Efficient Adaptation to Flood Risk

Online Appendix

Winston P. Hovekamp
Yale University

Katherine R. H. Wagner
University of British Columbia

January 12, 2023

A Data

Flood Insurance Policies and Claims—The administrative records on flood insurance policies and claims are identical to the data used by Wagner (2022). We outline the restrictions imposed both here and in this other paper to arrive at the final analysis sample.

The full data set includes the universe of flood insurance policies and claims written by the NFIP between 2001 and 2017 for 20 Atlantic and Gulf Coast U.S. states. These microdata are from FEMA's BureauNet database, which the NFIP itself uses to track its internal operations. The 20 states are Alabama, Connecticut, Delaware, Florida, Georgia, Louisiana, Maine, Maryland, Massachusetts, Mississippi, New Hampshire, New Jersey, New York, North Carolina, Pennsylvania, Rhode Island, South Carolina, Texas, Vermont, and Virginia. These states together account for approximate 85% of total flood insurance policies written in the U.S. (NRC, 2015).

The main analysis focuses on approximately 12 million observations on single-family primary residences located in high-risk flood zones. Following the NFIP rating system, we consider a house to be high-risk if it is located any of A, numbered A, V, or numbered V flood zones, and low-risk otherwise. Low-risk houses are not required to be elevated and so the effect of elevation can't be assessed for this group. Damages for multi-family homes, mobile dwellings, and condominiums also are unlikely to be a comparable control group for

elevated single-family residences. These other structure types also face different building code requirements, insurance prices, and take-up incentives than single family primary residences.

Evaluating the importance of house elevation necessitates matching house characteristics available in the contracts data set (e.g., differences in building codes) with the claims and damages associated with them. We therefore use information available from both data sets to match policies to claims in a manner that respects the prices and elevation differences between houses. Specifically, we match claims to policies based on policy written date, house construction year, flood zone, and zip code. These variables almost perfectly uniquely identify claims and policies and the match rate is 99%. In addition to the variables used for merging policies and claims, the data also include information on coverage, premia, and damages. Other relevant variables are premium paid, coverage purchased for building and contents, minimum elevation required, amount claimed, the NFIP community identification number, the flood identification number assigned by FEMA, and the depth of water that flooded the house.

We impose several additional restrictions on the single-family primary residence sample to arrive at the data set used in the analysis. We exclude 1% of policies that are missing flood zone or construction data since these variables are needed to identify whether houses are required to be elevated or not. We also exclude 4% of policies with coverage that is negative, equal to 0, or exceeds the maximum amount available for purchase. In addition, we exclude policies with premia smaller than the 1st and larger than 99th percentile of the distribution of values because these are clearly miscoded relative to the NFIP rate schedule, where policies are not written including e.g., total premia in excess of $60,000 per year or $16,000 per $1,000 of coverage or less than less than $0.10 per $1,000 of insurance are

not written. We impose similar restrictions on the claims data, dropping the 7% of claims reporting damages or payouts that are zero or negative, or with realized payouts that exceed purchased coverage.[1]

Zip codes are a key variable on which we merge claims to policies and insurance data to features of the natural and built environment. The data for the years 2010-2017 are missing between 5 and 10% of zip codes, which were erroneously deleted when the data were anonymized. As in Wagner (2022), we recover these zip codes by building a concordance from zip code to "flood map panel identifier", which is the unit of analysis of FEMA hydrological studies and is typically fully contained within a zip code. We recover approximately 75% of missing zip codes by identifying others with the same flood map panel identifier and assigning them to the same zip code.

Flood Severity—One of the heterogeneity analyses we are interested in is the extent to which the benefits of adaptation vary with flood severity. This information is not available in the raw insurance data, and so we construct it using information on flood type and floodwater depth from the claims data.

Each claim is associated with a "flood event number" assigned by FEMA that identifies whether it was incurred during a uniquely identified presidentially declared disaster event (PDD) or during a "nuisance" flood. FEMA assigns innocuous nuisance floods an identifying number of zero, while catastrophic flooding receives a unique identifier corresponding to the disaster (e.g., Hurricane Katrina). We identify the maximum of the flood event numbers in each zip code $\times$ year to determine whether or not a catastrophic flood struck the zip code.

[1] Zero entries for damages or payouts indicate either that no payout was made or that the claim is still outstanding.

We assign zip codes without any claims to a third, "not flooded" category.

Each claim also includes information on the depth of the water that flooded the house during the flood event. We average flood depth in each zip code-year and bin the depths into quintiles, which results in four bins since approximately 40% of zip codes are not flooded in a given year. Before doing so, we first assign water depths of 0.0001 to claims with recorded depth of zero because depths are rounded to the nearest foot and then assign depths of zero to policies without claims, to distinguish small floods from no floods. We treat the 2% of observations with negative water depths as missing and impute the depth implied by claims made by the same type of house (i.e., elevated or not) in the same flood zone in the same flood event. We follow the same imputation procedure for the 7% of claims with recorded water depths exceeding 25 feet.

We use these measures of flood severity based on event type and floodwater depth to define six monotonically increasing water depth categories. The first comprises zip codes that are not flooded (i.e., quintiles one and two). The second comprises nuisance floods in the third quintile of flood depth. The remaining four categories split the fourth and fifth quintiles of flood depths into nuisance floods and catastrophes, respectively. Wagner (2022) discusses analysis that ascertains the validity of these flood severity measures.

Land Use—The aggregation of the land use data is discussed in Taylor and Druckenmiller (2022). The data are derived from the National Land Cover Database, which provides remotely-sensed information on the spatial extent of different types of land cover for the United States at 30-m resolution. The database identifies 21 different classes of land use, including wetlands (categories 91 and 92). The data are aggregated to the zip code level by intersecting the spatial data with zip code maps. We merge the zip code aggregates for

developed share and wetlands share in 2001 (our first year of data) with the flood insurance data set.

References

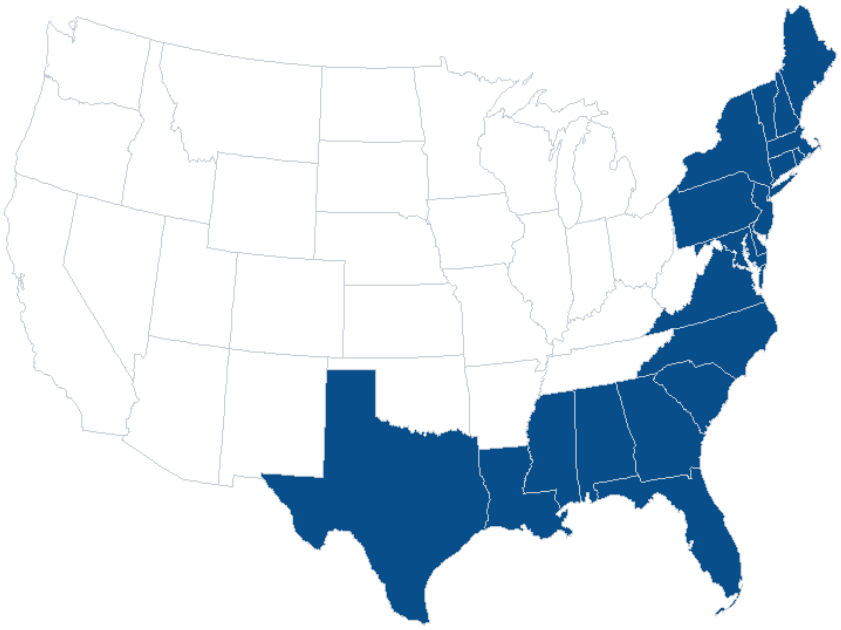
NRC (2015). Affordability of national flood insurance program premiums, report 1. Technical report, National Research Council.

Taylor, C. A. and Druckenmiller, H. (2022). Wetlands, flooding, and the clean water act. *American Economic Review*, 112(4):1334–1363.

Wagner, K. R. H. (2022). Adaptation and adverse selection in markets for natural disaster insurance. *American Economic Journal: Economic Policy*, 14(3):380–421.

B Figures and Tables

Figure A.1: States in Analysis



Notes: The 20 states included in the analysis are Alabama, Connecticut, Delaware, Florida, Georgia, Louisiana, Maine, Maryland, Massachusetts, Mississippi, New Hampshire, New Jersey, New York, North Carolina, Pennsylvania, Rhode Island, South Carolina, Texas, Vermont, and Virginia. These 20 states account for 83% of flood insurance policies written nationwide (NRC, 2015).

Table A.1: Heterogeneous Effects of Adaptation Policy on Flood Damages

	No Flood		Flood		Catastrophe	
	Average Cost	Any Claim	Average Cost	Any Claim	Average Cost	Any Claim
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Average Effect						
Adapted	0.000	0.000	-0.276***	-0.002***	-10.038***	-0.017***
	(0.000)	(0.000)	(0.070)	(0.001)	(1.686)	(0.002)
Panel B: Wetlands Heterogeneity						
Adapted	0.000	0.000	-0.250***	-0.001***	-8.662***	-0.014***
	(0.000)	(0.000)	(0.082)	(0.001)	(2.189)	(0.003)
Adapted × Wetland Fraction	0.000	0.000	-0.193	-0.003**	-4.169	-0.012
	(0.000)	(0.000)	(0.316)	(0.001)	(6.171)	(0.010)
Panel C: Developed Heterogeneity						
Adapted	0.000	0.000	-0.639***	-0.003***	-13.187***	-0.025***
	(0.000)	(0.000)	(0.139)	(0.001)	(2.497)	(0.004)
Adapted × Developed Fraction	0.000	0.000	0.770	0.003***	8.366*	0.019**
	(0.000)	(0.000)	(0.180)	(0.001)	(5.038)	(0.007)
Dep. Var. Mean	0.000	0.000	0.952	0.006	38.512	0.103
N	5,793,255		3,808,697		2,381,231	
Zip code × Year FE	✓	✓	✓	✓	✓	✓
Decade Built × Flood Severity Controls	✓	✓	✓	✓	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Outcome variables are average flood insurance cost (payouts per $1000 coverage) and an indicator variable equal to one if a claim is made. Adapted houses are built after communities are mapped and are required to be elevated. Wetlands fraction is the share of the zipcode covered by wetlands and developed fraction is the share of the zipcode devoted to urban infrastructure or dwellings. The dependent variable mean is for non-adapted houses. Decade built × flood severity controls are zip code × decade built × flood severity fixed effects and decade built × flood severity time trends. Flood severity is defined using flood water depth and flood event type (see text). Standard errors clustered by community are in parentheses.

Table A.2: Effects of Adaptation Policy on Flood Damages for Specific Disasters

	Hurricane Katrina		Hurricane Sandy		Hurricane Harvey	
	Average Cost	Any Claim	Average Cost	Any Claim	Average Cost	Any Claim
	(1)	(2)	(3)	(4)	(5)	(6)
Adapted	-14.042***	-0.029***	-4.040***	-0.006***	-7.334***	-0.010***
	(3.361)	(0.009)	(0.680)	(0.001)	(1.131)	(0.002)
Dep. Var. Mean	38.870	0.069	12.293	0.041	23.278	0.059
N	103,194		451,844		709,683	
Zip code × Decade Built FE	✓	✓	✓	✓	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Outcome variables are average flood insurance cost (payouts per $1000 coverage) and an indicator variable equal to one if a claim is made. Adapted houses are built after communities are mapped and are required to be elevated. The dependent variable mean is for non-adapted houses. Standard errors clustered by community are in parentheses.