156	0.1092	0.1919	-0.0363	0.1018
Observations	883,184	618,872	117,578	628,013	161,466	52,579	37,305	2,821

Notes. Table provides the correlations between non-compliance period and the demographic characteristics the PWS serves.

Table 2.3: Summary statistics by flooding incident 2010-2020

	No flooding within 1 month of violation period		Flooding within 1 month of violation period	
	Mean	SD	Mean	SD
Violation Report				
Non-Compliance Period	233.1	369.7	89.04	92.52
Surface water =1	0.0675	0.251	0.0656	0.248
Major vioaltion =1	0.565	0.496	0.428	0.495
Health based =1	0.125	0.330	0.179	0.383
PWS population served	2,459	22,099	2,045	15,125
County-level demographic				
% White	75.39	19.38	79.77	17.40
% Black	6.711	10.44	5.821	8.495
% Hispanic	12.23	14.92	9.000	11.96
% Below Poverty	13.83	5.262	12.93	4.826
% with HS diploma	86.86	5.770	87.50	5.071
% Homeowner	70.92	7.269	71.78	7.062
% Male	49.65	1.420	49.63	1.449
Median age	40.32	5.168	41.00	4.480
Median income	54,040	14,908	55,035	15,270
Median home value	180,730	89,916	192,045	97,700
Population	319,390	743,374	275,082	621,884
Urban = 1	0.570	0.495	0.527	0.499
Obeservations	614,623		13,160	

Notes. The table provides the mean attributes of violation reports and the population characteristics the PWS serves. Sample excludes the counties that did not experienced any PDD event during the studied period. County-level demographic data are from the American Community Survey (ACS) 5-year estimate from 2010 to 2020. PDD flooding includes categories such as flood, hurricane, and severe storm.

Table 2.4: Descriptive Results for Non-Compliance Period

	(1)	(2)	(3)	(4)	(5)	(6)
% Black	-1.507***	-1.601***	0.908***	1.358	0.809***	8.938***
	(0.0436)	(0.0429)	(0.0589)	(1.051)	(0.0568)	(1.029)
% Hispanic	0.626***	0.857***	0.00376	-2.704***	0.363***	6.346***
	(0.0395)	(0.0382)	(0.0543)	(0.582)	(0.0525)	(0.619)
% Below Poverty	-2.225***	1.732***	-6.327***	1.083***	-0.449**	3.906***
	(0.175)	(0.182)	(0.180)	(0.402)	(0.188)	(0.421)
% Homeowner	-0.927***	-0.779***	-0.853***	1.515***	-0.502***	-1.268***
	(0.0687)	(0.0661)	(0.0745)	(0.284)	(0.0719)	(0.293)
% with HS Diploma	2.254***	3.459***	-1.091***	-2.331***	1.035***	4.853***
	(0.111)	(0.110)	(0.131)	(0.397)	(0.129)	(0.428)
Median House Value	18.95***	-13.26***	15.62***	-7.906	-14.87***	28.99***
	(1.526)	(1.529)	(2.034)	(6.238)	(2.031)	(6.305)
Median Income	-90.92***	33.62***	-123.7***	-364.9***	37.64***	-81.26***
	(3.792)	(4.083)	(4.356)	(10.03)	(4.706)	(13.87)
Surface water	-76.83***	-62.97***	-87.68***	-64.03***	-73.73***	-60.19***
	(1.655)	(1.595)	(1.641)	(1.773)	(1.581)	(1.718)
System Size: Small	-37.02***	-49.72***	-12.68***	-13.57***	-24.89***	-25.94***
	(1.142)	(1.101)	(1.132)	(1.165)	(1.092)	(1.131)
System Size: Medium	-32.92***	-47.04***	-1.627	0.147	-16.40***	-12.45***
	(1.783)	(1.716)	(1.771)	(1.834)	(1.706)	(1.778)
System Size: Large	59.07***	39.33***	78.33***	74.81***	57.03***	60.02***
	(2.092)	(2.015)	(2.102)	(2.215)	(2.027)	(2.147)
System Size: Very Large	239.2***	217.7***	253.2***	249.3***	226.4***	241.0***
	(7.147)	(6.872)	(6.951)	(7.026)	(6.693)	(6.807)
Constant	911.1***	-223.6***	1,631***	4,389***	-45.43	249.7*
	(37.69)	(40.49)	(39.61)	(86.18)	(43.28)	(150.7)
Observations	882,924	882,924	882,924	882,888	882,924	882,888
R-squared	0.048	0.120	0.109	0.200	0.174	0.249
Year FE		Yes			Yes	Yes
County FE				Yes		Yes
State FE			Yes		Yes	

Notes. Dependent variable is noncompliance period (days). Standard errors are stated in parentheses and are clustered at PWS level. Median house value and median income are transformed with natural logs. *** p<0.01, ** p<0.05, * p<0.1

Table 2.5: Descriptive Results for Non-Compliance Period by Violation Type and County Characteristic

	(1) Major	(2) Health-based	(3) Chemicals	(4) Urban	(5) Rural
% Black	23.43***	-1.543**	2.590***	11.90	6.885
	(8.703)	(0.742)	(0.874)	(11.29)	(4.824)
% Hispanic	5.628	0.197	3.832***	10.62***	1.880
	(4.466)	(0.391)	(0.524)	(3.840)	(3.350)
% Below Poverty	8.297***	0.0933	1.004***	8.616**	1.140
	(2.769)	(0.348)	(0.318)	(3.544)	(2.012)
% Homeowner	-3.672*	0.0985	-0.636**	3.064	-3.688***
	(2.079)	(0.183)	(0.264)	(2.995)	(1.408)
% with HS Diploma	10.22***	-0.650*	0.522	9.383**	1.800
	(2.858)	(0.339)	(0.354)	(3.913)	(1.907)
Median House Value	1.292	-10.48***	-22.90***	-12.38	78.10*
	(50.18)	(3.765)	(5.308)	(45.50)	(44.29)
Median Income	2.831	-1.702	-31.43***	-146.9	-39.21
	(93.53)	(8.058)	(11.13)	(125.3)	(65.10)
Observations	618,599	117,387	301,905	491,176	391,712
R-squared	0.316	0.260	0.107	0.237	0.275

Notes. Dependent variable is noncompliance period (days). All the regressions include county fixed effect, year fixed effect, PWS system size controls, water source controls and PWS type controls. Standard errors are stated in parentheses and are clustered at PWS level. Median house value and median income are transformed with natural logs. Chemicals refers to the contaminants of arsenic, coliform, lead and copper, and nitrate. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.6: Impact of Flooding on Non-Compliance Period

	(1) All	(2) Major Violation	(3) Chemicals	(4) Urban	(5) Rural
PDD	-2,809***	-4,171***	-296.5	-3,644***	-1,288**
	(288.0)	(426.7)	(230.4)	(849.3)	(547.7)
PDD*% Black	-1.205***	-0.534	-0.424	-3.543***	0.656
	(0.407)	(0.626)	(0.333)	(1.124)	(0.774)
PDD*% Hispanic	-1.554***	-1.140*	-0.907***	-1.779**	-0.298
	(0.385)	(0.664)	(0.271)	(0.813)	(0.787)
PDD*% Below Poverty	14.41***	22.16***	1.809	17.94***	7.436***
	(1.456)	(2.218)	(1.137)	(4.114)	(2.317)
PDD*% Homeowner	-0.166	-0.343	0.228	-0.891	-0.824
	(0.558)	(0.917)	(0.411)	(1.688)	(1.299)
PDD*% with HS Diploma	3.011***	9.270***	0.297	2.775	4.571**
	(1.009)	(1.747)	(0.723)	(2.321)	(1.827)
PDD*Median House Value	-38.87***	-116.7***	5.418	-23.38	-45.86**
	(10.87)	(17.63)	(8.461)	(34.28)	(21.54)
PDD*Median Income	249.9***	397.1***	12.23	311.8***	118.0**
	(27.91)	(40.11)	(22.69)	(86.90)	(57.88)
% Black	4.296***	18.60***	9.307***	16.93	-4.552
	(1.251)	(1.712)	(1.180)	(13.67)	(5.578)
% Hispanic	6.034***	8.232***	-12.43***	15.86***	-3.502
	(0.799)	(1.116)	(0.747)	(4.539)	(5.391)
% Below Poverty	2.741***	4.922***	-0.339	7.772*	-0.286
	(0.541)	(0.749)	(0.500)	(4.698)	(2.502)
% Homeowner	1.672***	1.363***	-0.302	7.634**	-2.360
	(0.373)	(0.512)	(0.352)	(3.637)	(1.740)
% with HS Diploma	6.645***	11.83***	-2.123***	6.634	4.003
	(0.528)	(0.720)	(0.499)	(4.799)	(2.471)
Median House Value	-16.76**	-50.21***	-95.79***	18.83	-83.10
	(7.446)	(10.35)	(6.719)	(47.43)	(53.48)
Median Income	-91.79***	2.539	-19.77	-189.2	-29.30
	(17.49)	(24.11)	(16.10)	(151.1)	(83.30)
Observations	627,780	429,913	176,338	360,546	267,234
R-squared	0.231	0.303	0.170	0.228	0.242

Notes. Dependent variable is noncompliance period (days). Sample excludes the counties that did not experienced any PDD event during the studied period. All the regressions include the county fixed effect, year fixed effect, PWS system size controls, water source controls and PWS type controls. Standard errors are stated in parentheses and are clustered at PWS level. Median house value and median income are transformed with natural logs. Chemicals refers to the contaminants of arsenic, coliform, lead and copper, and nitrate. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.7: Robustness: Public Assistance program on County Characteristics

	(1) Log(Funding)	(2) # of Projects
% Black	-0.00854***	-0.0930
	(0.00306)	(0.134)
% Hispanic	0.0104***	0.355**
	(0.00333)	(0.146)
% Below Poverty	0.0405***	1.915***
	(0.0142)	(0.625)
% with HS Diploma	0.0342***	0.372
	(0.00900)	(0.396)
% Homeowner	-0.0247***	-0.839***
	(0.00663)	(0.291)
% Male	0.0382**	0.833
	(0.0176)	(0.774)
Median Age	0.0542***	3.179***
	(0.00948)	(0.417)
Median House Value	-0.726***	-30.39***
	(0.125)	(5.505)
Median Income	1.552***	91.84***
	(0.386)	(16.97)
Population	0.697***	22.82***
	(0.0369)	(1.622)
Urban=1	-0.0257	-9.616**
	(0.0908)	(3.992)
Constant	-8.475**	-1,009***
	(3.826)	(168.1)
Observations	2,598	2,598
R-squared	0.304	0.202

Notes. Dependent variables are log transformed PA grant funding in dollar amount and number of funded projects. County-level demographic data are from the American Community Survey (ACS) 5-year estimate from 2010 to 2020. Median house value and median income are transformed with natural logs. All the regressions include year fixed effect. Standard errors are stated in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.8: Impact of Flooding on Non-Compliance Period with PA controls

	(1) All	(2) Major Violation	(3) Chemicals	(4) Urban	(5) Rural
PDD	-2,211***	-2,940***	-269.9**	-2,835***	-1,080*
	(503.6)	(969.1)	(137.7)	(811.6)	(556.4)
PDD*% Black	-0.924	-0.194	-0.413**	-2.369**	0.596
	(0.647)	(1.220)	(0.201)	(1.053)	(0.764)
PDD*% Hispanic	-1.436***	-0.816	-0.940***	-1.116	-0.446
	(0.498)	(1.157)	(0.186)	(0.805)	(0.790)
PDD*% Below Poverty	10.51***	15.18***	1.454*	12.11***	6.027**
	(2.109)	(4.351)	(0.752)	(4.011)	(2.370)
PDD*% Homeowner	-1.394	-2.415	-0.0407	-2.286	-1.375
	(1.002)	(2.166)	(0.260)	(1.593)	(1.311)
PDD*% with HS Diploma	2.326	7.384**	0.251	2.186	4.006**
	(1.439)	(3.509)	(0.420)	(2.300)	(1.857)
PDD*Median House Value	-39.39**	-131.3***	6.987	-42.01	-37.63*
	(19.46)	(42.99)	(4.633)	(32.41)	(21.90)
PDD*Median Income	214.8***	331.2***	12.42	286.2***	92.07
	(52.67)	(98.89)	(13.96)	(85.89)	(58.75)
% Black	4.172	18.39*	9.303***	16.75	-4.600
	(6.437)	(10.42)	(1.743)	(13.67)	(5.577)
% Hispanic	5.927*	8.090	-12.43***	15.44***	-3.423
	(3.324)	(5.252)	(1.172)	(4.535)	(5.395)
% Below Poverty	2.790	4.963	-0.325	7.770*	-0.259
	(2.307)	(3.488)	(0.670)	(4.698)	(2.502)
% Homeowner	1.640	1.334	-0.307	7.570**	-2.353
	(1.659)	(2.563)	(0.526)	(3.637)	(1.740)
% with HS Diploma	6.656***	11.87***	-2.114***	6.639	4.022
	(2.226)	(3.460)	(0.683)	(4.801)	(2.471)
Median House Value	-17.03	-51.05	-96.06***	19.23	-83.33
	(35.08)	(54.92)	(9.404)	(47.43)	(53.50)
Median Income	-91.26	2.621	-19.72	-190.5	-29.19
	(74.63)	(116.2)	(22.27)	(151.0)	(83.29)
Log(Funding)	0.526	6.686	-1.078	-4.798	7.375**
	(2.884)	(5.241)	(0.894)	(4.295)	(3.395)
# of Projects	-0.293***	-0.409***	-0.0729***	-0.329***	-0.192***
	(0.0340)	(0.0697)	(0.0123)	(0.0447)	(0.0472)
Observations	627,780	429,913	176,338	360,546	267,234
R-squared	0.231	0.303	0.170	0.229	0.242

Notes. Dependent variable is noncompliance period (days). All the regressions include the county fixed effect, year fixed effect, PWS system size controls, water source controls and PWS type controls. Standard errors are stated in parentheses and are clustered at PWS level. Median house value, median income and PA fundings are transformed with natural logs. Chemicals refers to the contaminants of arsenic, coliform, lead and copper, and nitrate. *** p<0.01, ** p<0.05, * p<0.1

2.7 Figures

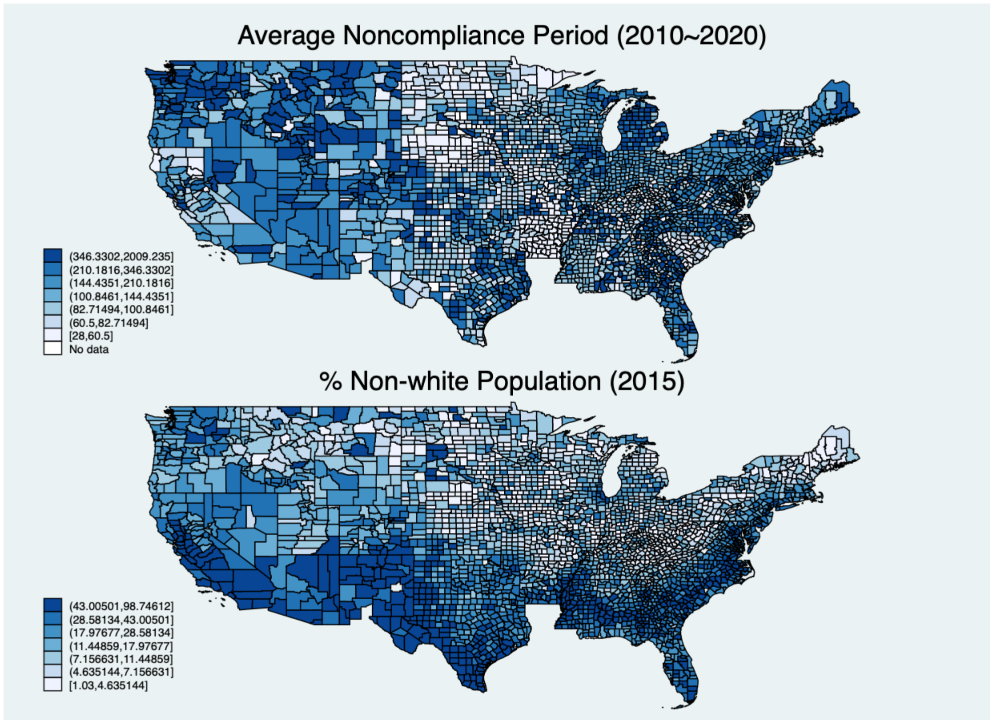

Figure 2.1: Average Non-compliance Period and % Non-white Population by county

Notes. Top graph provides the mean non-compliance period of each violation reports in each county. Bottom graph shows the county-level demographic data of percentage non-white population is from American Community Survey (ACS) 5-year estimate from 2010 to 2020.

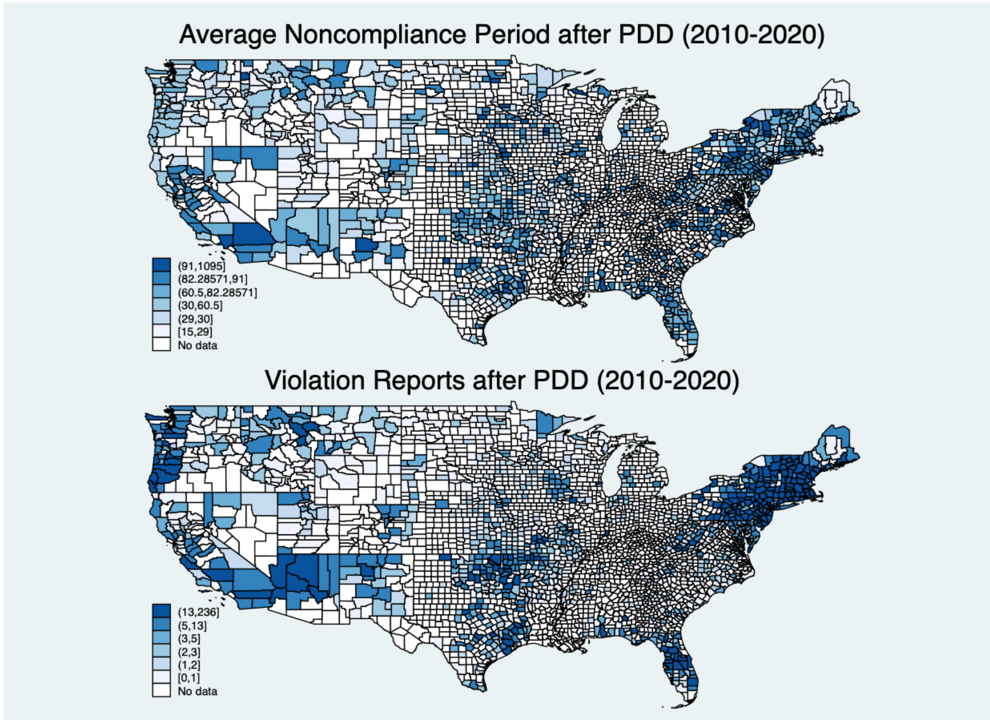

Figure 2.2: Average Non-compliance Period and Violation Report Counts after PDD by county

Notes. Top graph provides the mean non-compliance period of each violation reports in flooded counties. Bottom graph shows the number of violation reports in flooded counties.

Chapter 3 Impact of Flooding and Free School Meals on Student Performance

3.1 Introduction

Flood events are the most common and costly natural disasters in the U.S., affecting millions of individuals each year. According to the National Centers for Environmental Information, the U.S. has witnessed over $67.8 billion in flood damages since 2010 (Smith, 2020). In Kentucky, flooding is the state's most frequent and costly natural disaster[1]. At the end of July 2022, several counties in Eastern Kentucky were hit by severe flash floods resulting from a week-long heavy rain. The "1-in-1000 year"[2] flood event claimed more than 30 lives and destroyed hundreds of homes, schools, and roads in the area. Multiple counties are forced to delay the school start dates and many students will have to switch school districts.

For families in regions affected by natural disasters, the impacts extend beyond mere disruption of daily routines; they encompass a cascade of challenges that profoundly affect children's lives. Disasters often lead to missed school days and hindered academic progress, depriving children of crucial learning opportunities and social interactions. Moreover, they heighten exposure to various stressors like illness, family turmoil, domestic violence, and substance abuse, all of which can severely impact children's physical health and emotional well-being in both the short and long terms. Many researchers have studied the effect of different types of natural disasters on academic performance. These negative effects of disaster may have severe consequences

[1] Estimates are based on the Storm Events Database from NOAA/National Centers for Environmental Information (NCEI). Digital data are available at `http://www.ncdc.noaa.gov/stormevents/ftp.jsp`

[2] According to the United States Geological Survey (USGS), the term "1,000-year flood" means that a flood of that magnitude (or greater) has a 1 in 1,000 chance of occurring in any given year. In terms of probability, the 1,000-year flood has a 0.1% chance of happening in any given year.

for children's physical health and emotional and intellectual well-being in both the short and long terms (Mudavanhu, 2014; Gibbs et al., 2019; Thamtanajit, 2020). Thus, low school enrollment, low class attendance and high student drop-outs are recurring problems in child education among poor households especially in areas of high food insecurity.

Food and nutrition assistance programs play a pivotal role in safeguarding households' investments in education by alleviating financial burdens associated with schooling. These programs not only provide essential nutrition but also encourage parents to enroll their children in school and ensure regular attendance throughout their academic years. Literature has shown that food insecurity have a negative impact on students' academic performance and mental wellness (Alaimo et al., 2001; Ashiabi, 2005). Studies focused on food assistance program such as National School Lunch Program (NSLP) and School Breakfast Program (SBP), found positive effects on education and achievement (Kleinman et al., 2002; Imberman and Kugler, 2014; Schwartz and Rothbart, 2020; Gordanier et al., 2020; Cohen et al., 2021), also improves health outcomes (Peterson, 2014; Rothbart et al., 2023). In the aftermath of natural disasters, school districts become crucial focal points for providing nutrition support to students through the programs. These initiatives are pivotal in mitigating the immediate effects of food insecurity among children, ensuring they receive balanced meals despite the upheaval caused by disasters. The school meal programs not only enhance nutritional intake but also reduce food insecurity, thereby promoting better health outcomes and supporting educational achievement among vulnerable student populations.

This study aims to contribute to the existing literature on the impact of natural disasters on students' education by examining the effects of severe flooding on student achievement. Specifically focused on the Kentucky public school system, the research investigates how large regional flood events influence academic performance among students. In addition to assessing the direct impact of flooding, the study also evaluates the role of food security programs, such as the NSLP and SBP, in supporting students' academic resilience post-disaster. The findings indicate that the programs plays a crucial role in mitigating some of the adverse effects of large regional floods on students who participate in the free or reduced-price meal programs. This suggests that food security programs can help disadvantaged students maintain their academic performance despite the challenges posed by flooding events.

Furthermore, this research highlights the potential of school nutrition programs as a critical component of emergency preparedness and response strategies. While these programs have historically responded to emergencies, more research is still needed to examine their emergency preparedness. Understanding how food assistance programs impact education outcomes in times of crisis can provide valuable insights into developing and enhancing programs that effectively mitigate the disruptive effects of natural disasters on education. By exemplifying these mechanisms, the study aims to inform future educators, policymakers, and emergency responders on best practices for supporting student resilience and academic continuity in the aftermath of flood events. This knowledge can guide the development of proactive measures and policies to ensure that vulnerable student populations receive the necessary support to thrive academically amidst disaster recovery efforts.

The paper proceeds as follows. The first section introduces the background and related literature. The second section shows details on the data and variables of interest. The third section presents the research design and identification strategies. The fourth section presents the results. Finally, the fifth section concludes and discusses the limitations of the study.

3.2 National School Meals Program

The national school meals programs, National School Lunch Program (NSLP) and the School Breakfast Program (SBP), are federally-assisted meal programs in public and nonprofit private schools. Originating from the National School Lunch Act of 1946, signed by President Harry Truman, to "safeguard the health and well-being of the Nation's children and to encourage the domestic consumption of nutritious agricultural commodities and other foods." The program aims to provide low-cost or free lunches to eligible students. The NSLP is the second largest food and nutrition assistance program in the United States, following only SNAP (Supplemental Nutrition Assistance Program), and it has been vital in supporting the nutritional needs of its school-age population. About 35.9 million children received a school lunch in 2021 [3]and 43% of US public school students attend schools where a majority of students are eligible for free/reduced-price lunch [4]. Although schools are not required to offer NSLP and SBP meals, 94 percent of schools, both public and private. Of this group, the percentage of students who attended high-poverty schools was highest among Hispanic students (38%), Black students (37%), and American Indian/Alaska Native students (30%), and the percentage of students receiving free or reduced-price lunch often used to measure how many students live in poverty.

Eligibility for the NSLP and SBP is primarily determined based on household income. Students from households with incomes at or below 130% of the federal poverty income threshold qualify for free lunch under the program. Additionally, students who are enrolled in other federal assistance programs such as Temporary Aid to Needy Families (TANF), the Supplemental Nutrition Assistance Program

[3]Survey of Income and Program Participation (SIPP) provides national data about the receipt of free/reduced-price school meals. `https://www.census.gov/data/tables/2021/demo/public-assistance/sipp-receipts.html`

[4]U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Public Elementary/Secondary School Universe Survey," 2021–22 `https://nces.ed.gov/ccd/pubschuniv.asp`

(SNAP, formerly known as the Food Stamp Program), and the Food Distribution Program on Indian Reservations (FDPIR) are automatically certified for free lunch through administrative records. For students from households earning between 130% and 185% of the poverty threshold, they qualify for reduced-price lunch, where the meal cost is subsidized but still requires a minimal contribution from the family. At the beginning of each school year, schools distribute school meal applications to parents or guardians. These applications are submitted to the school food authority (SFA), where families self-report their household's total income for the most recent full month, the household size, and whether any household members participate in federal food and nutrition assistance programs. Based on this information, the SFA determines the eligibility of each student for free or reduced-price meals. This process ensures that eligible students receive access to nutritious meals that support their health and well-being throughout the school year, regardless of their financial circumstances. The national school meals programs not only aims to alleviate hunger but also plays a crucial role in promoting academic achievement and overall student success by ensuring that children have access to balanced meals during the school day.

3.3 Data

Education data Education assessment data collected from 2013 to 2018 by the Kentucky Department of Education (KDE) provides a comprehensive overview of student performance across various subjects and grade levels in Kentucky schools. [5] Since 2012, Kentucky has administered the Kentucky Performance Rating for Educational Progress (K-PREP) tests annually. These tests cover reading, mathematics, science, social studies, and writing, spanning elementary through high school grades. Held in May each year, the tests are scheduled by counties, and student performance

[5] School Report Card data downloaded from `https://openhouse.education.ky.gov/Home/SRCData`

is categorized into four achievement levels: novice, apprentice, proficient, and distinguished. At the lowest performance level, novice, students demonstrate minimal understanding of the content standards for their grade level. This level of performance indicates significant academic challenges and may signal a need for additional support and intervention to improve understanding. The second performance level, apprentice, demonstrates partial proficiency and may require further instruction and practice to achieve proficiency. Students at the third level, proficient, demonstrate a solid understanding of the content standards and can apply their knowledge effectively. They meet grade-level expectations and demonstrate competency in the subject area being tested. The distinguished level represents the highest level of achievement. Students at this level not only meet but exceed grade-level expectations. They demonstrate advanced understanding, critical thinking skills, and the ability to apply knowledge in complex ways. The school-level assessment data includes detailed information such as school names, school districts, grades tested, subjects assessed, and percentages of students falling into each achievement category. Data is further disaggregated by factors including gender, race, and eligibility for free/reduced-price meals. The dataset includes enrollment numbers and test participation rates for each school. I also included the number of full-time equivalent teachers at each school. This allows for the calculation of student-to-teacher ratios, an important indicator of classroom resource allocation and educational support.

Table 3.1 provides a description of the education assessment data. Each observation is a school-level test performance by grades and by subjects. 17.58% of all students in elementary schools are at novice level, 31.01% are at apprentice level, and 51.41% are at proficient or distinguished level. When focusing on students receiving free or reduced-price meals, a subset often used as an indicator of economic disadvantage, 21.21% are at novice level, 34.42% are at apprentice level, and 44.36% are at proficient or distinguished level. These patterns persist into middle and high school levels, where a higher percentage of students eligible for free/reduced-price meals perform at the novice and apprentice levels compared to their peers. This disparity shows the impact of economic disadvantage on educational outcomes, as students from economically disadvantaged backgrounds often face additional barriers that can affect their academic performance.

Flood event data For large regional floods, I use Presidential Disaster Declaration (PDD) Floods events and NFIP redacted claims as data sources. The PDD system is a formalized process to request and receive federal assistance following large natural disasters. PDD Summaries from FEMA provides information on all approved federal disaster declaration requests, including data on the disaster type, disaster event start and end dates, and affected counties.[6] NFIP redacted claims data [7] provides claim transactions on property type, date of loss, flood zone, and the amount paid on claims. I match the date of loss and the census tract of each claim to the incident period of PDD floods to determine if the flood damage is caused by a large regional flood. Since PDD floods are determined at the county level, not all communities within a county are affected by the flood. I construct a variable to identify which school districts in PDD counties are "hit" by each flood. I consider a school district to be hit if there are at least \$100,000 in building claims linked to the PDD floods within the school district. With the flood events and testing period dates, I match if the school year/semester is affected by any large regional floods.

Table 3.2 provides the education assessment data by flooded and non-flooded schools. The levels of academic performance of flooded schools and that of control schools are similar for both all students and students on free/reduced price meal.

[6]Additional information on the PDD data is available here: `https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2`

[7]Additional information on the NFIP redacted claims data are available here: `https://www.fema.gov/openfema-data-page/fima-nfip-redacted-claims-v1`

3.4 Empirical Framework

I am interested in the impact of large regional flooding on students test performance, which can be specified in the following equation:

$$Level_{igst} = \beta_0 + \beta_1 PDD_{ct} + \alpha_t + \delta_{igs} + \epsilon_{igst} \quad (3.1)$$

where $Level_{ict}$ is the percentage of students at 4 different achievement categories: novice, apprentice, proficient, and distinguished in school i at grade g at year t. PDD_{it} is a dummy variable equals to 1 if a PDD hit the school district within the school year (August to May) or within the school semester (January to May). α_t is year fixed effects, δ_{igs} is school-grade fixed effects which capture the effects of unobserved time-invariant cohort factors, and ϵ_{ict} is the idiosyncratic error term that changes across time for each cohort. The estimated β_1 captures the effect of a large regional flood event on the education outcome of the students. If a severe flood adversely affects the examination scores, the percentage of students scoring at novice and apprentice in schools that were affected by the flood should be higher than those of the schools that were not affected by the flood.

3.5 Empirical Results

3.5.1 Main Estimates

Table 3.3 shows the results from equation 3.1. For all students, the schools that experienced a large flooding event within the school year has the percentage of students scoring at novice level increases by 0.34 percentage point, which is 0.03 standard deviation increase in the percentage of students scoring at novice level. In contrast, the results for students with free/reduced price meal in those schools show insignificant change in test performance at all four levels. This suggest that, although students with free/reduced price meals are considered more disadvantaged, the program mitigates some of the adverse effects from the large regional floods on those students.

Free/reduced price meal can help disadvantaged students maintain their academic performance in the face of natural disasters like flooding, while other students, lacking this level of structured support, might experience a decline in their test performance. Bottom panel shows that the negative effects on all students from the PDD events hit the school within a semester of the test are larger, but the effects on free/reduced price meal students remain statistically insignificant.

Even though in equation 3.1 I included school-grade-subject fixed effects to control for unobserved time-invariant characteristics of the cohorts, the challenge relevant to my estimation is to account for omitted variables that are time-varying. To account for potential bias, I estimate additional specifications that include student-to-teacher ratio as the additional time-varying school-specific control in the robustness check in Table 3.4. In all cases, the results are similar to my baseline results and robust to the inclusion of the student-teacher ratio as an additional explanatory variable.

3.5.2 Heterogeneous Impacts of Flooding on students

Thamtanajit (2020) has shown that disasters often have a disproportional effect on students in elementary and middle school comparing to high school . Figure 3.1 represents the effect of flooding events student achievement across different school levels. In elementary schools affected by a flooding event, there is statistically significant increase in the percentage of students eligible for free/reduced-price meals who achieve at the distinguished level, signaling an unexpected positive outcome amidst adversity. Concurrently, there is a significant decrease in the percentage of these students scoring at the apprentice level, suggesting a shift towards higher academic achievement among economically disadvantaged students following the disaster. Conversely, middle school students exhibit a different pattern of impact. Following a flooding event, there is a significant increase in the percentage of middle school students scoring at the novice level. Additionally, there is a decrease in the percentage of middle school students achieving at the distinguished level, highlighting a decline in academic excellence among all students and students on free/reduced price meals post-disaster. For high school students, there are increases in students in novice level for both all students and free/reduced price meal ones. For those eligible for free/reduced-price meals, there is a smaller decrease in the percentage of students achieving at the proficient level compared to their peers.

Previous research indicates that both natural disasters and food assistance programs can have differential effects on students' math and reading test scores. Figure 3.2 reports the effect of flooding events on student achievement by subjects: mathematics and readings. Mathematics and reading are subjects systematically tested across multiple grade levels, providing a robust dataset for analysis. Unlike science, social studies, and writing, which are tested less frequently, math and reading assessments span from third to eighth grade, offering a comprehensive view of academic performance trends. The results highlight distinct patterns in how flooding events affect student achievement in these subjects. Specifically, students attending schools affected by flooding experience a notable decrease in the percentage scoring at proficient levels in mathematics. This decline suggests that disruptions caused by natural disasters can impair students' ability to grasp and apply mathematical concepts effectively, potentially due to interruptions in learning continuity or environmental stressors. Moreover, the impact appears less pronounced for students receiving free or reduced-price meals, indicating a mitigating effect of food assistance programs on math achievement post-disaster. In contrast, the results shows no significant effect of flooding events on reading scores for both student groups. This finding suggests that reading comprehension, which may rely more on continuous practice and less on cumulative learning, is less immediately disrupted by environmental factors such as school closures or displacement caused by natural disasters.

Figure 3.3 illustrates the impact of flooding events on student achievement across urban and rural counties. I matched school district data with county-level demographic data collected from the American Community Survey (ACS) 5-year estimates from 2013 to 2018. A county is classified as urban if at least 50% of its population lives in urban areas. Rural areas frequently experience higher rates of food insecurity compared to urban areas, meaning that students in these areas could benefit more from food assistance programs. Research by Gordanier et al. (2020) demonstrates that the universal free-lunch program, Community Eligibility Provision (CEP), improves math and reading scores, with the effects being more pronounced in rural schools compared to urban ones. The graph indicates that urban students in schools experiencing a PDD event within one school year show a significant increase in the percentage scoring at the novice level for all students and those on free/reduced-price meals. In rural areas, there is a substantial increase in the percentage of students scoring at the distinguished level, with a more significant effect for students on free/reduced-price meals. Additionally, there is a notable decrease in the percentage scoring at the novice level for students receiving free/reduced-price meals. These improvements can be attributed to several factors. First, flooding events often lead to emergency responses that include increased access to food assistance programs, providing crucial nutritional support that helps students focus better in school. Furthermore, post-disaster periods typically see heightened levels of community and government intervention, bringing in additional resources such as educational support and assistance for individual households. Lastly, recovery efforts often prioritize the most vulnerable populations, including those on free/reduced-price meals, directly impacting their academic performance by addressing both immediate and long-term needs.

3.6 Limitation

The analyses of associations between student academic performance and the impact of flooding events are limited by the spatial resolution of available data at the school level. Specifically, the data only allow for observations of overall enrollment, and the percentage of students scoring at each achievement level by grade and subject. This aggregated data lack the granularity needed to track individual student outcomes, such as whether a student dropped out or was placed in another school after a PDD event. Consequently, if a significant number of students failed to take the test due to family displacement, transportation issues, or being forced to work to support their households, the analysis results could be biased and underestimates the true impact of flooding events on academic performance. Despite these limitations, the analysis represents a valuable step toward understanding the educational outcomes linked to environmental factors and food assistance programs, highlighting the need for more detailed data collection and targeted interventions to support vulnerable student populations in the wake of natural disasters.

When using school-level data to compare the entire student population with students who receive free or reduced-price meals, the results should be interpreted carefully due to the inherent differences between these groups. Students eligible for free/reduced-price meals typically come from lower-income households, which can be associated with various socio-economic disadvantages, including limited access to resources, less educational support at home, and higher levels of stress. These factors can significantly influence academic performance and overall well-being, making it essential to consider these disparities when analyzing the data. The main results show the comparisons of the changes in test performance between the general student population and under-performing students. Recognizing these differences allows for more nuanced interpretations and understanding any observed disparities in educational outcomes within the context of underlying socio-economic factors. Furthermore, these comparisons highlight the importance of tailored interventions. Educators and policymakers can design more effective support systems by understanding the unique needs and obstacles of students from low-income households. This approach not only helps improve academic outcomes but also addresses broader issues of equity and access within the education system.

3.7 Discussion and Conclusion

This paper examines the impact of national school meals programs on the test scores of Kentucky students post-disaster. The study fills a critical gap in the existing literature, which has separately shown evidence of the impact of disasters and food assistance programs on student achievements. Specifically, this study investigates how national school meals programs aid economically disadvantaged students after disasters, providing a unique perspective on the intersection of these two factors. The results indicate that schools experiencing a significant flooding event within the school year see the percentage of students scoring at the novice level increase by 0.34 percentage points. However, students receiving free or reduced-price meals in these schools show an insignificant change in test performance. This suggests that the national school meals program plays a critical role in mitigating some of the adverse effects of large regional floods on these students. This stabilization effect highlights the importance of food security in maintaining academic performance during crises. Further analysis reveals heterogeneous effects based on school level and urban/rural distinctions. Elementary school students benefit more from the programs following a disaster compared to their middle and high school counterparts. Additionally, the gains from the school meals programs are larger in rural schools than in urban ones. The findings advocate for the continued and potentially expanded implementation of these programs, especially in vulnerable rural communities and among younger students, to foster resilience and academic success in the face of natural disasters.

A new strand of literature has focused on food security programs during the COVID-19 pandemic. As lockdown and school closure policies were implemented in response to the coronavirus, the government introduced several emergency feeding programs to address food security and support food assistance programs. These measures included initiatives such as the Pandemic Electronic Benefit Transfer (P-EBT), which provided funds to families to replace the value of missed school meals, and the expansion of free meal distribution sites to ensure that children continued to have access to nutritious food despite school closures. While no existing literature examines explicitly the impact of these programs on student achievement yet, studies on emergency feeding programs for students during the pandemic reveal significant parallels to this study. For instance, research indicates that these emergency programs played a crucial role in preventing increased food insecurity among low-income families and maintaining children's access to essential nutrients (Jablonski et al., 2021; McLoughlin et al., 2020). Both scenarios underline the critical role of school-based meal programs in ensuring continuous access to nutritious food for vulnerable students amid crises. The findings from the pandemic context highlight how government interventions on food security for children can mitigate adverse effects on student nutrition and well-being, providing a blueprint for effective responses to various emergencies. These interventions help stabilize household food supplies, reduce the stress and anxiety associated with food insecurity, and ensure that children remain healthy and ready to learn. The experience from the COVID-19 pandemic demonstrates the importance of flexibility and rapid response in food assistance programs, which can be crucial for maintaining educational outcomes during times of crisis. By drawing lessons from these emergency measures, policymakers can develop robust strategies that ensure food security and support student achievement in future emergencies, whether natural disasters, economic downturns, or public health crises.

Overall, the results suggest that food assistance programs such as the National School Lunch Program (NSLP) and School Breakfast Program (SBP) can help economically disadvantaged students mitigate some of the negative impacts of flooding events and maintain their academic performance. Understanding the mechanisms through which food assistance programs influence educational outcomes is crucial for devising effective strategies to mitigate the adverse effects of natural disasters on education. Research shows that consistent access to nutritious meals not only supports physical health but also enhances cognitive function, concentration, and overall academic performance. By analyzing how these programs provide stability, routine, and stress alleviation for students and their families, policymakers and educators can identify key pathways for implementation. This knowledge enables the development of targeted interventions that ensure food security during and after natural disasters, such as mobile meal delivery systems, expanded eligibility criteria, and robust community partnerships. By leveraging these insights, programs can be tailored to maintain educational continuity and support student well-being, ultimately fostering resilience and academic success in the face of crises. Integrating these strategies into disaster response plans can ensure that vulnerable students continue receiving the nutritional support necessary for their academic and personal development, even in challenging circumstances.

3.8 Tables

Table 3.1: Summary statistics by school level

	All Students						Free/Reduced Price Meal					
	Elementary		Middle		High		Elementary		Middle		High	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Novice (%)	17.58	11.65	15.88	9.36	13.01	7.91	21.21	12.15	20.48	10.50	16.80	9.04
Apprentice (%)	31.01	11.33	33.94	12.63	30.65	9.10	34.42	11.71	37.98	12.38	35.04	9.16
Proficient (%)	36.37	11.12	36.26	10.25	46.40	10.73	33.90	11.89	32.63	10.97	41.83	11.48
Distinguished (%)	15.02	11.08	13.89	9.59	9.91	6.93	10.46	8.57	8.89	6.84	6.32	5.15
Enrollment	72.45	34.93	157.50	107.68	195.98	119.59	45.61	23.73	93.49	66.21	101.98	58.39
Observation	33,238		14,453		1,293		28,689		12,951		1,226	

Notes. Table provides the mean the percentage of students at 4 different achievement categories and enrollment by school level.

Table 3.2: Summary statistics by flooded schools

	Flooded Schools				Non-Flooded Schools			
	All Student		Free/Reduced Price Meal		All Student		Free/Reduced Price Meal	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	S.D.	Mean
Novice (%)	19.35	12.90	23.60	12.62	16.80	10.82	20.69	11.54
Apprentice (%)	29.77	10.72	33.33	11.00	32.01	11.80	35.65	12.01
Proficient (%)	34.91	10.92	32.00	11.17	36.72	10.97	33.85	11.72
Distinguished (%)	15.96	11.59	11.04	8.67	14.46	10.54	9.79	8.02
Enrollment	117.49	90.74	72.04	58.25	99.29	78.45	60.26	46.63
Observation	3,066		2,518		45,918		40,348	

Notes. Table provides the mean the percentage of students at 4 different achievement categories and enrollment by flooded and non-flooded schools.

Table 3.3: Effect of Flood events on students performance

	All Students				Free/Reduced Price Meal			
	Novice	Apprentice	Proficient	Distinguished	Novice	Apprentice	Proficient	Distinguished
PDD within 1 school year	0.344*	-0.117	-0.266	0.0394	0.312	-0.296	-0.211	0.193
	(0.192)	(0.263)	(0.225)	(0.207)	(0.237)	(0.293)	(0.261)	(0.180)
Constant	16.64***	31.65***	36.82***	14.89***	20.85***	35.53***	33.76***	9.859***
	(0.0380)	(0.0520)	(0.0444)	(0.0408)	(0.0467)	(0.0578)	(0.0515)	(0.0355)
Observations	42,865	42,865	42,865	42,865	42,866	42,866	42,866	42,866
R-squared	0.487	0.201	0.311	0.402	0.386	0.112	0.263	0.263
	All Students				Free/Reduced Price Meal			
	Novice	Apprentice	Proficient	Distinguished	Novice	Apprentice	Proficient	Distinguished
PDD within 1 school semester	0.395*	-0.223	-0.463	-0.0579	0.328	-0.265	-0.786	0.113
	(0.166)	(0.204)	(0.201)	(0.175)	(0.209)	(0.237)	(0.241)	(0.162)
Constant	16.59***	31.60***	36.86***	14.95***	20.82***	35.52***	33.78***	9.882***
	(0.0314)	(0.0386)	(0.0381)	(0.0332)	(0.0397)	(0.0449)	(0.0456)	(0.0307)
Observations	42,508	42,508	42,508	42,508	42,509	42,509	42,509	42,509
R-squared	0.709	0.639	0.583	0.677	0.634	0.562	0.527	0.548

Notes. Dependent variable is the percentage of students at 4 different achievement categories: novice, apprentice, proficient, and distinguished. All the regressions include school-grade and year fixed effects. Standard errors are stated in parentheses and are clustered at school level. *** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Robust: Effect of Flood events on students performance

	All Students				Free/Reduced Price Meal			
	Novice	Apprentice	Proficient	Distinguished	Novice	Apprentice	Proficient	Distinguished
PDD within 1 school year	0.314*	-0.0515	-0.225	-0.0368	0.240	-0.242	-0.175	0.175
	(0.162)	(0.199)	(0.197)	(0.171)	(0.204)	(0.231)	(0.235)	(0.158)
Student-teacher ratio	0.386***	0.160**	-0.253***	-0.293***	0.517***	0.301***	-0.419***	-0.399***
	(0.0654)	(0.0805)	(0.0795)	(0.0691)	(0.0827)	(0.0936)	(0.0950)	(0.0640)
Observations	42,508	42,508	42,508	42,508	42,509	42,509	42,509	42,509
R-squared	0.709	0.639	0.583	0.678	0.635	0.562	0.527	0.549
	All Students				Free/Reduced Price Meal			
	Novice	Apprentice	Proficient	Distinguished	Novice	Apprentice	Proficient	Distinguished
PDD within 1 school semester	0.352*	-0.0674	-0.155	-0.0685	0.252	-0.274	-0.0652	0.126
	(0.166)	(0.204)	(0.201)	(0.175)	(0.209)	(0.237)	(0.240)	(0.162)
Student-teacher ratio	0.386***	0.160**	-0.253***	-0.293***	0.517***	0.301***	-0.420***	-0.398***
	(0.0654)	(0.0805)	(0.0795)	(0.0691)	(0.0827)	(0.0936)	(0.0950)	(0.0640)
Observations	42,508	42,508	42,508	42,508	42,509	42,509	42,509	42,509
R-squared	0.709	0.639	0.583	0.678	0.635	0.562	0.527	0.549

Notes. Dependent variable is the percentage of students at 4 different achievement categories: novice, apprentice, proficient, and distinguished. All the regressions include school-grade and year fixed effects. Standard errors are stated in parentheses and are clustered at school level.

3.9 Figures

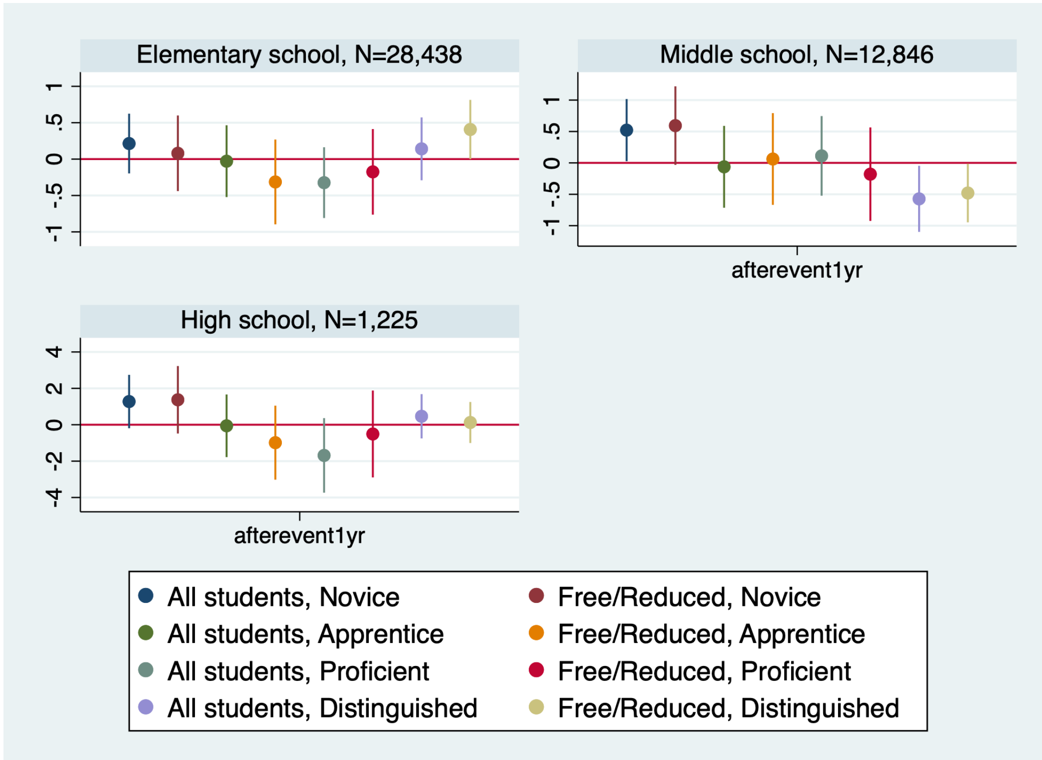

Figure 3.1: Heterogeneous effects by school level

Notes. Dependent variable is the percentage of students at 4 different achievement categories: novice, apprentice, proficient, and distinguished. All the regressions include school-grade and year fixed effects. Standard errors are clustered at school level.

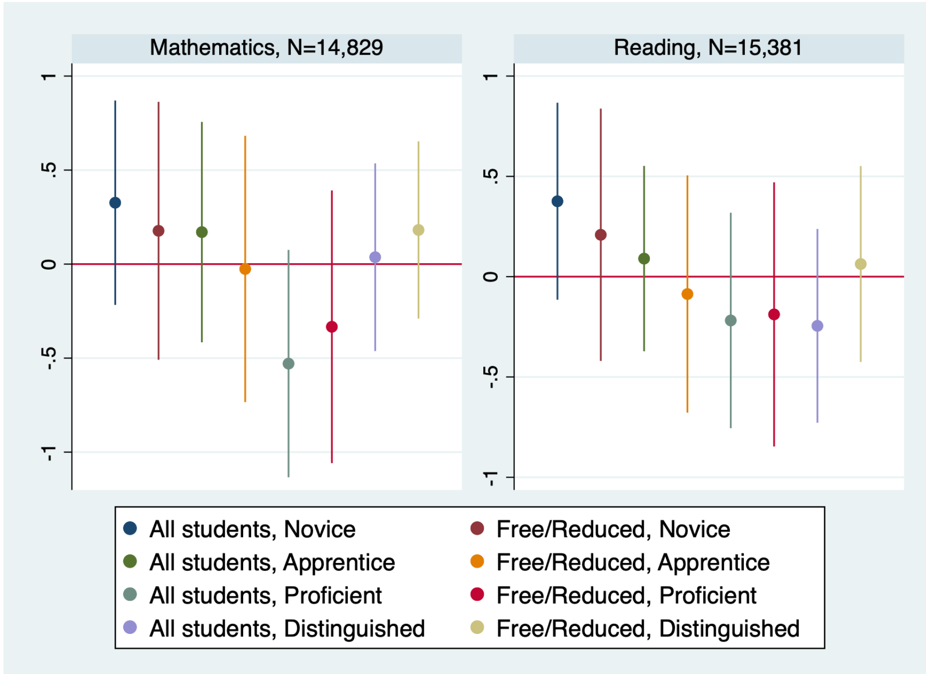

Figure 3.2: Heterogeneous effects by subjects

Notes. Dependent variable is the percentage of students at 4 different achievement categories: novice, apprentice, proficient, and distinguished. All the regressions include school-grade and year fixed effects. Standard errors are clustered at school level.

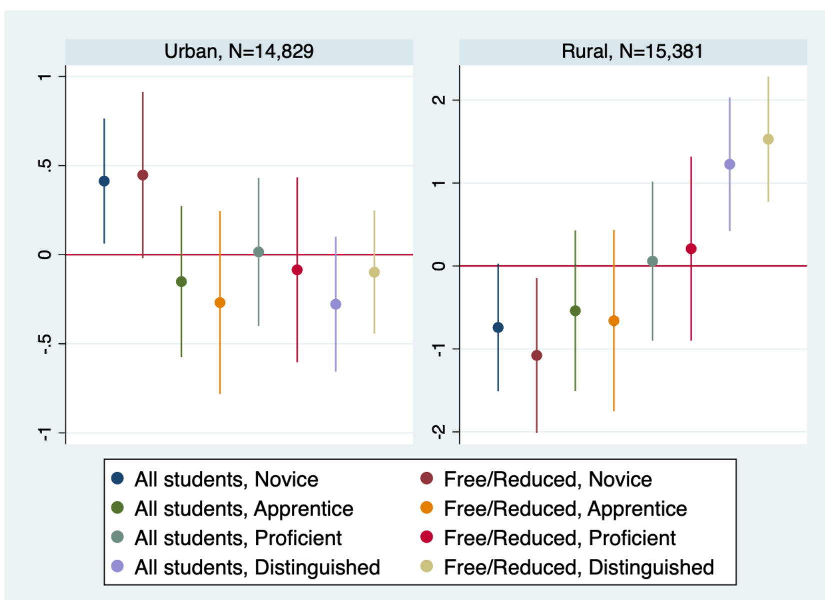

Figure 3.3: Heterogeneous effects by Urban/Rural Counties

Notes. Dependent variable is the percentage of students at 4 different achievement categories: novice, apprentice, proficient, and distinguished. Counties with more than 50 percent of the population living in urban areas are classified urban. All the regressions include school-grade and year fixed effects. Standard errors are clustered at school level.

Bibliography

Alaimo, K., C. M. Olson, and E. A. Frongillo Jr (2001). Food insufficiency and american school-aged children's cognitive, academic, and psychosocial development. *Pediatrics 108*(1), 44–53.

Allaire, M., H. Wu, and U. Lall (2018). National trends in drinking water quality violations. *Proceedings of the National Academy of Sciences 115*(9), 2078–2083.

Andrade, L., J. O'Dwyer, E. O'Neill, and P. Hynds (2018). Surface water flooding, groundwater contamination, and enteric disease in developed countries: A scoping review of connections and consequences. *Environmental pollution 236*, 540–549.

Ashiabi, G. (2005). Household food insecurity and children's school engagement. *Journal of children and poverty 11*(1), 3–17.

Atreya, A., S. Ferreira, and W. Kriesel (2013). Forgetting the flood? an analysis of the flood risk discount over time. *Land Economics 89*(4), 577–596.

Bakkensen, L. A., X. Ding, and L. Ma (2019). Flood risk and salience: New evidence from the sunshine state. *Southern Economic Journal 85*(4), 1132–1158.

Balazs, C., R. Morello-Frosch, A. Hubbard, and I. Ray (2011). Social disparities in nitrate-contaminated drinking water in california's san joaquin valley. *Environmental Health Perspectives 119*(9), 1272–1278.

Balazs, C. L., R. Morello-Frosch, A. E. Hubbard, and I. Ray (2012). Environmental justice implications of arsenic contamination in california's san joaquin valley: a cross-sectional, cluster-design examining exposure and compliance in community drinking water systems. *Environmental Health 11*(1), 1–12.

Balazs, C. L. and I. Ray (2014). The drinking water disparities framework: on the origins and persistence of inequities in exposure. *American journal of public health 104*(4), 603–611.

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

Billings, S. B., E. A. Gallagher, and L. Ricketts (2022). Let the rich be flooded: the distribution of financial aid and distress after hurricane harvey. *Journal of Financial Economics 146*(2), 797–819.

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

Bin, O. and S. Polasky (2004). Effects of flood hazards on property values: evidence before and after hurricane floyd. *Land Economics 80*(4), 490–500.

Blaikie, P., T. Cannon, I. Davis, and B. Wisner (2014). *At risk: natural hazards, people's vulnerability and disasters*. Routledge.

Brown, M. J. and S. Margolis (2012). Lead in drinking water and human blood lead levels in the united states.

Chen, J., H. Wu, H. Qian, and Y. Gao (2017). Assessing nitrate and fluoride contaminants in drinking water and their health risk of rural residents living in a semiarid region of northwest china. *Exposure and Health 9*(3), 183–195.

Cohen, J. F., A. A. Hecht, G. M. McLoughlin, L. Turner, and M. B. Schwartz (2021). Universal school meals and associations with student participation, attendance, academic performance, diet quality, food security, and body mass index: a systematic review. *Nutrients 13*(3), 911.

Council, N. R., D. on Earth, L. Studies, C. on Life Sciences, and S. on Arsenic in Drinking Water (1999). Arsenic in drinking water.

Cutter, S. L. (2012). *Hazards vulnerability and environmental justice*. Routledge.

Cutter, S. L., B. J. Boruff, and W. L. Shirley (2012). Social vulnerability to environmental hazards. In *Hazards vulnerability and environmental justice*, pp. 143–160. Routledge.

Delpla, I., T. Benmarhnia, A. Lebel, P. Levallois, and M. J. Rodriguez (2015). Investigating social inequalities in exposure to drinking water contaminants in rural areas. *Environmental Pollution 207*, 88–96.

Domingue, S. J. and C. T. Emrich (2019). Social vulnerability and procedural equity: Exploring the distribution of disaster aid across counties in the united states. *The American Review of Public Administration 49*(8), 897–913.

Drakes, O., E. Tate, J. Rainey, and S. Brody (2021). Social vulnerability and short-term disaster assistance in the united states. *International Journal of Disaster Risk Reduction 53*, 102010.

(EPA), U. E. P. A. (2013). Providing safe drinking water in america: 2013 national public water systems compliance report.

Finch, C., C. T. Emrich, and S. L. Cutter (2010). Disaster disparities and differential recovery in new orleans. *Population and environment 31*, 179–202.

Gayer, T., J. T. Hamilton, and W. K. Viscusi (2000). Private values of risk tradeoffs at superfund sites: housing market evidence on learning about risk. *Review of Economics and Statistics 82*(3), 439–451.

Gibbs, L., J. Nursey, J. Cook, G. Ireton, N. Alkemade, M. Roberts, H. C. Gallagher, R. Bryant, K. Block, R. Molyneaux, et al. (2019). Delayed disaster impacts on academic performance of primary school children. *Child development 90*(4), 1402–1412.

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

Gordanier, J., O. Ozturk, B. Williams, and C. Zhan (2020). Free lunch for all! the effect of the community eligibility provision on academic outcomes. *Economics of Education Review 77*, 101999.

Gruber, J. S., A. Ercumen, and J. M. Colford Jr (2014). Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PloS one 9*(9), e107429.

Hallstrom, D. G. and V. K. Smith (2005). Market responses to hurricanes. *Journal of Environmental Economics and Management 50*(3), 541–561.

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

Hino, M. and M. Burke (2020). Does information about climate risk affect property values? Technical report, National Bureau of Economic Research.

Hopenhayn, C. (2006). Arsenic in drinking water: impact on human health. *Elements 2*(2), 103–107.

Howell, J. and J. R. Elliott (2018). As disaster costs rise, so does inequality. *Socius 4*, 2378023118816795.

Howell, J. and J. R. Elliott (2019). Damages done: The longitudinal impacts of natural hazards on wealth inequality in the united states. *Social problems 66*(3), 448–467.

Imberman, S. A. and A. D. Kugler (2014). The effect of providing breakfast in class on student performance. *Journal of Policy Analysis and Management 33*(3), 669–699.

Jablonski, B. B., J. Casnovsky, J. K. Clark, R. Cleary, B. Feingold, D. Freedman, S. Gray, X. Romeiko, L. S. Olabisi, M. Torres, et al. (2021). Emergency food provision for children and families during the covid-19 pandemic: examples from five us cities. *Applied Economic Perspectives and Policy 43*(1), 169–184.

Johnson, G. S. (2008). Environmental justice and katrina: A senseless environmental disaster. *Western Journal of Black Studies 32*(1), 42.

Kleinman, R. E., S. Hall, H. Green, D. Korzec-Ramirez, K. Patton, M. E. Pagano, and J. M. Murphy (2002). Diet, breakfast, and academic performance in children. *Annals of Nutrition and Metabolism 46*(Suppl. 1), 24–30.

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

MacDonald, D. N., J. C. Murdoch, and H. L. White (1987). Uncertain hazards, insurance, and consumer choice: evidence from housing markets. *Land Economics 63*(4), 361–371.

McDonald, Y. J. and N. E. Jones (2018). Drinking water violations and environmental justice in the united states, 2011–2015. *American Journal of Public Health 108*(10), 1401–1407.

McLoughlin, G. M., J. A. McCarthy, J. T. McGuirt, C. R. Singleton, C. G. Dunn, and P. Gadhoke (2020). Addressing food insecurity through a health equity lens: a case study of large urban school districts during the covid-19 pandemic. *Journal of Urban Health 97*, 759–775.

Mishra, A., A. Alnahit, and B. Campbell (2021). Impact of land uses, drought, flood, wildfire, and cascading events on water quality and microbial communities: A review and analysis. *Journal of Hydrology 596*, 125707.

Mudavanhu, C. (2014). The impact of flood disasters on child education in muzarabani district, zimbabwe.

Nogueira, G., C. V. Nakamura, M. C. Tognim, B. A. Abreu Filho, and B. P. Dias Filho (2003). Microbiological quality of drinking water of urban and rural communities, brazil. *Revista de Saúde Pública 37*, 232–236.

Patel, A. I., C. E. Hecht, A. Cradock, M. A. Edwards, and L. D. Ritchie (2020). Drinking water in the united states: implications of water safety, access, and consumption. *Annual Review of Nutrition 40*(1), 345–373.

Peterson, C. (2014). Investigating the historic long-term population health impact of the us national school lunch program. *Public health nutrition 17*(12), 2783–2789.

Rajapaksa, D., C. Wilson, S. Managi, V. Hoang, and B. Lee (2016). Flood risk information, actual floods and property values: a quasi-experimental analysis. *Economic Record 92*, 52–67.

Rothbart, M. W., A. E. Schwartz, and E. Gutierrez (2023). Paying for free lunch: the impact of cep universal free meals on revenues, spending, and student health. *Education Finance and Policy 18*(4), 708–737.

Ryder, S. S. (2017). A bridge to challenging environmental inequality: Intersectionality, environmental justice, and disaster vulnerability. *Social Thought & Research: A Continuation of the Mid-American Review of Sociology*, 85–115.

Schaider, L. A., L. Swetschinski, C. Campbell, and R. A. Rudel (2019). Environmental justice and drinking water quality: are there socioeconomic disparities in nitrate levels in us drinking water? *Environmental Health 18*(1), 1–15.

Schwartz, A. E. and M. W. Rothbart (2020). Let them eat lunch: The impact of universal free meals on student performance. *Journal of Policy Analysis and Management 39*(2), 376–410.

Shr, Y.-H. J. and K. Y. Zipp (2019). The aftermath of flood zone remapping: The asymmetric impact of flood maps on housing prices. *Land Economics 95*(2), 174–192.

Smith, A. B. (2020). U.S. billion-dollar weather and climate disasters, 1980-present (NCEI Accession 0209268). *NOAA National Centers for Environmental Information*.

Smith, A. H., E. O. Lingas, and M. Rahman (2000). Contamination of drinking-water by arsenic in bangladesh: a public health emergency. *Bulletin of the World Health Organization 78*(9), 1093–1103.

Stillo, F. and J. MacDonald Gibson (2017). Exposure to contaminated drinking water and health disparities in north carolina. *American Journal of Public Health 107*(1), 180–185.

Strosnider, H. (2017). Rural and urban differences in air quality, 2008–2012, and community drinking water quality, 2010–2015—united states. *MMWR. Surveillance Summaries 66*.

Switzer, D. and M. P. Teodoro (2018). Class, race, ethnicity, and justice in safe drinking water compliance. *Social Science Quarterly 99*(2), 524–535.

Ten Veldhuis, J., F. Clemens, G. Sterk, and B. Berends (2010). Microbial risks associated with exposure to pathogens in contaminated urban flood water. *Water research 44*(9), 2910–2918.

Thamtanajit, K. (2020). The impacts of natural disaster on student achievement: Evidence from severe floods in thailand. *The Journal of Developing Areas 54*(4).

Tversky, A. and D. Kahneman (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology 5*(2), 207–232.

Viscusi, W. K. (1991). Economic theories of decision making underuncertainty: Implications for policy analysis. In *Policy analysis and economics*, pp. 85–109. Springer.

Ward, M. H., R. R. Jones, J. D. Brender, T. M. De Kok, P. J. Weyer, B. T. Nolan, C. M. Villanueva, and S. G. Van Breda (2018). Drinking water nitrate and human health: an updated review. *International journal of environmental research and public health 15*(7), 1557.

Yard, E. E., M. W. Murphy, C. Schneeberger, J. Narayanan, E. Hoo, A. Freiman, L. S. Lewis, and V. R. Hill (2014). Microbial and chemical contamination during and after flooding in the ohio river—kentucky, 2011. *Journal of Environmental Science and Health, Part A 49*(11), 1236–1243.

Yi, D. and H. Choi (2020). Housing market response to new flood risk information and the impact on poor tenant. *The Journal of Real Estate Finance and Economics 61*(1), 55–79.

Vita

Fang-Yu Yeh

Education

M.S. in Economics, University of Kentucky, 2019

B.S. in Electrical Engineering, National Taiwan Normal University, Taiwan, 2016

Employment

Graduate Teaching Assistant, University of Kentucky 2018-2024

Research Assistant, University of Kentucky 2022-2023

Awards and Fellowsips

Jerry Ellig Award for Best Student Paper, Benefit-Cost Analysis Annual Conference, 2023

Max Steckler Fellowship, University of Kentucky, 2023

Gatton Doctoral Fellowship, University of Kentucky, 2018-2021