Floods

Dev Patel

January 28, 2024*

Please Click Here for the Latest Version

Abstract

Floods threaten a quarter of the world's population, most of whom live in poor countries. How do floods impact economic development, and how do households adapt? To answer these questions, I first combine methods from geophysics and machine learning in the analysis of satellite data to detect inundation at a granular geographic level anywhere every day for the past two decades. Using this approach in Bangladesh, I find that floods cause a persistent decline in economic activity and force structural change by pushing employment out of agriculture, spurring migration, and shifting children into school. Places with recent exposure to floods experience less harm after subsequent inundation. Using a simple model of experience-driven adaptation, I derive empirical tests for two mechanisms underpinning this pattern and find evidence for both. In a survey of rural farmers, I first show that past flood exposure increases the perceived marginal benefit of adaptation investment by raising households' beliefs about future disaster risk and damages. I next find that the marginal cost of coping with floods via temporary urban migration declines in inundation experience. Consistent with this "learning-by-doing" channel, reduced mobility frictions identified from quasi-random variation in Colonial-era transportation networks mediate the differential treatment effects of past flood exposure. Together, my results indicate that endogenous adaptation will significantly reduce the damage from future flooding.

*Department of Economics, Harvard University: devpatel@g.harvard.edu. I am especially grateful to Emily Breza and Gautam Rao for their kindness, guidance, and encouragement. I thank Aakash Ahamed, Sarfaraz Alam, John Conlon, Melissa Dell, Ed Glaeser, Siddharth George, Claudia Goldin, Asim Khwaja, Gabriel Kreindler, Steven Lee, Ross Mattheis, Louise Paul-Delvaux, Justin Sandefur, Suproteem Sarkar, Corey Scher, Clara Sievert, Brit Sharoni, Evan Soltas, Arvind Subramanian, Elie Tamer, and Davide Viviano for helpful discussions. I am indebted to all of the farmers who participated in this research for their time and insights and my field team for their exceptional efforts. Roshni Islam and Noshin Othoi provided superb research assistance. This research underwent human subjects review at Harvard University. I gratefully acknowledge financial support from the Chae Family Economics Research Fund, the Center for International Development/Malcolm Wiener Center for Social Policy, the Harvard Climate Change Solutions Fund, the Harvard Mellon Urban Initiative, the Institute for Humane Studies under grants IHS016778 and IHS017455, the Mind, Brain, and Behavior Initiative, a National Science Foundation Graduate Research Fellowship under grant DGE1745303, the National Science Foundation Doctoral Dissertation Research Improvement Grant under grant 2242263, the Warburg Fund, and The Weiss Fund for Research in Development Economics at the University of Chicago. Click here for a guide to the flood detection method developed in this paper.

> "Lord, / when you send the rain / think about it, please, / a little? / Do / not get carried away / by the sound of falling water, / the marvelous light / on the falling water..."
>
> —*Untitled*, James Baldwin

1 Introduction

Nearly one in four people face significant flood risk, 89% of whom live in low- or middle-income countries (Rentschler et al., 2022).[1] Climate scientists forecast that under even conservative projections of global warming, much of the world will experience higher frequency and greater severity of flooding, already the world's most common natural disaster (Brunner et al., 2021). Floods pose the most significant threat to the world's poor, who disproportionately live where climate change will likely spur the largest increases in flooding (IPCC, 2022) and where such catastrophes already cause the most harm (Kahn, 2005).

How do floods impact economic development? This question requires accurate measures of floods. I develop a method to detect local flooding at a daily level anywhere in the world that corrects issues with current approaches. I implement the procedure in Bangladesh—ground zero for much of global warming's most devastating effects—to estimate the impact of floods on economic outcomes and households' adaptation to these most common natural disasters.

Limitations of existing data have hindered our understanding of the incidence and impacts of flooding.[2] Flood databases are typically created in one of two ways. First, some governments collect so-called *in situ* measurements of water-level data from sensors along rivers and oceanfronts. Converting these vertical shifts into flood measurements requires complex hydrological modeling relying on strong assumptions, an empirical challenge exacerbated by these sensors' sparse coverage. The second group uses news articles or government publications of flooding events.[3] But these sources lack a consistent definition of

[1] An estimated 1.81 billion people face inundation depths over 0.15 meters in the event of a 1-in-100-year flood.

[2] As an example of the policy impacts of these measurement issues, the organization GiveDirectly worked with Google Research's Flood Forecasting Initiative to send cash to poor households in Mozambique in the days ahead of severe floods. The only available data source for validating the flood model came from 50 years of government river level measurements, which feature drawbacks discussed in Section 2.4. Floods ultimately never arrived in the 11 villages that were predicted to be impacted and therefore sent cash transfers, while severe flooding did occur in other nearby areas that did not receive payments. GiveDirectly specifically highlighted increasing the predictive power of the flood forecasting models (which necessarily requires historical data on flooding for training) to improve service delivery moving forward (Shun and Lummis, 2023).

[3] Researchers frequently use the Dartmouth Flood Observatory Archive (Brakenridge, 2023) and the EM-DAT International Disaster Database (Guha-Sapir et al., 2023), both of which rely on these sources. Some

"flood", underreport floods in poorer areas, and largely omit short-term—yet economically important—inundation events.

I develop a method to measure floods without knowing *ex ante* where or when an inundation event occurs by combining methods from machine learning and geophysics in the analysis of satellite data. This strategy allows me to create a standard, objective definition of flooding held constant across space and over time while avoiding the reporting bias issues that generate undercoverage of small floods and those in developing countries. I exclusively use publicly available data available for the entire globe. In contrast to *in situ* data from river stations, my method can detect pluvial, fluvial, and coastal floods. My analysis primarily uses two classes of satellite instruments. Modern radar-based sensors exhibit astounding and granular accuracy in the detection of surface water, yet only orbit infrequently, potentially missing many short- and medium-length floods. By contrast, optical sensors can be thwarted by the clouds through which radar-based instruments can "see", creating significant bias when relying on these data alone to measure floods. What they lack in quality, however, they make up for in quantity, photographing every place on earth twice a day.

I use machine learning to combine the best features of both types of data. First, I use frontier techniques from geophysics to analyze the backscatter data collected by radar-based sensors to precisely measure local surface water from space. Notably, I measure water coverage each time the satellite passes over every location rather than restricting to the times and places where floods have been reported in another source. To address the undercoverage arising from the infrequent orbits of these instruments, I use machine learning to emulate the satellite information using other remote sensing data. Although cloud coverage makes direct analysis of the optical data difficult, this "missing" data could nevertheless contain meaningful signal. My machine learning approach lets the data tell me how these optical measures capture flooding conditions. Because these other remote sensing data are available both for the days when satellites carrying radar-based sensors orbit overhead and every other day for the past 20 years, I can detect flooding throughout the period. The resulting time series features both the accuracy of the radar-based instruments and the frequency of optical ones. Finally, I construct an objective definition of flooding based on precise deviations from

researchers have augmented this set of times and locations by incorporating *in situ* or remote sensing data to more precisely measure the coverage and duration of these flooding events. Tellman et al. (2021) provide an impressive recent example of this approach in constructing the Global Flood Database, using optical images from the MODIS instrument to map flooding events from the Dartmouth Flood Observatory (DFO) database. Even this approach can omit a significant share of floods, however. Due to cloud cover, flash floods, and inaccuracies in the DFO data—among other reasons—Tellman et al. (2021) only map 913 of 3,054 flooding events using their satellite data approach. In the case of Bangladesh, for instance, this amounts to only 23.71 percent of DFO flooding events. In another example of incorporating satellites into an existing record of flooding, the United Nations' UNOSAT Flood Portal calculates inundation extents for major flooding events that required U.N. assistance (UNOSAT, 2019).

local surface water while flexibly controlling for seasonal variation. I survey farmers to measure their recall of recent flooding and calibrate a threshold to convert the (enhanced) satellite measurements into a commonly-understood definition of a "flood."

To illustrate the effectiveness of this approach and test its validity, I use the procedure to measure local inundation in Bangladesh.[4] I estimate flood exposure every day for the past two decades for each of Bangladesh's 5,158 unions—a small administrative unit with a typical size of 10-20 square kilometers (about 2500 to 5000 acres). The prediction algorithm performs extremely well and explains nearly 90 percent of the variation in surface water as detected by the gold-standard satellite technology. I validate this measure of flooding against several other sources, including a database I built based on online news articles, digitized records from government flooding reports, and *in situ* data I obtained from the Bangladesh Water Development Board's river monitoring stations on water height every three hours. Across a battery of validation tests, my satellite-derived flooding measure aligns closely with each of these alternatives.

Having constructed the database, I turn to the main question of interest: how do floods impact economic development? Compared with many other natural disasters, floods are particularly well-suited from an econometric standpoint to identifying the causal impact of environmental shocks due to the local idiosyncratic variation in their incidence. The long-run and granular nature of my flood measure allow me to construct control groups to flooded areas that faced equivalent underlying flood risk but happened not to experience a flood during the treatment year. I create strata combining flooded and non-flooded areas with identical past flooding patterns, and within these groups, further restrict comparisons using a flood risk measure generated from a machine learning algorithm trained on underlying geographic characteristics. Using a stacked difference-in-differences design in an event-study specification, I estimate the causal impact of experiencing a flood.

I find that floods have large, negative, and persistent economic impacts, which can be seen even from space. Using satellite data on nighttime luminosity, I show that floods cause significant decreases in brightness—a measure of electricity use and a proxy for economic activity. These gaps last for at least seven years, the latest period for which I can observe treatment effects in the series. These overall impacts mask significant heterogeneity based on previous flooding history. I find strong empirical evidence that past inundation exposure mitigates damages. These effects are significantly stronger for places with above median luminosity prior to flooding, suggesting that this pattern is not mechanical due to a floor

[4]Ranked eighth in population across the world, fourth in population density (excluding countries with fewer than one million people), and tenth lowest in average elevation, Bangladesh faces likely the most consequential flood risks of anywhere. Going forward, the country will bear particularly increased likelihood of flooding under global warming projections (IPCC, 2022).

effect in measuring brightness. Furthermore, quantile regression results show consistent effects throughout moments of the distribution. Turning to granular remote sensing data tracing out the footprint of every building in Bangladesh, I document a significant decline in physical capital after flooding events as measured by the share of surface area covered by buildings.

To understand the mechanisms underpinning these findings, I turn to large-scale administrative data covering nearly 19 million individuals and find evidence that floods increase rates of structural transformation. Workers move out of agriculture into offices and businesses, and migration increases. Parents send their children to school at higher rates. The evidence suggests higher human capital investment reflects both a short-run decrease in agricultural productivity, lowering the opportunity cost of household labor, and an increase in households' long-run expectations about the relative returns to education. Using satellite data to measure the greenness of vegetation, I show a decline in agricultural productivity in the season immediately following a flood. In a survey I conduct with 2,279 rice farmers across southwestern Bangladesh, I find suggestive evidence that conditional on the underlying geographic propensity for flooding, parents who perceive higher flood risk also expect their children to attain more education and hold salaried instead of agricultural jobs.

What explains the heterogeneous treatment effects by past flood exposure? I use a simple model to illustrate two mechanisms underpinning why undergoing more floods in the past reduces their future harm, and I find empirical support for both. Households increase their beliefs about future flood risk, raising their perceived marginal benefit of investment. Separately, the adaptation technology itself changes, e.g., after paying a fixed cost after the first flood, the marginal return to investment shifts. Temporary urban migration—an important margin of adaptation—exhibits diminishing fixed costs with repeated travel. Using Colonial-era transportation networks for identification, I find suggestive evidence that spatial frictions mediate the heterogeneous effects by past flooding experience.

I build on four primary strands of literature. I contribute to a rich body of work focused on detecting flooding and add to a fast-growing literature combining machine learning with remote sensing data to capture important economic outcomes.[5] I demonstrate how this powerful combination can be used to predict other satellite data itself in the context of surface water, a principal that can be applied to a host of other settings where the best measurement technology leaves gaps in coverage, either spatially or temporally. I also add to the literature on the impact of climate shocks—particularly in low- and middle-income

[5]See, for example, Rolf et al. (2021), Walker et al. (2022), and Ratledge et al. (2022), among many others. See Guiteras et al. (2015), Tellman et al. (2021), Brakenridge (2023), and Guha-Sapir et al. (2023) for examples of flood measurement.

countries—by providing new evidence on the consequences of flooding.[6] Third, I relate to the rich economics literature examining the economics of adaptation to climate change.[7] My results point to market access and migration specifically as key channels facilitating households' ability to cope with the economic threat of natural disasters, and suggest that policies focused on alleviating these types of frictions may be particularly beneficial as global warming continues. Finally, I build on a long literature examining structural transformation in development economics. The transition of the labor force out of agriculture has long been considered a central step in long-run economic growth (e.g. Kuznets (1957)). Canonical models of this process focus on the *push* of positive agricultural productivity shocks and the *pull* from positive manufacturing shocks, largely motivated by England's experience during the industrial revolution.[8] Recent empirical evidence has highlighted the important role of factor biased technical change Bustos et al. (2016), openness to trade Moscona (2019), and financial frictions Albert et al. (2023) in shaping the consequences of technological change in agriculture. I build on this work by examining the local legacy of sharp negative agricultural productivity shocks typical of natural disasters associated with climate change to illustrate scope for such catastrophic events to shift labor across sectors.

2 Limitations of Existing Flood Databases

Evaluating the economic consequences of flooding presents a significant challenge because of systematic biases in existing datasets. In this section, I discuss these issues and present evidence of their relevance.

[6]For a few examples from this long literature, see Hsiang and Jina (2014), Elliott et al. (2015), and Pelli et al. (2023) on the impact of cyclones, Deschênes and Greenstone (2007), Guiteras (2009), Dell et al. (2012), Burke et al. (2015), Burgess et al. (2017), Shah and Steinberg (2017), Bharadwaj and Mullins (2021), and Somanathan et al. (2021) on temperature and precipitation, Hornbeck and Naidu (2014), Desmet et al. (2018), Kocornik-Mina et al. (2020), Gandhi et al. (2022), and Balboni et al. (2023) on floods, and Jayachandran (2009), Greenstone and Hanna (2014), Arceo-Gomez and Oliva (2014), Hanna and Oliva (2015), Pande et al. (2015), Barwick et al. (2018), Chang et al. (2019), Adhvaryu et al. (2019), and Rangel and Vogl (2019) on air quality.

[7]For a small section of this growing body of work, see, among others, Deschênes and Greenstone (2011), Boustan et al. (2012), Barreca et al. (2016), Burke and Emerick (2016), Balboni (2019), Barwick et al. (2019), Greenstone et al. (2022), Carleton et al. (2022), Lane (2022), Hornbeck (2023), Aker and Jack (2023), Emerick et al. (2016), Ito and Zhang (2020), McGuirk and Nunn (2020), Khanna et al. (2021), Taylor and Druckenmiller (2022), Berkouwer and Dean (2022), Moscona (2022), Moscona and Sastry (2023), and Hsiao (2023). For a recent review, see Kala et al. (2023).

[8]See Nurkse (1953), Lewis (1954), Rostow (1960), Baumol (1967), Murphy et al. (1989), Kongsamut et al. (2001), Gollin et al. (2002), Ngai and Pissarides (2007), and Gollin et al. (2021).

2.1 Inconsistent Definitions of "Flood"

What is a flood? According to the Intergovernmental Panel on Climate Change (IPCC, 2022), "floods are the inundation of normally dry land". A dataset of floods therefore requires objective standards of both "inundation" and "normally" to ensure consistent measurement across places and time. Given the substantial seasonal shifts in rivers and other bodies of water across much of the world—especially in developing countries—normally dry areas at some points of the year may be regularly submerged during other times, and these patterns may shift across years.[9] Measuring floods critically requires accounting for this permanent water consistently across time and space.

I conduct three empirical exercises to illustrate the consequences of ambiguous definitions of flooding in existing databases.[10] I begin by comparing the Dartmouth Flood Observatory Archive (Brakenridge, 2023) to the EM-DAT International Disaster Database (Guha-Sapir et al., 2023)—the two most widely used flood sources, both of which rely on newspaper articles and government reports to document inundation events.[11] Focusing on all flooding events since 2000, I calculate the number of flooding events in each dataset at the country-by-year level. Of the 2,086 country-years for which at least one of these two databases records a flood, they agree on the total number of flooding events just 26.61 percent of the time, despite relying on the same source material.[12] To characterize this gap another way, the DFO reports 814 more flooding events at the country-year level than does EM-DAT, while EM-DAT lists 764 floods missing from the DFO archive.

I next show suggestive evidence of inconsistent measurement not only *across* databases but *within* datasets over time. I calculate the flooding event with the smallest area for each calendar year using the DFO archive. The smallest reported flood ranges from 11.96 square kilometers in the Czech Republic in 1998 to 4,161.40 square kilometers in the United States in 2016. It seems implausible that across all floods in the world, the true variation across

[9] To illustrate the importance of these shifts in where water is supposed to be, I map the extent of water bodies that exist all year in blue against those that exist for just part of the year in red for Bangladesh in Appendix Figure A.2, the underlying data of which I take from the Global Surface Water dataset (Pekel et al., 2016).

[10] Newspapers and governments do not use a common standard when declaring a flood. For example, the Flood Forecasting and Warning Centre of the Bangladesh Water Development Board bases its main classification on river heights crossing a "danger-level". Meanwhile, the Federal Emergency Managemeny Agency of the United States (relying on the National Flood Insurance Program definition) considers a flood to be "a general and temporary condition of partial or complete inundation of 2 or more acres of normally dry land area or of 2 or more properties."

[11] For examples of papers in economics using one of these two sources, see, among others, Mueller et al. (2014), Kocornik-Mina et al. (2020), Jia et al. (2022), Gandhi et al. (2022), and Balboni et al. (2023).

[12] Figure 1 plots these discrepancies, comparing the number of flooding events at the country-by-year level in each database. Points have been jittered to visualize the mass at each comparison, and red dots signify when the two datasets agree while blue denotes disagreement.

years in the smallest inundation event spans such a large range.

Finally, I compare inundation events as reported by a government to those in the DFO and EM-DAT databases. I focus on Bangladesh, which has made their flooding documents public such that they theoretically fall under the sample of input sources. I digitize information from the annual reports between 2010 and 2020 on fluvial flooding events, for which the government collects the most accurate data.[13] I match these records to river stations from the Bangladesh Water Development Board (BWDB), thereby constructing a complete record of where the government designated floods based on river water heights. During this same period between 2010 and 2020, the DFO and EM-DAT databases include 18 and 16 flooding events, respectively. I find these sources understate the number of floods reported by the Bangladeshi government by a factor of three to 11, depending on how I classify the water-level records.[14]

Guiteras et al. (2015) provides a notable exception to most prior work in also letting satellite data inform them about the location and timing of floods instead of relying on an external source.[15] This addresses the issue of inconsistent definitions because they are able to simply define a flood based purely off of the remote sensing data itself. Their approach differs from mine however in the remote sensing data series used—in particular, their paper was published just as the Sentinel-1 satellite launched, so they could not use that radar-based data source for their estimates. Instead, they had to rely exclusively on the MODIS instrument, an optical sensor vulnerable to bias due to cloud coverage. Because of this limitation, their series uses 16-day composite MODIS images constructed from cloud-free observations. Their resulting measure of surface water relies on the relative values of constructed indices of surface vegetation and surface water. By incorporating remote sensing data unavailable to them, my approach builds on their paper in several important ways. First, I rely on the radar-based measurement as primary method to detect surface water, which yields far more accurate measurements by relying on radiation backscatter as opposed to constructed indices from spectral bands. Second, the 16-day composite underlying their data selects the best available pixel value from all MODIS passes over the period. This

[13]For example, one entry reads: "The Dharla at Kruigram registered several peaks during the monsoon 2010. It crossed its danger level (DL) on 20th July and continues to flow above DL till 12:00 hours of 24th July, 2010 (4 days). It attained its highest level 26.83m on 22nd July at 18:00 hours, which was 33cm above the DL (26.50 m)."

[14]Comparing the DFO and EM-DAT events to those in the BWDB data requires taking a stance on the definition of a flooding "event". In the end—regardless of the aggregation—the takeaway remains the same. For instance, if I define unique flooding events at the division-by-month level, the BWDB classifies 179 distinct events. Ignoring the country's seven divisions and counting unique months gives 58 floods, while ignoring months and just looking at different divisions annually yields 62 inundation events.

[15]A separate approach has been to use deviations in rainfall as a measure of flooding, as in Felbermayr and Groschl (2014). This requires strong assumptions about the link between local rainfall and local flooding.

likely creates a bias against detecting floods, which may be more likely to occur on cloudy days featuring worse pixel quality. This could be one reason why they document such a low correlation between rainfall and flooding in their data. Third, my measure estimates floods every day, capturing flash floods which farmers report being economically relevant yet will likely be often omitted from 16-day analyses. This could help to explain why in their paper, they document a satellite-derived flood rate of just 11.8 percent among the farmers' who self-report experiencing a flooding experience.

2.2 "Missing" Floods in Poor Areas

The reliance on news articles and government reports—the main sources for databases like the Dartmouth Flood Observatory Archive (DFO) and EM-DAT International Disaster Database—generates significant underreporting of flooding events in poorer and more remote parts of the world.[16]

To illustrate the extent of "missing" floods, I examine how a country's income relates to the number of floods in existing databases. Given improved infrastructure in richer nations, one might expect that the true relationship between flooding and income is negative: with worse dams and river embankments, poor countries may be more likely to experience floods.[17] I link the total number of floods since 2000 in the DFO and EM-DAT datasets to information on total land area and per capita income in purchasing power parity terms from the World Bank's World Development Indicators (World Bank, 2023). I then estimate Equation 1 separately for each dataset at the country level, where $f(Latitude, Longitude)$ is a quadratic polynomial of the coordinates of each country's centroid.[18]

$$\frac{\sum Floods_i}{Area_i} = \alpha + \beta \log(Income_i) + \Gamma f(Latitude_i, Longitude_i) + \varepsilon_i \tag{1}$$

[16]Existing databases acknowledge varying data quality across places yet nevertheless have been widely used to study natural disasters in developing countries. For instance, the Dartmouth Flood Observatory Archive describes its dataset as follows: "The statistics presented in the Dartmouth Flood Observatory Global Archive of Large Flood Events are derived from a wide variety of news and governmental sources. The quality and quantity of information available about a particular flood is not always in proportion to its actual magnitude, and the intensity of news coverage varies from nation to nation. In general, news from floods in low-tech countries tend to arrive later and be less detailed than information from 'first world' countries."

[17]Of course, poor countries differ from rich ones in many important ways—including geographic—which might also impact flooding. To partly address this concern, I adjust for the total land mass of the country and control for a flexible function of each country's centroid. Nevertheless, this exercise is merely suggestive.

[18]I report heteroskedasticity-robust standard errors, and limit the sample to those countries where either the DFO or EM-DAT databases report at least one flood since 2000. If poor countries are more likely to have zero floods reported even when they did experience at least one, then this sample restriction would bias the results towards finding no bias against poor countries.

Contrary to intuition, the analysis shows that existing databases actually report *more* floods in richer countries than in poorer ones. Figure 2 presents these results in binned scatter plots, along with the coefficient β from equation 1 and the associated standard error. The DFO data exhibit a strong positive relationship between income and reported floods (*p-value* =0.012), while the EM-DAT data show a positive link of similar magnitude though below standard thresholds of statistical significance (*p-value* =0.244). These results suggest "missing" floods from these databases, consistent with sparse news coverage of flooding in poorer places. One might expect a similar bias within countries.

Instead of traditional news reports, alternative sources such as social media could be used and may be less susceptible to this systematic undercoverage Lai et al. (2022). As one test of the feasibility of this kind of approach, I construct a dataset of every publicly available, geo-coded tweet in Bangladesh between 2016 and 2022 using the Harvard Center for Geographic Analysis Geotweet Archive (Lewis and Kakkar, 2016).[19] The final sample includes 2,693,031 tweets that I then classify as having mentioned a flood based on a keyword match for "flood" or "bonna" (the Bangla word for flood). Throughout this entire period, just 829 tweets mention floods. Figure A.3a shows that the number of tweets increases over time, potentially leading to inconsistent results if, for instance, people in rural areas began using Twitter later. Flood tweets occur on 394 distinct dates in 104 unique unions (out of 5,158 total)—however, the vast majority of flood tweets occur in Dhaka. Figure A.3b maps this spatial distribution, indicating just how much of the country remains uncovered by this data source, particularly in more remote and poorer areas. Even non-traditional sources that nevertheless rely on human reporting cannot accurately identify flooding across space and time.

2.3 Focus on Large Floods

Existing flood datasets like the DFO, EM-DAT, and UNOSAT records focus primarily on large-scale inundation events. For instance, among all flooding events in the DFO database since 2000, the number of days between the recorded begin date and end date is one day or less 7.45 percent of the time and three days or less in 24.35 percent of events. Does the omission of smaller, shorter term floods impact our understanding of the economic consequences of floods? The question is difficult to answer because armed solely with existing data, one cannot measure either the magnitude of missing flash floods or accurately assess their

[19]The database includes all tweets which feature either (1) GPS-based longitude/latitude generated by the originating device or (2) place-name-centroid-based longitude and latitude based on the user-defined place designation run through Twitter's bounding box. Approximately one to two percent of all tweets fall under this category.

consequences. If small-scale floods occur rarely and have little impact, then this feature of existing datasets causes less concern.

To empirically examine whether omitting small floods matters, I simply ask vulnerable households directly. I survey 2,021 farmers across 250 villages in the Khulna division of Bangladesh—a coastal region on the front lines of climate change.[20] Farmers perceive a much greater risk of flash floods than longer lasting flooding events.[21] Not only do farmers expect flash floods to occur frequently, but they also incur significant economic costs when they happen.[22] Although longer-duration floods tend to cause more damage, shorter-term events nevertheless can result in significant harm to both farmers' property and their crops.[23] A database accurately capturing all flooding events relevant for people's lives should therefore include local, short-term floods.

2.4 Modeling Uncertainty of *in situ*-Based Estimates

In part due to the challenges involved with using news articles and government reports to measure floods, direct measurement of water has emerged as an appealing alternative source. This *in situ* data typically comes from stations capturing water height along rivers and coasts. This section presents two important limitations to estimating flooding from these measurements.

[20]From the 642 unions in the Global Administrative Areas (2018) data, I excluded 32 urban areas and then selected nine unions from which I have government salinity station data, 37 unions with water stations from the Bangladesh Water Development Board, 48 unions that are also included in the Bangladesh Integrated Household Survey sampling frame, and 121 unions that are also included in the 2016-2017 Bangladesh Labor Force Participation Survey sampling frame. This yields 185 unique unions. Both the Bangladesh Integrated Household Survey and the Bangladesh Labor Force Participation Survey were designed to be representative, and just 29 of the 185 initially selected unions fall outside of both of those survey's sampling frames as exclusively part of the government salinity or water stations. Then, I randomly sort the remaining unions, and choose the next 65. Enumerators visited each union and did an initial listing of 50 households who were planning on harvesting rice during the upcoming Boro season and made the primary agricultural decisions on their land. In almost all unions, this goal of 50 households was achieved and typically within a single village. From this initial list, farmers were randomly ordered to be selected for an interview. Initially, 10 households were selected per union, though this number was revised down to nine given survey length concerns after the first week. On average, 9.1 farmers were surveyed in each union. Before a household was deemed unavailable and a replacement household was selected from the randomized listing order, enumerators attempted to contact them multiple times over multiple days via their phone number collected during the listing. A random half of farmers answered the questions about flood damages discussed above in fall 2022; the other half answered them in summer 2023.

[21]I ask, "How many years do you think it would take for a [TIME] long flood to happen in this village?", where [TIME] is one-day, three-day, week, and month. Figure A.4 plots the distribution of responses, showing that farmers expect short-term floods far more frequently than longer ones.

[22]I ask farmers to recount every flooding event they have experienced, along with its duration and damage to both housing and crops. Table 1 presents these results.

[23]Of course, farmers may disproportionately recall the most severe floods, yet even under this selection, this pattern clearly demonstrates that even flash floods can be economically meaningful.

The conversion of the *in situ* vertical measurement of water to a *horizontal* indicator of flooding presents a major challenge. Researchers have thus far relied on hydrological models requiring strong assumptions about items such as elevation, river flow, and soil type. A rich body of work illustrates the sensitivity of flood estimates to modeling parameters.[24] The calibration-validation process used to inform these models can itself by biased against poor countries, where both data and the volume of research lag far behind levels found in richer places.

Second, the sparsity of *in situ* data—especially in poor countries—further exacerbates these reliability issues. Take, for instance, the Khulna division in Bangladesh. Home to 17.5 million people and accounting for 15% of Bangladesh's total land, Khulna features just 47 places where *in situ* data are regularly collected—and it is the division where river height measurements are perhaps most densely collected.[25] Estimating flooding in places far away from monitoring stations relies on strong modeling assumptions that limit the reliability of the estimates. Further bias in measurement occurs when monitoring sites are located in more urban and richer areas.

3 Measuring Floods

To avoid the pitfalls of existing approaches, my method uses globally available satellite data to detect flooding events. To illustrate this procedure, I apply the method to estimate flooding inundation in Bangladesh for each of the country's 5,158 unions. Consider a set of polygons $\mathcal{I}$, indexed by i (unions in the case of Bangladesh).[26] Let $S_{i,t}$ denote the share of surface water in polygon i on day t.[27] This value can be decomposed into the sum of permanent water $P_{i,t}$ and floods $F_{i,t}$, the object of interest. To estimate this value, I take four main steps. See Appendix Section A for details.[28] Thanks to the petabytes of data

[24] See, for instance, Werner (2004), Gaume and Borga (2008), Lumbroso and Gaume (2012), Komi et al. (2017), and Khojeh et al. (2022).

[25] Figure A.5 maps the rivers in this area in blue and the government's river monitoring stations that collect water level data with red stars.

[26] Analyzing the remote sensing data can be quite demanding, depending on the size and number of polygons. In general, having more polygons improves the accuracy of the prediction function by increasing the training sample size used by the supervised learning algorithm, and smaller polygons improves the computational speed involved with processing the remote sensing data per polygon.

[27] The estimand of interest in this procedure is the amount of flooding in a given polygon. Some researchers or policymakers may instead be interested in mapping out the precise flooding boundaries. This can be easily incorporated into my approach using one of many existing remote sensing techniques that delineate inundation extents *conditional* on knowing a flood occurred after first using my approach to identify the existence of a flood.

[28] To illustrate these steps, I show examples from the Holokhana union in Bangladesh (pictured in Appendix Figure A.7), which experienced extensive flooding in July 2020. I present visualizations of the satellite data for this union both for a week during these floods and the same calendar week in 2021 for comparison.

distributed publicly by Google Earth Engine, all remote sensing data necessary for this approach is available for the entire globe.

Step #1 I begin by estimating surface water $S_{i,t}$ using extremely accurate yet infrequent radar-based satellite data. Denote the radar-based measures by $A_{i,t}$. Synthetic aperture radar (SAR) images constitute an appealing data source for detecting flooding because they can "see through clouds" that otherwise can obstruct optical sensors. SAR instruments operate both day and night and during all weather conditions which make them particularly well suited to the inclement weather that typically coincides with flooding.[29] On days when the radar data are available ($A_{i,t} \neq \emptyset$), I assume this measure provides an unbiased estimate of true surface water $S_{i,t}$ according to equation 2 where $\varepsilon_{i,t}$ is a mean-zero error term $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^{\varepsilon})$.

$$\underbrace{S_{i,t}}_{\substack{\text{Surface} \\ \text{Water}}} = \underbrace{F_{i,t}}_{\text{Flooding}} + \underbrace{P_{i,t}}_{\substack{\text{Permanent} \\ \text{Water}}} = \underbrace{A_{i,t} + \varepsilon_{i,t}}_{\substack{\text{Radar-Based} \\ \text{Surface Water}}} \quad \forall \quad A_{i,t} \neq \emptyset \tag{2}$$

To prepare and analyze the SAR data, I implement techniques from the methodological frontier of geophysics, described in detail in Appendix Section A.1. From a high-level, I use methods from computer vision to distinguish surface water from the raw radar backscatter data collected by the satellite, taking steps to ensure consistent and accurate measures by holding constant the exact orbital position of the satellite and applying filters for noise, among others.

Step #2 I next fill in gaps in the time series of surface water from Step 1. See Appendix Section A.2 for details on this step. I begin by constructing a second set of remote sensing data which I denote by the vector $\mathbf{X}_{i,t}$. Components of the vector must satisfy two requirements. First, they must exist at a sufficiently high-frequency data such that they overlap both with the ground truth measure $A_{i,t}$ and all other periods t of interest. Second, they must plausibly predict flooding. I build two complementary data sources based on remote sensing data that satisfy these requirements, providing daily coverage since July 4^{th}, 2002. The first set of inputs comes from optical images capturing data on spectral bands across seven wavelengths from satellites passing over the earth every day. I prepare these data using methods from the forefront of geophysics and then construct a high-dimensional vector of

[29] These instruments can detect surface water at a fine geographic resolution even amid the cloud cover that thwarts optical instruments. However, the satellites equipped with these tools orbit infrequently, potentially missing many short- and medium-length floods. Perhaps due to this limitation, researchers using the radar instruments have thus far combined them with news databases or other flooding reports to improve the precision of inundation extent boundaries for a single flooding event, thereby incorporating the same biases as the original sources.

predictors from each image. Using these data alone to accurately predict flooding is not possible due to the correlation of cloud cover—which thwarts these sensors' ability to detect what's occurring on the ground—and flooding events.[30] The idea behind my approach is that even though clouds may distort the data, this distortion itself can still contain helpful signal—it just needs to be properly extracted, which I do in a data-driven manner using machine learning. The second set of inputs to $\mathbf{X}_{i,t}$ comes from rainfall data based on models combining remote sensing and weather gauge data to estimate precipitation worldwide every 30 minutes with data available since June 1^{st}, 2000.

Equipped with this data, success of the method hinges on the ability to predict the accurate surface water measurement $A_{i,t}$ using the high-frequency predictors $\mathbf{X}_{i,t}$ constructed from these data sources. This task is particularly well-suited to supervised machine learning. I estimate a prediction model that takes as an input the vector $\mathbf{X}_{i,t}$ and produces an estimate of the accurate flood inundation on days without missing data, which I denote by $\widehat{A}_{i,t}$. Let $g(\cdot)$ denote this function, which yields the relationship in equation 3 where $\mu_{i,t}$ is a mean-zero error term $\mu_{i,t} \sim \mathcal{N}(0, \sigma^{\mu})$.

$$A_{i,t} = \widehat{A}_{i,t} + \mu_{i,t} = g(\mathbf{X}_{i,t}) + \mu_{i,t} \tag{3}$$

I find that a histogram gradient boosted regression tree fits best among several potential algorithms, and that overall the model performs extremely well, explaining about 90% of the variation in the radar-based measure. Appendix Section A.2 provides details benchmarking this performance.

By applying this same model out of sample, I can use the high-frequency inputs $\mathbf{X}_{i,t}$ to predict $\widehat{A}_{i,t}$ on days when the radar sensor did not collect data, resulting in a vector of surface water for each polygon i, which relates to the true surface water $S_{i,t}$ according to equation 4.[31]

$$S_{i,t} = g(\mathbf{X}_{i,t}) + \mu_{i,t} + \varepsilon_{i,t} \tag{4}$$

[30] To examine the scope for bias in using optical data alone, I follow the frontier methods for analyzing MODIS data from Tellman et al. (2021) to detect surface water and overlap this measure with "permanent" water as measured by the Global Surface Water database relying on LANDSAT images (Pekel et al., 2016). Given that this permanent water exists year-round, the analysis using the optical MODIS data should detect water on all of the same pixels. By contrast, I find that in Bangladesh, the optical methods can only detect water on 3.28 percent of these permanent pixels on average for unions with any permanent water. This accuracy exhibits strong seasonal variation, however, falling to less than one percent during the summer monsoon season, precisely when floods occur most frequently. For methodological details and a graph of this pattern, see Appendix Figure A.1.

[31] Potential differential selection of the training sample does not raise concerns in this setting because the precise timing of the orbiting radar satellites used in this study do not relate to conditions on the ground.

Step #3 To create a database of flooding, I must remove normal water inundation from this overall surface water measure to isolate unexpected deviations. This "permanent" water relies critically on removing seasonal fluctuations in bodies of water, a force which poses particular relevance in countries featuring monsoons or water run-off from mountain ranges, both of which generate enormous variation in inundation across the calendar year.[32] To separate "permanent" water from flooding, I residualize $S_{i,t}$ on a flexible function of seasonality $f(i,t)$. The core assumption requires the seasonal value to be an unbiased predictor of permanent water $P_{i,t}$. The rich time series of daily surface water across two decades allows me to remove even subtle shifts in bodies of water that emerge over time. I estimate equation 5 separately for each polygon, where $\mathbf{W}_{i^*,t} \times \mathbf{H}_{i^*,t}$ denotes calendar-week by half-decade fixed effects for polygon-of-interest i^*, $\mathbf{W}_{i^*,t} \times Y_{i^*,t}$ allows for linear trends in calendar-weeks, and $\mathbf{Y}_{i^*,t}$ captures year fixed effects.[33]

$$S_{i^*,t} = F_{i^*,t} + P_{i^*,t} = F_{i^*,t} + f(i^*,t) + \nu_{i^*,t} = F_{i^*,t} + \mathbf{W}_{i^*,t} \times \mathbf{H}_{i^*,t} + \mathbf{W}_{i^*,t} \times Y_{i^*,t} + \mathbf{Y}_{i^*,t} + \nu_{i^*,t} \quad \forall i^* \in \mathcal{I} \tag{5}$$

In other words, for each observation, I remove the average amount of surface water in that polygon for that same calendar week in the same half-decade along with any linear trends occurring during that calendar week year-to-year and annual shocks to bodies of water that impact every day in that year equally. This approach removes seasonal variation that itself I allow to vary over time (perhaps due to the impacts of global warming, for instance) while also accounting for sudden local shifts in surface water (such as the construction of a new dam that creates a new lake in a polygon). The resulting residuals $(F_{i,t} + \nu_{i,t})$ can be interpreted as the percentage point deviation from the expected amount of surface water in that polygon on that day. See Appendix Section A.3 for further discussion of this step.

Step #4 To convert the resulting continuous measure of surface water deviation $\widehat{F_{i,t}}$ into a definition of flooding consistent with most notions of the term, I incorporate survey data featuring recall of recent floods across 250 different areas. By calibrating a threshold τ to match these recounted inundation events to define a flood as days with $\widehat{F_{i,t}} > \tau$, I classify the satellite flooding residual using an objective standard that nevertheless captures on-the-

[32] As an illustrative example of this dynamic, figure A.11 plots how the river height changes for a single sample river station over the course of 2022 based on data I obtained from the Bangladesh Water Development Board. Shifts due to changing tides and monsoon rains can be clearly seen in the time series, resulting in more than a tripling of river height over the course of the year. Climate change has likely exacerbated this problem by shifting these seasonal patterns over time.

[33] Based on visual assessments of the residual plots, this approach removes seasonality more effectively than the Baxter-King and Hodrick-Prescott filters commonly-used in time series econometrics.

ground experiences.[34]

Converting this daily surface water deviation into a binary indicator for flooding features three important advantages. First, by construction, the residuals form a nearly continuous support up to the largest flooding inundation area. Very small deviations (which could also simply reflect noise inherent in the estimating procedure) should likely not be considered flooding, while huge deviations certainly would be. However, the continuous measure on its own makes this definition quite fuzzy. Second—for the purposes of scientific communication—researchers and policymakers may find communicating a "27 percentage point increase in surface water" significantly more difficult than simply "a flood occurred". Third, the common usage of the word "flood" carries an important meaning such that someone might "know it when they see (or experience) it." Ideally, the dataset would reflect that same definition.

To address these concerns, I take a data-driven approach using survey data I collect from 644 farmers across 149 unions in the Khulna division of Bangladesh.[35] I asked each farmer to recount every flooding event they have experienced, when it occurred, and its duration. I use their responses to pick a threshold above which I define surface water deviations as "flooding." Denote this threshold cut-off by τ. To address potential memory issues (Guiteras et al., 2015), I restrict to the time period since April 2022 under the assumption that farmers can perfectly recall experiencing inundation events within the past eight months.[36] I classify a union-month as a flooding month if *any* farmer in that location recalls a flooding event, and I calculate the maximum number of days of inundation reported.

To estimate τ, I choose a threshold of surface water deviations to match the pattern of flooding events recalled by farmers. After implementing the flooding estimation procedure at the union-level for Bangladesh, for any given τ, I calculate the number of flooded days in any union during any month. Comparing this with the farmers' answers generates three complementary statistics: the share of union-months of which I perfectly estimate the number of flood days, the share of which I underestimate, and the share of which I overestimate. Figure 3 presents the results of this exercise for different thresholds. A cut-off of approximately .2 appears to work best in this setting for defining a flood—equivalent to a

[34]Instead of predicting a continuous measure of surface water, an alternative approach involves first defining flooding events directly from the continuous measure of the radar-based data and then predicting that binary outcome, effectively moving step 3 prior to step 2. My preferred procedure predicts the continuous outcome because with the much longer and daily panel resulting from step 2, the function $f(i,t)$ can be substantially more flexible than with the far fewer observations available after step 1. However, I report results from this alternative approach when relevant in the Appendix.

[35]This is the same sample from the discussion of flash floods in section 2, though due to an error in the tablet survey elicitation, a smaller portion of farmers answered questions about flood recall accurately.

[36]Indeed, the distribution of years of recalled floods—shown in Figure A.12—matches a recency bias pattern.

20 percentage point increase in surface water. At the polygon-month level, this threshold perfectly captures essentially all flood days. In the absence of equivalent calibration data for other settings, this benchmark will likely perform well, particularly for similarly sized polygons.[37]

Validation To help assess the quality of this measure of flooding, I compare my estimates from Bangladesh to four complementary other sources: (1) basic contextual knowledge from the region, (2) a new database of flooding I construct from newspaper articles, (3) annual reports on flooding from the government, and (4) *in situ* water-level measurements from water stations across the country. Appendix Section A.4 provides details on these tests, showing consistent evidence that my new measure accurately captures floods. The overall patterns of inundation in my data pass intuitive geographic and temporal checks, aligning with major, and well-documented flooding events. Event study specifications show that the surface water I estimate spikes at the same moments that newspapers report flooding. Government reports of inundation as well as river height data both closely track my satellite and machine-learning derived estimates as well.

4 Estimating the Impact of Floods

Equipped with this flooding database, I turn to the main question of interest: how do floods impact economic development? This section describes my econometric approach to estimate the causal effect of experiencing a flooded state in period t on outcome Y in periods $t+\tau$, where τ traces out the impulse response function. Let i index units and $Flood_{i,t}$ equal 1 if i experienced a flood in period t and 0 otherwise. Let $Y_{i0,t+\tau}$ denote the outcome for unit i in period $t+\tau$ without experiencing a flood in period t, and $Y_{i1,t+\tau}$ denote the same for experiencing a flood (noting that only one of these outcomes is ultimately realized). I denote the flood history for p years prior to period t by the vector $\mathbf{Flood}_{i,t-1}^{p} = \{Flood_{i,t-p}, Flood_{i,t-p+1}, \ldots, Flood_{i,t-1}\}$. For ease of exposition, consider the case of defining treatment annually (experiencing a flood at any time during the year), such that t indexes years.

To estimate this object, I adopt a stacked difference-in-differences specification to combine flooding events occurring at different times in an event study. This overall empirical strategy rests on two assumptions: first, no anticipation—that is, that units do not respond

[37]To partially account for noise around this threshold τ, I take a final processing step to fill in "missing" floods using gaps in the time series. I assume that any period of up to five days for which a polygon exhibits residual surface water above this threshold τ both before and after the gap reflects measurement error, and change the binary status of these days to match floods.

in anticipation of future flooding status—and second, parallel trends, whereby in the absence of treatment, the outcomes of flooded units would have evolved similarly to those of unflooded ones. Both of these assumptions seem plausible given the quasi-random nature of floods conditional on place, following the rich literature on climate shocks (Dell et al., 2014). My approach follows in the spirit of local projections techniques from time series (Cengiz et al., 2019; Dube et al., 2023). In doing so, I address the concerns about negative weights amid heterogeneous treatment effects recently raised in the literature while maintaining flexibility for arbitrary weighting schemes and controls.[38]

I seek to identify the average treatment effect on the treated (ATT) of experiencing a flood on outcomes in year τ given by equation 6 where w_t denotes the weights on each cohort t. In the primary specification I use, these weights on cohort-specific treatment effects are always positive and proportional to group size and treatment variance. This object captures the impact of experiencing a flood among all flooded units relative to those same units not experiencing a flood in period t. This estimand includes flooded units that experience subsequent floods at any arbitrary time after the initial flooded year y, allowing for the possibility that flooding in one period increases the likelihood of future flooding (for instance, by damaging a river bank) as part of the estimated treatment effect.[39]

$$ATT^{\tau} = \sum_{t} w_t \left(\mathbb{E}[Y_{i1,t+\tau} - Y_{i0,t+\tau} | Flood_{i,t} = 1] \right) \tag{6}$$

This object includes treatment effects for units that regularly experience flooding—in other words, this object includes the marginal impact of an additional flood irrespective of past inundation exposure.

Constructing the Control Group The core intuition underpinning this identification strategy involves comparing flooded units with a set of "clean" controls and then aggregating these effects together. For each treatment year y, I begin with the sample of all units such that $Flood_{i,y} = 0$. I make several further restrictions on this sample to construct a control group.

First, mis-classifying treated units as control units presents one particularly important concern in this context because my definition of treatment stems from an underlying con-

[38] In fact, many recent estimators proposed in the literature are either specific sub-cases of this general approach or very closely related to one simply by choosing specific combinations of weights or base periods for the local projection.

[39] The empirical evidence suggests this channel might play an important role: conditional on having a flood in a year prior to 2022, the probability of experiencing a flood the following year is 36.48 percent, compared to a base rate of 21.28 percent overall in the sample of unions with at least one flood.

tinuous measure of surface water deviation.[40] As a result, noise in the satellite imagery may result in measurement error in the treatment variable. To protect against this concern, I only consider control units whose maximum surface water residual in the treatment year y never exceeds 10 percentage points (compared to the flooding definition of 20 percentage points).

Second, geographic spillovers pose a potential threat to interpreting treatment effects. In my default specification, I therefore additionally exclude any control unit that falls within 10 kilometers of any treated unit.[41]

Third, I require the vector of past flood history $\mathbf{Flood}_{i,y-1}^{pre}$ to be identical between treatment and control units, where *pre* denotes the period length after which I assume no dynamic treatment effects. If this assumption does not hold, my results will be biased (and likely I would observe a pre-trend given the spatial concentration of flooding). For each pattern of past flood exposure in the flooded group, I create a stratum to which I add control units with the same history. I exclude any treated units for whom no perfect match exists, and estimate all treatment effects within these strata. This ensures that lingering impacts from previous flooding exposure do not influence the results. In my default specification, I set *pre* equal to six years such that I have five years of identical flood history before treatment and control units diverge.

Fourth, I restrict to control units with the same expected flood risk as treated units in their strata. While matching on previous flood history $\mathbf{Flood}_{i,y-1}^{pre}$ does help considerably with this issue, a Bayesian would nevertheless note that because the treated units by construction experienced a flood in year y, they likely faced a higher *ex ante* chance of flooding relative to their control counterparts despite having identical past inundation exposure. To address this concern, I use a supervised machine learning algorithm to develop a measure of flood risk by predicting flooding experience using a host of underlying geographic features, including existing hydrological models of inundation likelihood.[42] Within strata, I further

[40] Appendix Figure A.23 visualizes this issue by highlighting the mass around the discontinuity of the flooding definition threshold.

[41] I calculate these distances based on union centroids. On average, each union features 37 other unions that fall within 10 kilometers. As robustness checks, I additionally consider 0, 20 and 30 kilometers, the latter two of which have an average number of spillover unions of 76 and 142 unions, respectively.

[42] To estimate flood risk, I first calculate the daily flood hazard rate based on the full panel of flooding experience. To account for the long tail of the distribution, I assign each union its percentile rank in this distribution of hazard rates, though results are broadly similar using the raw value. This variable constitutes the main flood risk outcome that I predict in the algorithm. As inputs, I calculate the mean and standard deviation of elevation in each union, binned latitude and longitude at the tenth of a degree level, the length of major rivers through that union, and the government's estimates of flood hazard, proneness, average drainage characteristics, and flooding depth from Bangladesh Agricultural Research Council (BARC) shapefiles. I then randomly split the sample of 5,158 unions into a training dataset and a testing dataset, with 20 percent reserved for the latter. Finally, I train a random forest algorithm to predict flood rank using the full set of

exclude any control unit with a predicted flood risk outside of the support of the treatment group's predicted risk.

Finally, when possible, I identify control units with the same pre-treatment trends in the outcome Y.[43] Like as in many similar designs, differential trends between flooded and non-flooded areas present an important potential threat to identification. For example, if coastal areas simultaneously exhibit differential exposure to floods in one year and differential economic growth (perhaps due to trade, for instance), then a flooding specification comparing coastal areas to non-coastal areas might yield biased point estimates. Because control units might be systematically different than flooded areas even after conditioning on flood risk, one would prefer to select from among this potential pool of controls those units most similar to those places that ultimately experienced flooding. To systematically create a comparable control group, I incorporate synthetic control weights into this difference-in-differences design. This is particularly appealing by removing researcher degrees-of-freedom in estimating pre-trends (such as adjusting covariates, choosing time periods, or in this case, selectively choosing among a large number of control units), thereby addressing the pretesting concerns noted by Roth (2022). Arkhangelsky et al. (2021) show that relative to traditional approaches, incorporating synthetic control often performs better and has appealing robustness properties. I calculate weights on units such that the average outcome for the treated units is approximately parallel to the weighted average for control units. Unlike in traditional synthetic control (Abadie et al., 2010), the difference-in-differences design allows me to choose weights on units that allow treatment and control to vary up to a constant.[44] I also report specifications without these synthetic control weights. In these specifications—in order to recover the ATT in equation 6—I weight based on their propensity

inputs. Applying this model out of sample to the hold-out unions, I can explain 0.71 percent of the variation in true historical incidence ranking using my predicted measure. The fact that this R^2 is less than one can be viewed as an advantage in this case because I already match on previous flood history $\mathbf{Flood}^{pre}_{i,y-1}$. This measure, by contrast, captures a data-driven notion of similarity.

[43] I only estimate these weights in the case of true panel data in which I observe repeated outcomes for the same unit of observation. In the case of a repeated cross-section of individual or household data, I do not estimate the synthetic control weights.

[44] I assign treated units weights equal to the inverse of the number of control units (the initial weights of which sum to 1), and then scale all of these weights such that they sum to the total number of observations. To satisfy the requirements of Arkhangelsky et al. (2021), I must restrict to a balanced panel and to events featuring flooding occurring after the first period with observed outcomes when including the synthetic control weights. Departing from Arkhangelsky et al. (2021), in my preferred specification I do not weight time periods to match pre-treatment and post-treatment outcomes in the control group. This is for two reasons. First, I am interested in the entire impulse response function in the wake of a flood instead of just the cumulative effect. Second, given the quasi-random timing of floods conditional on place, imbalanced timing effects seem particularly unlikely in this setting. However, in the appendix, I also report the synthetic difference-in-differences results with these timing weights for each calendar year y. In those regressions, I calculate standard errors via bootstrapping with 500 draws.

scores, which I estimate using a logistic regression of a host of geographic variables and all floods up to and including the treatment year.[45]

After having constructed these flooded and control units, my approach involves stacking flooding events together in relative time indexed by t instead of calendar time, which avoids the issues with heterogeneous dynamic treatment effects. I set $t = 0$ to denote the year when a flood occurs. With this sample, I estimate equation 7 (shown below for horizons five years before and after a flood), clustering standard errors both at the unit and calendar-year levels.[46] The coefficients of interest are β^τ—the treatment effect of moving from a non-flood state to a flood state at a horizon τ.[47] The treatment variable $Flood^\tau_{i,t,y}$ equals 1 if unit i experiences a flood τ years from t, where $t = 0$ corresponds to y. I include unit fixed effects ψ_i and stratum by cohort by relative time fixed effects $\delta_{s(i,y),t}$. To estimate the cumulative impact of floods as a summary measure of their consequences, I estimate equation 7 replacing each individual horizon term β^τ with simply an indicator for the observation occurring after the treatment event ($\tau \geq 0$).

$$Y_{i,t,y} = \alpha + \sum_{\tau=-5}^{5} \beta^\tau Flood^\tau_{i,t,y} + \psi_i + \delta_{s(i,y),t} + \varepsilon_{i,t,y} \tag{7}$$

One important feature of this specification is the shifting composition of units used to estimate treatment coefficients across horizons τ. This allows me to use all available data—even if I only observe outcomes over a limited time period—but makes interpreting the event study plots more challenging. For example, a visual "pre-trend" may be due to either a differential path of outcomes in to-be flooded areas in the years leading up to treatment, a shifting sample composition over the horizons, or some combination of the two. To separate these forces, I can restrict τ and y such that I observe each horizon for each event, creating a constant sample to remove any composition effect.

Representativeness This specification restricts the sample of flooding events in two important ways. First, because the satellite data restricts the time period over which I can measure flooding, I cannot estimate treatment effects of any flood that occurs in the first *pre* years of the satellite series. Intuitively, this stems from the fact that I cannot know whether

[45] For geographic inputs, I use the mean and standard deviation of elevation within each union, average distance to major rivers, and a host of flood risk measures produced by the Bangladeshi government based on hydrological modeling. See Appendix Section B for more details. With propensity score p, I assign flooded units a weight of one and control units a weight of $p/(1-p)$.

[46] For the synthetic control weights, I estimate the unit weights within calendar year and then apply them to the corresponding observation in the full regression.

[47] To account for collinearity, I normalize $\beta^{-1} = 0$. I estimate each horizon in a single regression, equivalently to using seemingly unrelated regressions, which allows me to easily test for differential treatment effects across horizons.

a flood occurred in the years prior to when I first measure floods, risking contamination from dynamic treatment effects of unobservable inundation events. Second, in order to construct plausible comparison groups, I must exclude flooding events for whom no control group exists or that violate another condition of constructing the sample. At an annual level, in the case of the marginal impact of flooding, with a path length of 5 years, I successfully match 62.52 percent of 7,893 all flooding events with at least one control unit with the exact same flood history $\mathbf{Flood}^5_{i,t-1}$. In order to ensure that potentially heterogeneous dynamic treatment effects do not bias my results, my estimator therefore departs from the estimand described in equation 6 by considering a less representative sample.

Placebo Tests In the spirit of Fisher (1935), I conduct a placebo test to further assess the robustness of my results. I assume the underlying data generating process draws from the distribution of flooding histories across unions and—holding fixed the sequence of flooding within place over time—randomly reassign units to different past paths. I then compare the β^τ I estimate from the true data to the distribution of treatment effects under the placebo assignments and calculate an exact p-value from this comparison.

Spatial Correlation I take two approaches to addressing potential spatial correlation. First, this design-based placebo approach naturally lends itself to a method to incorporate potential spatial correlation into inference. Floods are not randomly assigned across geographic areas, with inundation in one area likely correlated with inundation in a nearby area.[48] To assess the potential implications of this feature, I allow for correlation at an increasingly larger spatial scale and assess how the results shift. Let g^l denote different administrative levels indexed by $l = \{1, 2, 3, 4\}$, with 1 being the lowest level (unions, the geographic unit at which I estimate flooding) and 4 being the highest (indexing Bangladesh's seven divisions). To account for spatial correlations in flooding, I calculate the distribution of flood histories at each geographic aggregate g^l. In the case of unions ($g^l = 1$), this is simply the bootstrapping procedure above. At higher levels, I have a distribution of histories within each administrative unit. I then take a two-step bootstrapping procedure: first, randomly reassigning each geographic aggregate to a set of histories from the distribution of aggregates, and then within each of these aggregates, randomly assigning flood paths to unions drawing from the distribution of the newly assigned set. Note that the conservativeness of this test does not monotonically increase with l—in fact, at the extreme of the entire country, the procedure becomes equivalent to the union-level version. As a second approach

[48]To visualize this, Appendix Figure A.24 plots a binned scatter plot showing how, conditional on one union experiencing a flood in a given year, the likelihood of other unions in the country also experiencing a flood decays with distance.

to accounting for spatial correlation, I estimate standard errors allowing for spatial dependence across unions of an unknown form up to a fixed radius, following Conley (1999). In practice in this context, the spatial covariance is negative after making all of the restrictions to the sample discussed previously, resulting in *more* precise standard errors after making these adjustments. Due to this fact, I plot the wider confidence intervals in the event study plots, though I also report the Conley-adjusted p-values in the tables.[49]

Alternative Inference To relax the asymptotic assumptions underpinning the traditional standard errors I report, I additionally calculate confidence intervals via bootstrapping. The relative scarcity of floods raises concerns about corner cases, especially in situations where I only observe outcomes for a relatively small number of unions. To avoid this pitfall, instead of the standard non-parametric bootstrap, I implement the Bayesian bootstrap (Rubin, 1981) with 500 draws. In each iteration, I draw weights from an exponential distribution, normalizing their values to sum to the total number of observations, and re-estimating each statistic with those weights. I treat each cohort y as a different stratum and draw weights within strata, and then multiply these weights by the original weights used in the regression specification.

Interpretation of Point-Estimates Measurement error in my definition of flooding creates an important caveat for interpreting the coefficients from these specifications. Let $F_{i,t}$ denote the true (unobserved) flooding in location i at time t and $\widehat{F_{i,t}}$ denote my measure. The impact of flooding I estimate above based on the observed data $\widehat{F_{i,t}}$ gives a point estimate of $\widehat{\beta}$. Let β capture the true flooding impact if I had perfect data. One could imagine a hypothetical "first-stage" regression predicting the truth $F_{i,t}$ with the observed $\widehat{F_{i,t}}$, yielding a coefficient γ. If my measure perfectly predicted the truth (with a $\gamma = 1$), then $\widehat{\beta} = \beta$ would capture the true impact of floods. Measurement error certainly exists, however, implying that the effect sizes I estimate understate the consequences of flooding in proportion to the

[49] I use distance cutoffs of 10, 20, and 30 kilometers for the spatial kernel by adapting code from Hsiang (2010) and Baum-Snow and Han (Forthcoming) to account for the computation demands of my data. In the specifications using repeated cross-section individual observations from the administrative data, the computational demands of estimating the GMM make estimating these standard errors infeasible. However, in specifications collapsing the data to the union level, I can implement this procedure (albeit without the same effective frequency weights or individual controls), and all of the results remain the same. In specifications with smaller samples, due to the handling of fixed effects when I observe multiple observations per panel unit per period (as in the case of the individual-level repeated cross section), there can be very slight differences in the point estimates when estimating the Conley standard errors. However, these discrepancies are always incredibly tiny and never change any result in any meaningful way. Further, I find very similar results estimating spatial standard errors from Heldring and Calderon (2020).

predictive relationship between the observed flood measure and the truth ($\beta = \widehat{\beta}/\gamma$).[50] As a benchmark given the predictive ability of the supervised machine learning algorithm I estimate to measure floods, a conservative upper bound on γ would be .8, suggesting that all effect sizes I report should be scaled upwards by 25 percent.

5 The Impact of Floods on Economic Development

Using this empirical strategy, I estimate the causal effect of experiencing flooding on a variety of important indicators of economic development.

First, I find a strong and persistent negative impact on economic activity as proxied by nighttime lights visible from space (Henderson et al., 2012).[51] Because luminosity varies so much between urban and rural areas, I standardize the outcome using the never-flooded sample within urban status.[52] Figure 4 plots an event-study with standardized average median luminosity as the outcome. The results show a significant decline in light intensity in the wake of an inundation event. Table 2 presents stacked difference-in-differences results aggregating across horizons. These negative impacts persist across a battery of robustness checks. The shifting sample across horizons presents one concern with interpreting these event study plots; Appendix Figure E.3 shows the same pattern holds across three fixed samples, each consisting of floods across three consecutive calendar years between 2015 and 2019. Appendix Figure E.4 plots event studies for alternative nighttime luminosity outcomes; the results remain similar. To visualize this without the machinery of the difference-in-differences specification, Figure 5 plots the raw trends between flooded and non-flooded areas for the 2017 treatment cohort, chosen to balance placebo and dynamic treatment effects because the nighttime luminosity series begins in 2013. This graph illustrates the clear separation between flooded and non-flooded areas at the time of inundation.

I estimate the impact of floods on physical capital using remote sensing data analyzed by Google to estimate the footprint of every building in Bangladesh. However, this information is only available for satellite images taken at the end of 2020, and so I must use a cross-sectional regression to estimate the impacts. I calculate the share of area in a union covered by a building, take the log of this value to account for the skewed distribution, and standardize the value to the distribution of those places that did not experience floods in 2020. I estimate the impact of floods in 2020 using the same strata as discussed before (therefore

[50] This measurement error stems, for example, from the discrete nature in which I define flooding from a continuous measure or noise in the satellite data.

[51] Appendix Section C.1 describes my construction of this outcome in further detail.

[52] I define urban areas based on whether the name of the union in the GADM dataset includes either "Paurashava" or "Ward".

only comparing places with the same previous flooding history and expected flooding risk), along with controls for urban areas and fixed effects for 50 quantiles of union area. Floods cause a .2 standard deviation decrease on this measure of physical capital. Encouragingly, adding increasingly stringent fixed effects for administrative regions (division, then district, and then upazila), this negative coefficient *increases* significantly, ultimately rising to .5 standard deviations when only comparing places within the same upazila.

To shed light on individual behavior in the wake of flooding, I study impacts on occupational choice and migration from the Bangladesh government's Sample Vital Statistics Registration system, a large-scale, nationally administrative survey used to calculate key demographic and statistical information. I obtain this data from the Bangladesh Bureau of Statistics. For further details, see Appendix Section C.3.

I find that floods force structural transformation. Labor shifts out of the agricultural sector, as shown in Figures 8 and 9 and Table 5. This employment shifts into businesses and office work, as shown in Figures 10, 11, 12, 13 and Tables 7 and 6. Relatedly, I see a significant increase in net migration.

Floods push children into school, as shown in Figures 14 and 15 for boys and girls, respectively. Two economic forces could be driving this pattern. First, the marginal productivity of the child's labor on the farm might decline in the wake of inundation. The immediate opportunity cost of having your children leaving the fields could fall. Consistent with this channel, I observe a decline in crop yield as measured by the greenness of vegetation as captured by satellite data. To validate this variable, I digitize agricultural production data at the lowest geographic level for which it is available (which is not the union-level) from the Bangladeshi government and show that greenness strongly predicts true output. For details, see Appendix Section C.5. Second, parents' long-run beliefs about the returns to education may shift. Table 12 presents results correlating parents' perceived flood risk with their expectations about their children's ultimate educational attainment and occupation at age 25. I control for true flood risk ventiles and the child's current education, age, and gender, clustering standard errors at the household level. I am underpowered in this analysis, in part because I only asked this to half of the sample during the baseline. Nevertheless, the coefficients are consistent with parents perceiving higher flood risk also believing in higher returns to education.

I do not find statistically significant impacts on several other important measures of economic well-being. Using the administrative vital statistics data, I cannot detect effects on households' self-reported insolvency, literacy, or infant mortality. I confirm this latter result using geo-coded data from four rounds of the Demographic and Health Surveys. I do however find that floods delay marriage for women.

Differential Treatment Effects by Flood History How does past exposure to flooding impact the effects of subsequent disasters? To bring the data to bear on these dynamics, I augment the specification in equation 7 to include an interaction term with a measure of past inundation experience as shown in equation 8. I define experience as number of flooded periods in the strata-specific past flood history.[53]

$$Y_{i,t,y} = \alpha + \sum_{\tau=-5}^{5} \left(\beta^{\tau} Flood^{\tau}_{i,t,y} + \delta^{\tau} Flood^{\tau}_{i,t,y} \sum_{z=y-pre}^{y-1} Flood_{s(i,z)} \right) + \mathbf{X}_{i,t,y} + \lambda_{s(i,y),t} + \varepsilon_{i,t,y} \quad (8)$$

The results show that previous experience with flooding significantly mitigates the negative consequences of inundation. Table 3 presents the results for my preferred measure of nighttime luminosity, showing that enduring one additional flooding event in the preceding half decade reduces the treatment effect by approximately one third on average.

What explains this pattern? One potential explanation is simply mechanical: if floods destroy a sufficient amount of physical capital, then a floor effect could explain why subsequent floods show smaller effects. I do not find compelling evidence that this channel can explain the results, however. First, when restricting to only those observations with above-median nighttime luminosity in the year prior to the flood, I observe even stronger mitigation effects from past experience. Second, Figure 6 shows the results from quantile regressions for every decile beween .1 and .9. Across these moments, results remain remarkably constant, which suggests the floor effects do not explain the differential impact by past experience.

Experience-Driven Climate Adaptation A separate explanation for the differential treatment effects by past flooding exposure involves experience-driven climate adaptation. To illustrate these channels, I begin with a model of the harm caused by global warming featuring adaptation. Let $D_{i,t}$ denote the climate damage experienced by unit i in period t. Household i maximizes their discounted present value of utility, choosing whether to give up consumption today in order to invest $x_{i,t}$ in a costly climate adaptation technology. Adaptation $0 \leq A_{i,t} \leq 1$ mitigates the maximum damage $M_{i,t}$ caused by a negative climate shock, which occurs with probability $P_{i,t}$.

$$\underbrace{\mathbb{E}[D_{i,t}]}_{\substack{\text{Expected} \\ \text{Damage}}} = \underbrace{P_{i,t}}_{\substack{\text{Event} \\ \text{Probability}}} \times \underbrace{M_{i,t}}_{\substack{\text{Maximum} \\ \text{Damage}}} \times \underbrace{A_{i,t}}_{\text{Adaptation}} \quad (9)$$

[53] For consistency with the previous results, I include the set of specifications using the same synthetic control weights, noting however that they become conceptually much less clear in this interacted specification.

I assume the adaptation mitigation factor $A_{i,t}$ is a decreasing function $g(\cdot)$ of the stock of investments S_t made by i as of period t, bounded between 0 and 1, and that $g(0) = 1$. To ensure a steady state solution for this infinitely lived agent, I assume the stock depreciates such that only investments made over a fixed period of time τ persist: $S_t = \sum_{s=t-\tau-1}^{t-1} x_s$. The Bellman equation in equation 10 gives the continuation value capturing households' expected payoff at each time t, where δ denotes the discount factor and y_t denotes exogenously determined income. I assume an interior solution without savings, so households face a budget constraint $x_t \leq y_t$ that does not bind.

$$V_t = \arg\max_{x_t} \left(U(y_t - P_{i,t} M_{i,t} g \left(\sum_{s=t-\tau-1}^{t-1} x_s \right) - x_t) + \delta V_{t+1} \right) \tag{10}$$

Households will invest in climate adaptation until the marginal benefit in future mitigation no longer exceeds the marginal cost in forfeited consumption today. Equation 11 presents this indifference condition for period t suppressing the individual subscripts, where all derivatives are taken with respect to x_t.

$$U'(y_t - P_t M_t g(S_t) - x_t) = -\delta P_t M_t g'(S_t - x_{t-\tau-1} + x_t) U'(y_{t+1} - P_{i,t+1} M_{i,t+1} g(S_t - x_{t-\tau-1} + x_t) - x_{i,t+1}) \tag{11}$$

To solve for the steady state equilibrium investment, I note that given constant values for income and climate risk ($y = y_t \quad \forall t, P = P_t \quad \forall t, M = M_t \quad \forall t$) households will invest the same amount in each period: ($x = x_t \quad \forall t$). This delivers an optimal investment level independent of the functional form of the utility function, shown in equation 12.

$$1 = -\delta P M g'(\tau x) \tag{12}$$

To gain intuition for this result, consider a simple characterization of the adaptation production function: $A_{i,t} = 1/(1 + S_t) = 1/(1 + \tau x_t)$. The optimal investment level under this specification equals $(\sqrt{\tau \delta P M} - 1)/(\tau)$. This solution produces intuitive comparative statics: as the expected potential damage from floods PM increases, so too does household investment in adaptation technology. The more households value the future (higher δ), the more they sacrifice consumption today. By contrast, the longer past investments persist (larger τ), the less investment occurs. Based on research from climate science, I assume global warming operates by increasing the frequency of future flooding through higher $P_{i,t}$ and greater severity of future flooding via a larger $M_{i,t}$. Even in the absence of an experience effect, this will increase adaption investment through the higher expected cost faced by

households.

How might experience shape this equilibrium? I consider two potential channels. First, households may update their beliefs about the likelihood of future natural disasters. When households optimize based on their perceived climate risk $\widehat{PM}$ as opposed to the truth, updating in the aftermath of experiencing an event will increase investment. This occurs both for Bayesian processing and for a host of non-Bayesian frameworks, in which households overreact to salient or recent events, for instance (see Patel (2023) for examples of this type of updating). Second, experiencing a natural disaster may cause learning about successful adaptation, increasing the returns to any given investment x. Letting g^* denote the new adaptation production function, I assume experience results in more mitigation per marginal investment: $g'(x) > g^{*\prime}(x) \quad \forall x$. For the class of production functions I consider above, let experience simply increase efficiency by a factor β, such that $A^* = 1/(c + \beta\tau x)$ in steady state. Although equilibrium investment may decrease because households achieve more per dollar, the improvement in the technology of the adaptation production function more than compensates for this effect, resulting in a net increase in overall climate damage mitigation.[54] This prediction has important implications for projecting future climate damages. As global warming increases the likelihood of flooding $P_{i,t}$, equilibrium investment may rise not only from the higher expected risk but also from the act of actually experiencing a flood, helping to offset the harm caused by climate change.

Tests of the Model's Mechanisms I test for these two mechanisms underpinning experience-driven climate adaptation empirically. First, consider households' beliefs about P and M. The model suggests that one way past floods shape future reactions is through households updating their expectations. To test this, I use the survey I conduct with nearly 2,300 households in rural Bangladesh and directly elicit these two objects P and M. I estimate equation 13, shown below for P but identical for M as well, where i indexes households and j indexes villages. $Flood_{j(i)}$ captures flooding in 2022, $FloodRisk_{j(i)}$ denotes fixed effects for predicted flood risk ventiles, and I cluster standard errors at the village level. I use a visual belief elicitation method to capture farmers perceptions about future flood risk. See Patel (2023) for details on this method.

$$P_i = \alpha + \beta Flood_{j(i)} + \Gamma FloodRisk_{j(i)} + \varepsilon_i \tag{13}$$

Table 11 shows the impact of experiencing a flood in 2022 on farmers' perceptions of flood risk. Experiencing a flood increases both the expected damage after undergoing a flood (M

[54] See Appendix D for this derivation.

in the model), and the likelihood of a flood (P in the model).

Second, to examine whether the adaptation technology $g(\cdot)$ changes in response to floods, I focus on a key adaptation featuring a fixed cost: migration. Migrating outside of the village once lowers barriers to future migration, as shown in a rich literature in economics. I assess the extent to which lowering this fixed cost mediates the differential treatment effects by past flooding exposure.

To identify place-level variation in migration frictions, I use an instrument rooted in historical travel networks. The intuition underpinning this empirical strategy uses idiosyncratic road placement to isolate quasi-exogenous variation in travel times. I digitize data on road networks during the colonial era from an 1877 publication by the British Military detailing routes throughout the Bengal Presidency (Dutton, 1877).[55] This publication provides information on the different paths by which one might travel among important locations within the British Raj. I use transportation networks from more than 140 years ago to better capture the types of random forces with much less concern for reverse causality. In particular, I use data from prior to the construction of the Colonial Indian railroad network (Donaldson, 2018) and most other significant transportation projects in modern Bangladesh. The roads underpinning my instrument are significantly more susceptible to factors like river width and the preferences of the British Colonial Government than today's important determinants of economic activity.

I digitize the bilateral total route lengths from this publication. I begin by geocoding every major city for which the publication reports starting routes.[56] I then classify a set of target sample nodes that includes all locations with longitudes between 87 degrees and 93 degrees, thereby covering all of Bangladesh and its closest neighbors. Figure C.4 maps both the 34 target cities in blue and the 131 other cities in red superimposed on the modern-day boundaries of India and Bangladesh. Then, I digitize all route lengths starting from sample nodes, including those that end in destinations beyond the target longitudes. In the case of multiple routes from origin i to destination j, I calculate the average distance. This results in a final sample of 356 direction-specific routes covering 225 unique city-pairs, shown in Figure C.6.

For the main instrument, I simply calculate the distance from the centroid of each union to the closest point on a straight line between all nodes in the actual historical route network. As a first-stage, I use my survey of farmers to show that this measure strongly predicts households' perceived likelihood of migrating in the event of a flood. Interacting this term

[55]For details on this data and its construction, see Appendix Section C.6.

[56]I exclude the historical cities of Bhurtpur, Bustee, Kishnagurh, and Gya for which I cannot find reliable coordinates.

with previous flooding exposure, I see that previous inundation decreases the impact of future disasters significantly more in places closer to these historical transportation routes.

To test whether this mechanism might mediate the differential treatment effects by exposure, I add additional interaction terms with the log of this distance to colonial routes. The results vary somewhat depending on the weights used in the regression, but the overall pattern suggests that past exposure mitigates floods' economic harm less the closer a union falls to the historical transportation network.

6 Conclusion

In this paper, I first developed an approach to detecting local flooding inundation by combining supervised machine learning with various sets of remote sensing data. In doing so, I expand the temporal coverage of modern satellite technology to construct a consistent, accurate measure of surface water. This method can be applied anywhere in the world. Future research in this area might assess the validity of these estimates in other contexts, particularly the calibration of the flooding threshold, and explore other potential dimensions of flood severity such as water height. Measuring flooding accurately can help researchers and policymakers alike better understand their causes and consequences, a particularly urgent task given the natural disaster projections under the current global warming trajectory.

I next turn to the impact of flooding on economic development, and find strong negative impacts on economic activity. Previous flooding exposure mitigates these effects, and I find evidence consistent with experience-driven adaptation: places with lower fixed costs to adaptation experience greater mitigation from past disasters, and households increase their beliefs about future flood risk in response to flooding. These results have important implications for projecting future climate damages under global warming as endogenous adaptation may partially compensate for the increased rates of flooding.

References

Abadie, Alberto, Alexis Diamond, and Jens Haunmueller, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 2010, *105* (490).

Adhvaryu, Achyuta, Prashant Bharadwaj, James Fenske, Anant Nyshadham, and Richard Stanley, "Dust and death: evidence from the West African Harmattan," 2019.

Aker, Jenny C. and Kelsey Jack, "Harvesting the Rain: The Adoption of Environmental Technologies in the Sahel," *Review of Economics and Statistics*, 2023.

Albert, Christoph, Paula Bustos, and Jacopo Ponticelli, "The Effects of Climate Change on Labor and Capital Reallocation," *Working Paper*, 2023.

Arceo-Gomez, Hanna and Paulina Oliva, "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City," *Economic Journal*, 2014, *126* (591), 257–280.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager, "Synthetic Difference-in-Differences," *American Economic Review*, 2021, *111* (12).

Balboni, Clare Alexandra, "In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities," *Working Paper*, 2019.

Balboni, Clare, Johannes Boehm, and Mazhar Waseem, "Firm Adaptation in Production Networks: Evidence from Extreme Weather Events in Pakistan," *Working Paper*, 2023.

Barreca, Alan, Karen Clay, Olivier Dêschenes, Michael Greenstone, and Joseph S. Shapiro, "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century," *Journal of Political Economy*, February 2016, *124* (1).

Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur, "The healthcare cost of air pollution: Evidence from the world's largest payment network," 2018.

— , — , Liguo Lin, and Eric Zou, "From fog to smog: The value of pollution information," 2019.

Baum-Snow, Nathaniel and Lu Han, "The Microgeography of Housing Supply," *Journal of Political Economy*, Forthcoming.

Baumol, William J., "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review*, 1967, *57* (3), 415–426.

Berkouwer, Susanna B. and Joshua T. Dean, "Credit, Attention, and Externalities in the Adoption of Energy-Efficient Technologies by Low-Income Households," *American Economic Review*, 2022, *112* (10), 3291–3330.

Bharadwaj, Prashant and Jamie T. Mullins, "Weather, Climate, and Migration in the United States," *Working Paper*, 2021.

Boustan, Leah Platt, Matthew E. Kahn, and Paul W. Rhode, "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century," *American Economic Review*, 2012, *102* (3), 238–244.

Brakenridge, G.R., "Global Active Archive of large Flood Events," *Dartmouth Flood Observatory, University of Colorado*, 2023.

Brunner, Manuela I., Daniel L. Swain, Raul R. Wood, Florian Willkofer, James M. Done, Eric Gilleland, and Ralf Ludwig, "An Extremeness Threshold Determines the Regional Response of Floods to Changes in Rainfall Extremes," *Communications Earth & Environment*, 2021, *2* (1).

Burgess, Robin, Olivier Dêschenes, Dave Donaldson, and Michael Greenstone, "Weather, Climate Change and Death in India," 2017. Working Paper.

Burke, Marshall and Kyle Emerick, "Adaptation to Climate Change: Evidence from US Agriculture," *American Economic Journal: Economic Policy*, 2016, *8* (3).

— , Solomon Hsiang, and Edward Miguel, "Global Non-Linear Effect of Temperature on Economic Production," *Nature*, 2015, *527*, 235–239.

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli, "Agricultural Productivity and Structural Transformation: Evidence from Brazil," *American Economic Review*, 2016, *106* (6), 1320–1365.

Carleton, Tamma, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Robert E. Kopp, Kelly McCusker, Ishan Nath, James Rising, Ashwin Rode, Samuel Seo, Arvid Viaene, Jiacan Yuan, and Alice Tianbo Zhang, "Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits," *Quarterly Journal of Economics*, 2022, *137* (4), 2037–2105.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *Quarterly Journal of Economics*, 2019, *134* (3).

Chang, T. Y., J. Graff Zivin, T. Gross, and M. Neidell, "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China," *American Economic Journal: Applied Economics*, 2019, *11* (1), 151–172.

Conley, T.G., "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Dell, Melissa, Benjamin Jones, and Benjamin Olken, "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, 2012, *4* (3).

— , — , and — , "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, 2014.

Deschênes, Olivier and Michael Greenstone, "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather," *American economic review*, 2007, *97* (1), 354–385.

— and — , "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US," *American Economic Journal: Applied Economics*, 2011, *3* (4), 152–185.

Desmet, Klaus, Robert E Kopp, Scott A Kulp, Dávid Krisztián Nagy, Michael Oppenheimer, Esteban Rossi-Hansberg, and Benjamin H Strauss, "Evaluating the economic cost of coastal flooding," Technical Report, National Bureau of Economic Research 2018.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 2018, *108* (4-5), 899–934.

Dube, Arindrajit, Daniele Girardi, Jordá. Óscar, and Alan M. Taylor, "A Local Projections Approach to Difference-in-Differences Event Studies," *Working Paper*, 2023.

Dutton, Charles, *Routes in the Bengal Presidency*, Calcutta: Office of Quartermaster-General in India, 1877.

Elliott, Robert J., Eric Strobl, and Puyang Sun, "The local impact of typhoons on economic activity in China: A view from outer space," *Journal of Urban Economics*, 2015, *88*, 50–66.

Elvidge, C. D., M. Zhizhin, T. Ghosh, FC Hsu, and J. Taneja, "Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthl Averages: 2012 to 2019," *Remote Sensing*, 2021, *13* (5).

Emerick, Kyle, Alain de Janvry, Elisabeth Sadoulet, and Manzoor H. Dar, "Technological Innovations, Downside Risk, and the Modernization of Agriculture," *American Economic Review*, 2016, *106* (6).

Felbermayr, Gabriel and Jasmin Groschl, "Naturally negative: The growth effects of natural disasters," *Journal of Development Economics*, 2014, *111*, 92–106.

Fisher, Ronald A., *The Design of Experiments*, Edinburgh: Oliver and Boyd, Ltd., 1935.

Gandhi, Sahil, Matthew E Kahn, Rajat Kochhar, Somik Lall, and Vaidehi Tandel, "Adapting to Flood Risk: Evidence from a Panel of Global Cities," 2022, (30137).

Garioud, A., S. Valero, S. Giordano, and C. Mallet, "On the Joint Exploitation of Optical and SAR Satellite Imagery for Grassland Monitoring," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2020, *XLIII-B3-2020*, 591–598.

Gaume, E. and M. Borga, "Post-Flood Field Investigations in Upland Catchments after Major Flash Floods: Proposal of a Methodology and Illustrations," *Journal of Flood Risk Management*, 2008, pp. 175–189.

Global Administrative Areas, "GADM database of Global Administrative Areas," *Version 3.4*, 2018.

Gollin, Douglas, Casper Worm Hansen, and Asger Mose Wingender, "Two Blades of Grass: The Impact of the Green Revolution," *Journal of Political Economy*, 2021, *129* (8), 2344–2384.

__ , Stephen Parente, and Richard Rogerson, "The Role of Agriculture in Development," *American Economic Review*, 2002, *92* (2), 160–164.

Greenstone, Michael and Rema Hanna, "Environmental regulations, air and water pollution, and infant mortality in India," *American Economic Review*, 2014, *104* (10), 3038–3072.

__ , Andrew Hultgren, Tamma Carleton, Michael Delgado, Diana R. Gergel, Trevor Houser, Solomon Hsiang, Amir Jina, Robert E. Kopp, Steven B. Malevich, Kelly E. McCusker, Terin Mayer, Ishan Nath, James Rising, Ashwin Rode, and Jiacan Yuan, "Estimating Global Impacts to Agriculture from Climate Change Accounting for Adaptation," November 2022.

Guha-Sapir, D., R. Below, and Ph. Hoyois, "EM-DAT: The CRED/OFDA International Disaster Database," *Universite Catholique de Louvain*, 2023.

Guiteras, Raymond, "The Impact of Climate Change on Indian Agriculture," *Working Paper*, 2009.

__ , Amir Jina, and A. Mushfiq Mobarak, "Satellites, Self-Reports, and Submersion: Exposure to Floods in Bangladesh," *American Economic Review, Papers and Proceedings*, 2015, *105* (5), 232–236.

Hanna, Rema and Paulina Oliva, "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City," *Journal of Public Economics*, 2015, *122*, 68–79.

Heldring, Leander and Luis Calderon, "Spatial Standard Errors for Several Commonly Used M-Estimators," *Working Paper*, 2020.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, 2012, *102* (2).

Hornbeck, R., "Dust Bowl Migrants: Environmental Refugees and Economic Adaptation," *The Journal of Economic History*, 2023, *83* (3), 645–675.

Hornbeck, Richard and Suresh Naidu, "When the Levee Breaks: Black Migration and Economic Development in the American South," *American Economic Review*, 2014, *104* (3), 963–990.

Hsiang, Solomon M., "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America," *Proceedings of the National Academy of Sciences*, 2010, *107* (35).

__ and Amir S. Jina, "The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence from 6,700 Cyclones," *Working Paper*, 2014.

Hsiao, Allan, "Sea Level Rise and Urban Adaptation in Jakarta," *Working Paper*, 2023.

Huffman, George J., "Algorithm Theoretical Basis Document Version 06: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG)," *National Aeronautics and Space Administration (NASA)*, 2019.

IPCC, *Climate Change 2022: Impacts, Adaptation and Vulnerability*, Cambridge, UK and New York, USA: Cambridge University Press, 2022.

Ito, K. and S. Zhang, "Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China," *Journal of Political Economy*, 2020, *128* (5), 1627–1672.

Jayachandran, Seema, "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires," *Journal of Human Resources*, 2009, *44* (4), 916–954.

Jia, Ruixue, Xiao Ma, and Victoria Wenxin Xie, "Expecting Floods: Firm Entry, Employment, and Aggregate Implications," Working Paper 30250, National Bureau of Economic Research 2022.

Kahn, Matthew E., "The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions," *Review of Economics and Statistics*, 2005, *87* (2), 271–284.

Kala, N., C. Balboni, and S. Bhogale, "Climate Adaptation," *VoxDevLit*, 2023, *7* (3).

Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song, "The productivity consequences of pollution-induced migration in China," *Working Paper*, 2021.

Khojeh, Shokoufeh, Behzad Ataie-Ashtiani, and Seiyed Mossa Hosseini, "Effect of DEM resolution in flood modeling: a case study of Gorganrood River, Northeastern Iran," *Natural Hazards*, July 2022, *112* (3), 2673–2693.

Kocornik-Mina, Adriana, Thomas K. J. McDermott, Guy Michaels, and Ferdinand Rauch, "Flooded Cities," *American Economic Journal: Applied Economics*, April 2020, *12* (2), 35–66.

Komi, Kossi, Jeffrey Neal, Mark A. Trigg, and Bernd Diekkrüger, "Modelling of flood hazard extent in data sparse areas: a case study of the Oti River basin, West Africa," *Journal of Hydrology: Regional Studies*, 2017, *10*, 122–132.

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie, "Beyond Balanced Growth," *Review of Economic Studies*, 2001, *68* (4), 869–882.

Kuznets, Simon, "Quantitative Aspects of the Economic Growth of Nations: II: Industrial Distribution of National Product and Labor Force," *Economic Development and Cultural Change*, 1957, *5* (4), 1–111.

Lai, Kelvin, Jeremy R. Porter, Mike Amodeo, David Miller, Michael Marston, and Saman Armal, "A Natural Language Processing Approach to Understanding Context in the Extraction and GeoCoding of Historical Floods, Storms, and Adaptation Measures," *Information Processing and Management*, 2022, *59* (1), 102735.

Lane, Gregory, "Adapting to Floods with Guaranteed Credit: Evidence from Bangladesh," *Working Paper*, 2022.

Lee, Jong-Sen, Jen-Hung Wen, T.L. Ainsworth, Kun-Shan Chen, and A.J. Chen, "Improved Sigma Filter for Speckle Filtering of SAR Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, 2009, *47* (1), 202–213.

Lewis, Benjamin and Devika Kakkar, "Harvard CGA Geotweet Archive v2.0," 2016.

Lewis, William Arthur, "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 1954, *22* (2), 139–91.

Lumbroso, Darren and Eric Gaume, "Reducing the uncertainty in indirect estimates of extreme flash flood discharges," *Journal of Hydrology*, 2012, *414-415*, 16–30.

McGuirk, Eoin F and Nathan Nunn, "Transhumant pastoralism, climate change, and conflict in africa," Technical Report, National Bureau of Economic Research 2020.

Moscona, Jacob, "Agricultural Development and Structural Change, Within and Across Countries," *Working Paper*, 2019.

__ , "Environmental Catastrophe and the Direction of Invention: Evidence from the American Dust Bowl," *Working Paper*, 2022.

__ **and Karthik Sastry**, "Does Directed Innovation Mitigate Climate Damage? Evidence from US Agriculture," *Quarterly Journal of Economics*, 2023.

Mueller, Valerie, Clark Gray, and Katrina Kosec, "Heat Stress Increases Long-Term Human Migration in Rural Pakistan," *Nature Climate Change*, 2014, *4*, 182–185.

Murphy, Kevin M., Andrei Shleifer, and Robert Vishny, "Income Distribution, Market Size, and Industrialization," *Quarterly Journal of Economics*, 1989, *104* (3), 537–564.

Ngai, L. Rachel and Christopher A. Pissarides, "Structural Change in a Multisector Model of Growth," *American Economic Review*, 2007, *97* (1), 429–443.

Nurkse, Ragnar, *Problems of Capital Formation in Underdeveloped Countries* 1953.

Pande, R., M. Greenstone, J. Nilekani, A. Sudarshan, A. Suganathan, and N. Ryan, "Lower Pollution, Longer Lives: Life Expectancy Gains if India Reduced Particulate Matter Pollution," *Economic and Political Weekly*, 2015, *L* (8).

Patel, Dev, "Learning About a Warming World: Attention and Adaptation in Agriculture," *Working Paper*, 2023.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, 2011, *12*, 2825–2830.

Pekel, Jean-Francois, Andrew Cottam, Noel Gorelick, and Alan S. Belward, "High-Resolution Mapping of Global Surface Water and Its Long-Term Changes," *Nature*, 2016, *540*, 418–422.

Pelli, Martino, Jeanne Tschopp, Natalia Bezmaternykh, and Kodjovi M. Eklou, "In the eye of the storm: Firms and capital destruction in India," *Journal of Urban Economics*, 2023.

Rangel, Marcos A and Tom S Vogl, "Agricultural fires and health at birth," *Review of Economics and Statistics*, 2019, *101* (4), 616–630.

Ratledge, Nathan, Gabe Cadamuro, Brandon de la Cuesta, Matthieu Stigler, and Marshall Burke, "Using Machine Learning to Assess the Livelihood Impact of Electricity Access," *Nature*, 2022.

Rentschler, Jun, Melda Salhab, and Bramka Arga Jafino, "Flood Exposure and Poverty in 188 Countries," *Nature Communications*, 2022, *13* (3527).

Rolf, Esther, Jonathan Proctor, Tamma Carleton, Ian Bolliger, Vaishaal Shankar, Miyabi Ishihara, Benjamin Recht, and Solomon Hsiang, "A Generalizable and Accessible Approach to Machine Learning with Global Satellite Imagery," *Nature Communications*, 2021, *12* (4392).

Rostow, W. W., *The Stages of Economic Growth: A Non-Communist Manifesto*, London: Cambridge University Press, 1960.

Roth, Jonathan, "Pre-Test with Caution: Event-Study Estimates After Testing for Parallel Trends," *American Economic Review: Insights*, 2022, *4* (3).

Rubin, Donald B., "The Bayesian Bootstrap," *The Annals of Statistics*, 1981, *9* (1).

Shah, Manisha and Bryce Millett Steinberg, "Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital," *Journal of Political Economy*, 2017, *125* (2), 527–561.

Shun, Maggie and Vera Lummis, "Sending Cash to Flood Survivors—4 Things We Got Right and Wrong," 2023.

Sirko, W., S. Kashubin, M. Ritter, A. Annkah, y.S.E. Bouchareb, Y. Dauphin, D. Keysers, M. Neumann, M. Cisse, and J.A. Quinn, "Continental-Scale Building Detection from High Resolution Satellite Imagery," *Working Paper*, 2021.

Somanathan, E., Rohini Somanathan, Anant Sudarshan, and Meenu Tewari, "The Impact of Temperature on Productivity and Labor Supply: Evidence from Indian Manufacturing," *Journal of Political Economy*, June 2021, *129* (6).

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

Tellman, B., J.A. Sullivan, and C. et al. Kuhn, "Satellite Imaging Reveals Increased Proportion of Population Exposed to Floods," *Nature*, 2021, *596*, 80–86.

UNOSAT, "UNOSAT Flood Dataset," 2019.

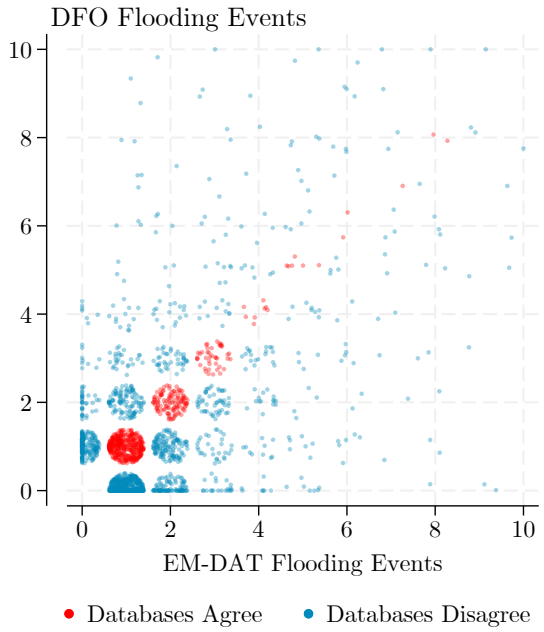
Vollrath, Andreas, Adugna Mullissa, and Johannes Reiche, "Angular-Based Radiometric Slope Correction for Sentinel-1 on Google Earth Engine," *Remote Sensing*, 2020, *12* (11).

Walker, Kendra, Ben Moscona, Kelsey Jack, Seema Jayachandran, Namrata Kala, Rohini Pande, Jiani Xue, and Marshall Burke, "Detecting Crop Burning in India using Satellite Data," *Working Paper*, 2022.

Werner, M. G. F., "A Comparison of Flood Extent Modelling Approaches through Constraining Uncertainties on Gauge Data," *Hydrological Earth Systems Sciences*, 2004, *8*, 1141–1152.

World Bank, "World Development Indicators," 2023.

Figure 1: Comparison of EM-DAT and DFO Flooding Databases



Note: Figure 1 plots a scatter plot of the number of flooding events at the country-by-year level in the EM-DAT database (Guha-Sapir et al., 2023)—along the x-axis—and the DFO archive (Brakenridge, 2023)—along the y-axis. Red dots denote when the two datasets agree on the number of flooding events, and blue signifies disagreement. Only floods since 2000 are considered, and only country-years for which at least one of the datasets indicates a flood. Points have been jittered to help visualize mass at each comparison.

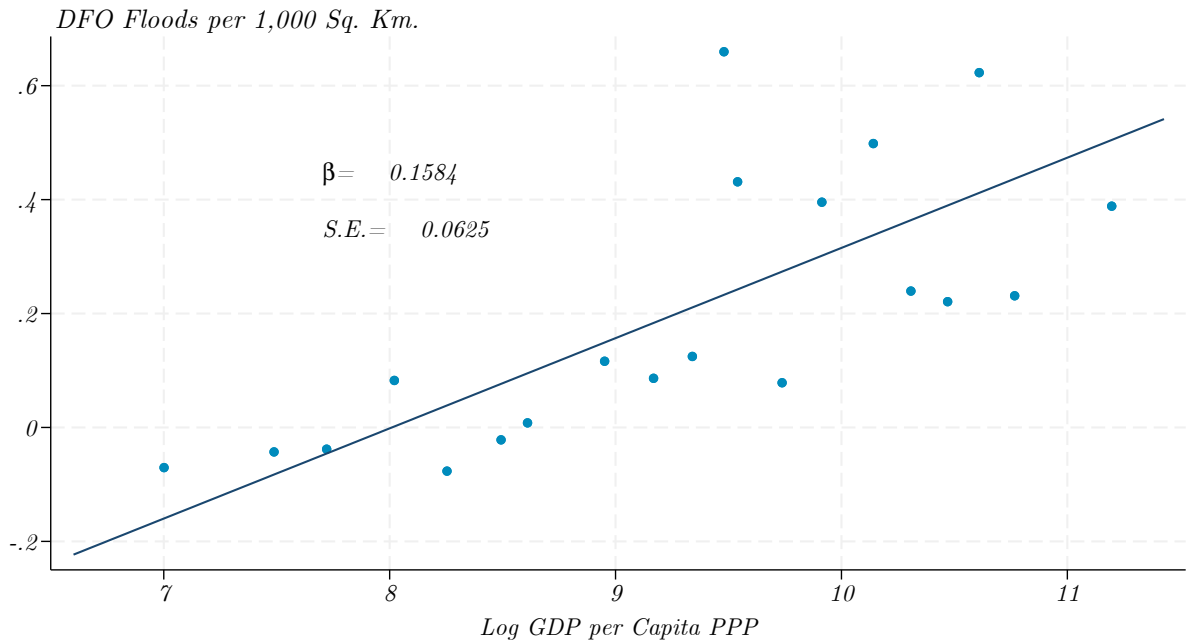
Table 1: Farmers' Reported Damages from Floods by Length

Duration	(1) Harm House	(2) House Damage	(3) Harm Crops	(4) Crop Damage
One Day	32.35	13.40	55.88	52.63
One to Three Days	51.18	35.50	78.29	57.75
Three Days to One Week	42.51	37.26	85.43	67.06
One Week to One Month	34.12	45.05	89.61	82.37
More than One Month	30.97	39.27	93.65	94.71

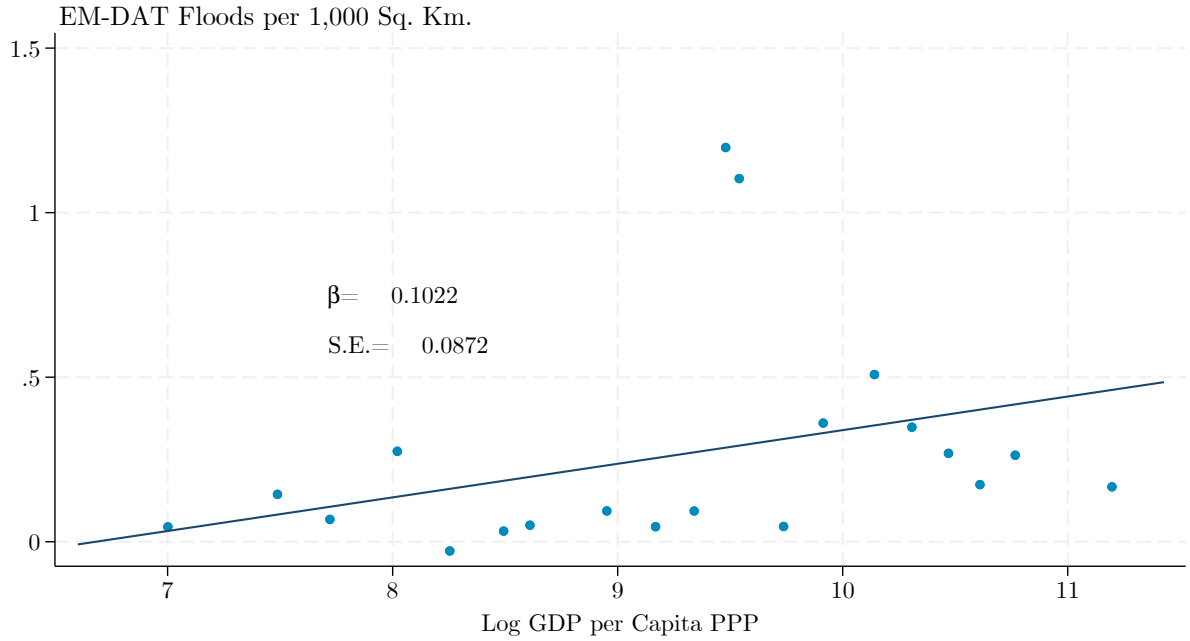
Note: Table 1 presents farmers' reports of damages incurred during all floods they recall, averaged by duration of the inundation event. Column (1) reports the percent of flooding events that damage the respondent's home. Column (2) gives—conditional on experiencing any property damage—the magnitude of the cost expressed as a percentage of household annual income. Column (3) shows the percent of events damaging the respondent's crops, and column (4) shows—conditional on crop damage—the amount of damage expressed as a percentage of total harvest.

Figure 2: Rich-Country Bias in Existing Flooding Databases

(a) DFO

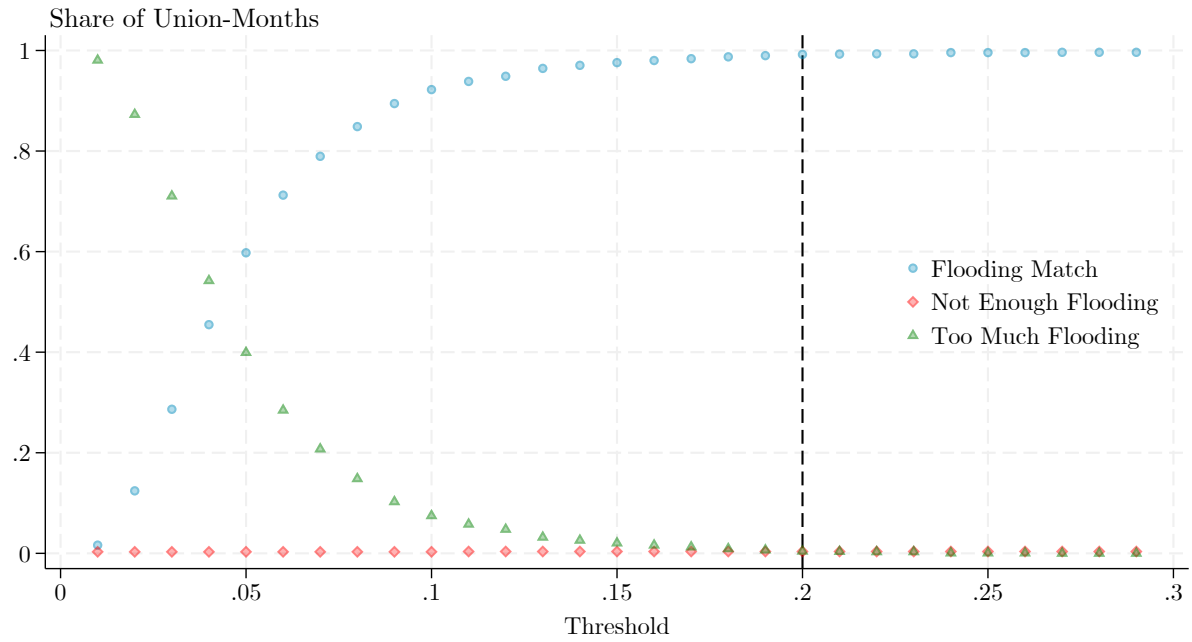


(b) EM-DAT



Note: Figure 2 presents binned scatter plots of the total number of floods since 2000 per land area (y-axis) against the log per capita income in purchasing power parity terms (x-axis) at the country-level. Both plots adjust for quadratic polynomials in the latitude and longitude of each country's centroid.

Figure 3: Choosing the Flooding Threshold τ



Note: Figure 3 compares the number of flood days in a union-month as estimated via the remote sending data for a given residual threshold τ against the number as reported by farmers in those places during the survey. The threshold above which I classify residuals as floods varies along the x-axis. The share of union-months is plotted along the y-axis, where the total sample includes all 250 unions from the survey sample and every month between April 2022 and the month prior to when farmers in that union were interviewed (either October, November, or December). The blue circles denote the share of union-months featuring a perfect match between the number of flood days via the remote sensing method and the farmers' answers. The green triangles show the portion of union-months for which the method overestimates the amount of flooding relative to the farmers' recall. The red diamonds capture the fraction of union-months when the algorithm estimates too few flood days.

Figure 4: The Impact of Floods on Nighttime Luminosity

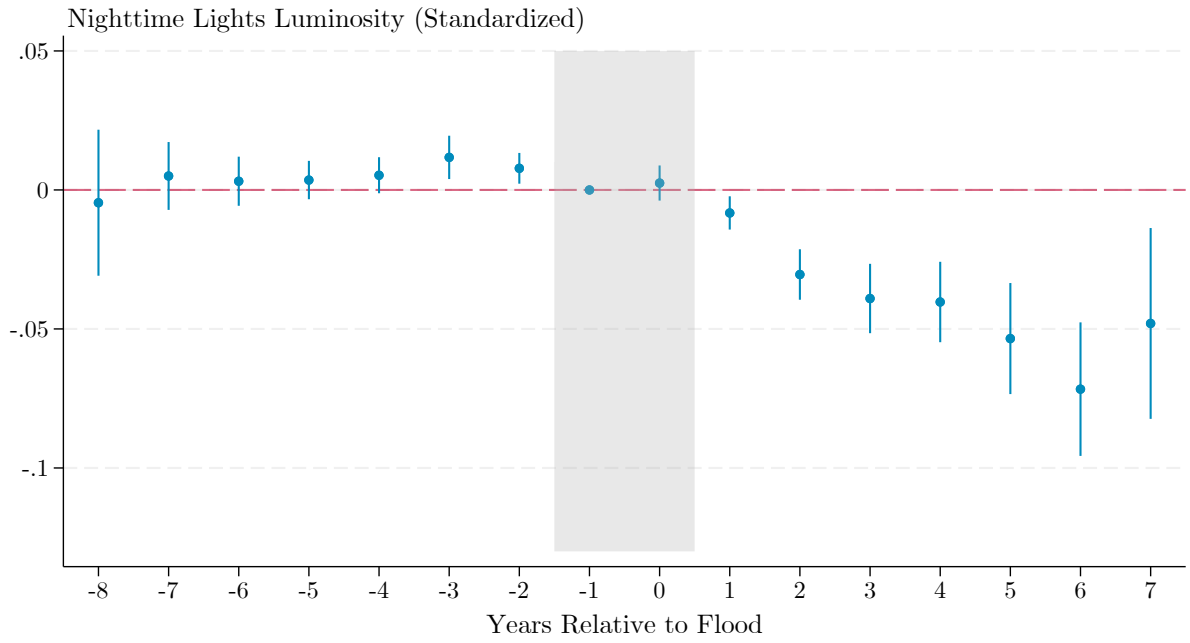
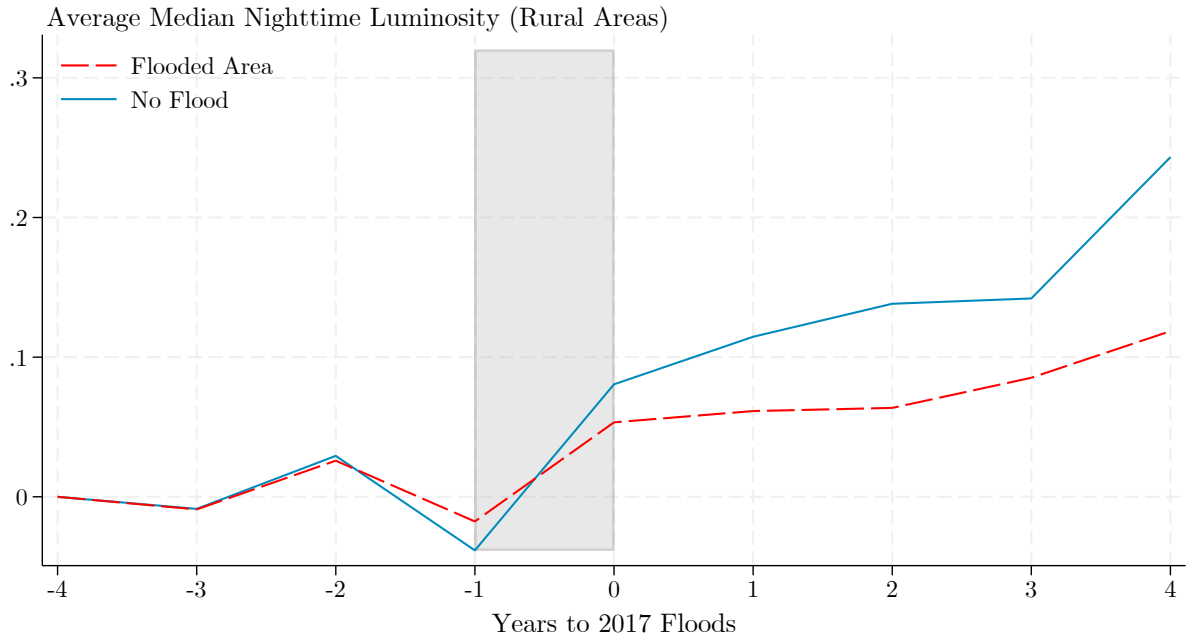


Figure 5: The Impact of Floods on Nighttime Luminosity—Raw Data Trends



Note: Figure 4 plots the coefficients from an event study on the impact of flooding. The outcome is the average value within each polygon of the median luminosity after masking out problematic observations in the satellite data, standardized to the never-flooded unit distribution. All regressions include union and strata by relative time fixed effects with standard errors clustered at the union level and union-by-cohort level synthetic control weights. Figure 5 plots the evolution of average median masked nighttime luminosity for flooded unions and control unions in the 2017 cohort, where the average values of each have been normalized to 0 at $t = -4$. The year 2017 was chosen because the nighttime luminosity data only begins in 2013. I standardized the levels to be 0 in period -4. I calculate weighted averages using propensity scores based on logistic regressions of geographic features using all years between 2002 and 2017. In both figures, the sample excludes control unions whose centroids fall within 10 kilometers of any treated union centroid.

Table 2: The Impact of Floods on Nighttime Luminosity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0257***	-0.0190***	-0.0348***	-0.0376***	-0.0228***	-0.0439***	-0.0472***	-0.0242***	-0.0519***
	(0.00462)	(0.00348)	(0.00410)	(0.00769)	(0.00407)	(0.00548)	(0.0119)	(0.00526)	(0.00659)
Sample	All	All	All	All	All	All	All	All	All
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean	-.13	-.13	-.13	-.127	-.127	-.127	-.14	-.14	-.14
Observations	24642	24642	24642	20502	20502	20502	16776	16776	16776
Clusters	2,738	2,738	2,738	2,278	2,278	2,278	1,864	1,864	1,864

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0303***	-0.0241***	-0.0420***	-0.0425***	-0.0291***	-0.0513***	-0.0478***	-0.0312***	-0.0567***
	(0.00600)	(0.00413)	(0.00518)	(0.00952)	(0.00515)	(0.00662)	(0.0139)	(0.00651)	(0.00749)
Sample	First	First	First	First	First	First	First	First	First
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean	-.117	-.117	-.117	-.116	-.116	-.116	-.133	-.133	-.133
Observations	21564	21564	21564	17757	17757	17757	14652	14652	14652
Clusters	2,396	2,396	2,396	1,973	1,973	1,973	1,628	1,628	1,628

Note: Table 2 presents stacked difference-in-differences estimates of the impact of flooding on average masked median luminosity, standardized to the never-flooded union distribution. All specifications include union fixed effects and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between synthetic control weights from Arkhangelsky et al. (2021), propensity score weights calculated from the strata, or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 3: The Heterogeneous Impact of Floods by Experience on Nighttime Luminosity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0280***	-0.0136***	-0.0192***	-0.0406***	-0.0159***	-0.0253***	-0.0483***	-0.0179***	-0.0321***
	(0.00533)	(0.00195)	(0.00207)	(0.00869)	(0.00256)	(0.00296)	(0.0132)	(0.00332)	(0.00382)
Post × Flood × Experience	0.00735	0.00614***	0.00801***	0.0117*	0.00763***	0.0130***	0.00514	0.00872***	0.0168***
	(0.00484)	(0.00146)	(0.00166)	(0.00707)	(0.00166)	(0.00204)	(0.0120)	(0.00213)	(0.00277)
Sample	All	All	All	All	All	All	All	All	All
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Conley Exp. p-value	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean	-.13			-.127			-.14		
Observations	59,040	117,504	117,504	41,625	89,892	89,892	28,719	68,301	68,301
Clusters	2,738	3,436	3,436	2,278	3,168	3,168	1,864	2,909	2,909

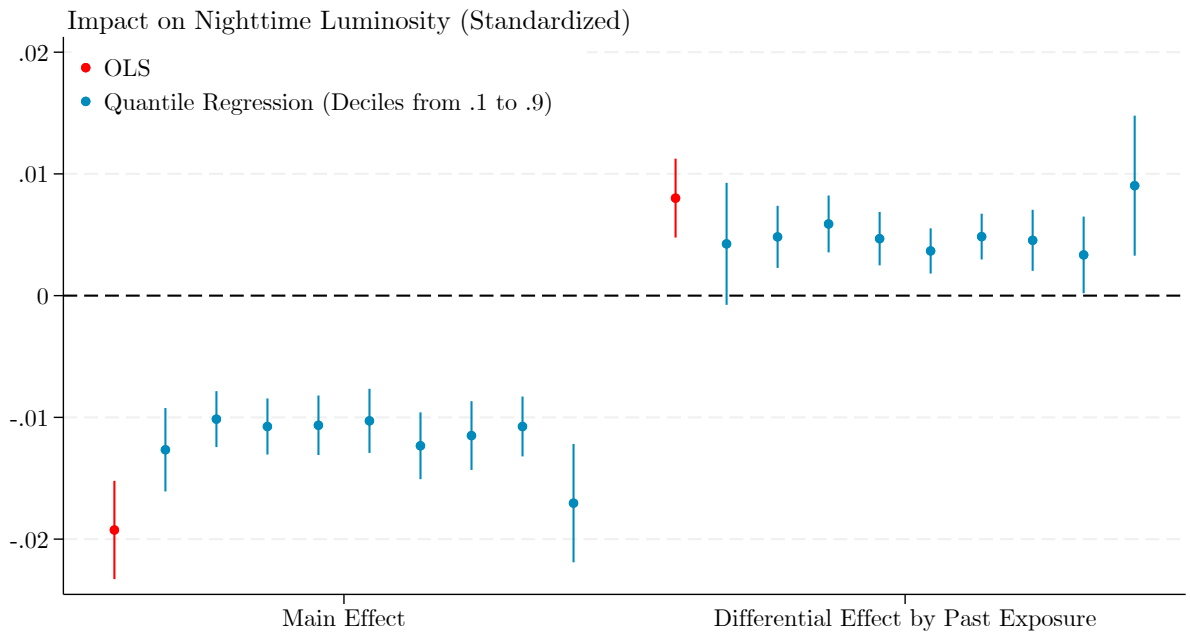
Note: Table 3 presents stacked difference-in-differences estimates of the impact of flooding on average masked median luminosity, standardized to the never-flooded union distribution. I define experience in the interaction term as the number of flooded years in that strata experienced in the five years leading up to treatment. All specifications include union fixed effects and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between synthetic control weights from Arkhangelsky et al. (2021), propensity score weights calculated from the strata, or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 4: The Impact of Floods on Built Structures

	(1)	(2)	(3)	(4)	(5)	(6)
Flood	-0.434***	-0.422***	-0.461***	-0.405***	-0.545**	-0.331**
	(0.0721)	(0.0532)	(0.122)	(0.0741)	(0.214)	(0.131)
Weights	Propensity	Propensity	Propensity	Propensity	Propensity	Propensity
Spillover Radius	10	10	20	20	30	30
Control Mean	.121	.121	.066	.066	-.005	-.005
Observations	974	974	683	683	497	497

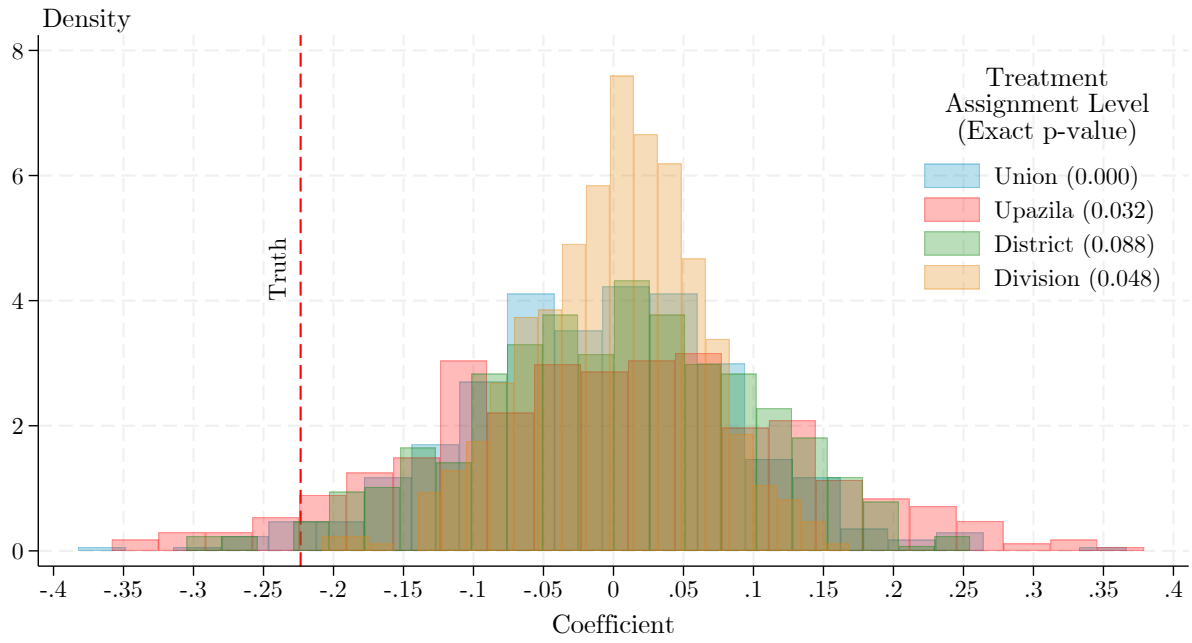
Note: Table 4 presents regressions of the impact of floods in 2020 on physical capital as detected by satellites at the end of the year. The data come from Google Open Buildings, covering the footprint of every structure in Bangladesh (see Appendix Section C.2 for details). To construct the outcome in this specification, I calculate the share of area in a union covered by a building, take the log of this value to account for the skewed distribution, and standardize the value to the distribution of those places that did not experience floods in 2020. The regression includes fixed effects for strata (therefore only comparing places with the same previous flooding history and expected flooding risk), a control for urban areas, and fixed effects for 50 quantiles of union area.

Figure 6: Quantile Regression Results of the Impact of Floods on Nighttime Luminosity



Note: Figure 6 presents results from quantile regressions of the impact of floods on nighttime luminosity and an interaction term with the number of previous floods experienced in the previous five years. The main OLS coefficient is plotted in red. From left to right, the impacts on the 10^{th} through 90^{th} percentiles of the distribution are shown in blue. All specifications cluster standard errors at the union level, and no weights are used.

Figure 7: The Impact of Floods on Built Structures: Placebo Test



Note: Figure 7 presents coefficients of the impact of flooding on built structures using draws from placebo flood distributions. The data come from Google Open Buildings, covering the footprint of every structure in Bangladesh (see Appendix Section C.2 for details). To construct the outcome in this specification, I calculate the share of area in a union covered by a building, take the log of this value to account for the skewed distribution, and standardize the value to the distribution of those places that did not experience floods in 2020. The regression includes fixed effects for strata (therefore only comparing places with the same previous flooding history and expected flooding risk), a control for urban areas, and fixed effects for 50 quantiles of union area. The treatment assignment level denotes the administrative level from which I draw the distribution of floods in the two-step assignment process.

Figure 8: The Marginal Impact of Floods on Agricultural Employment—Men

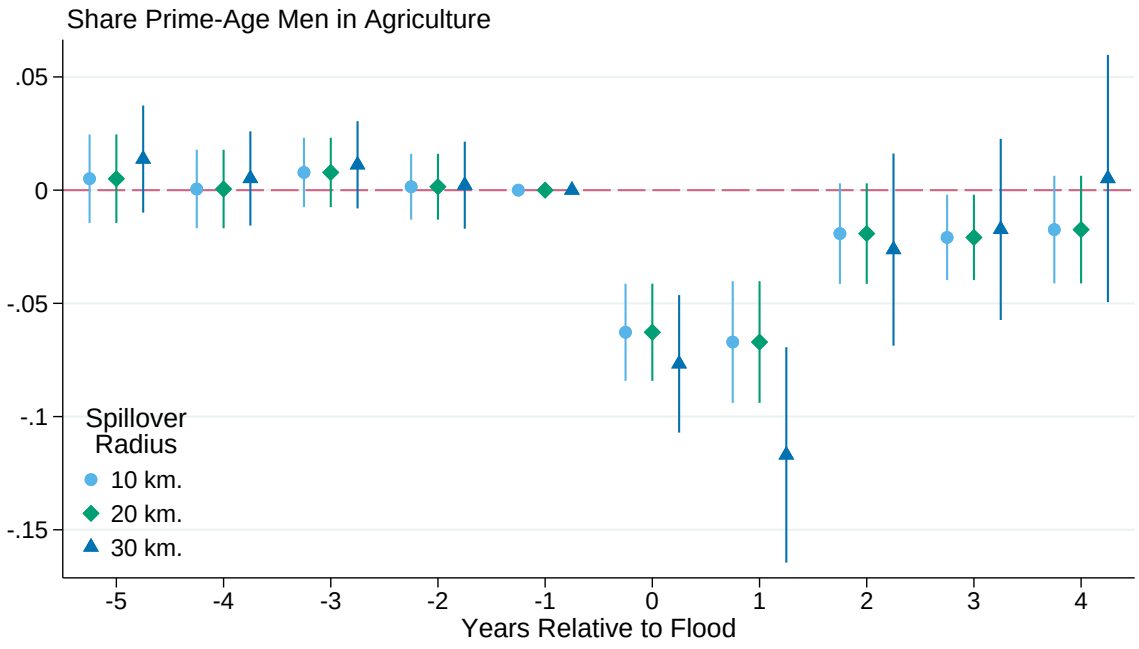
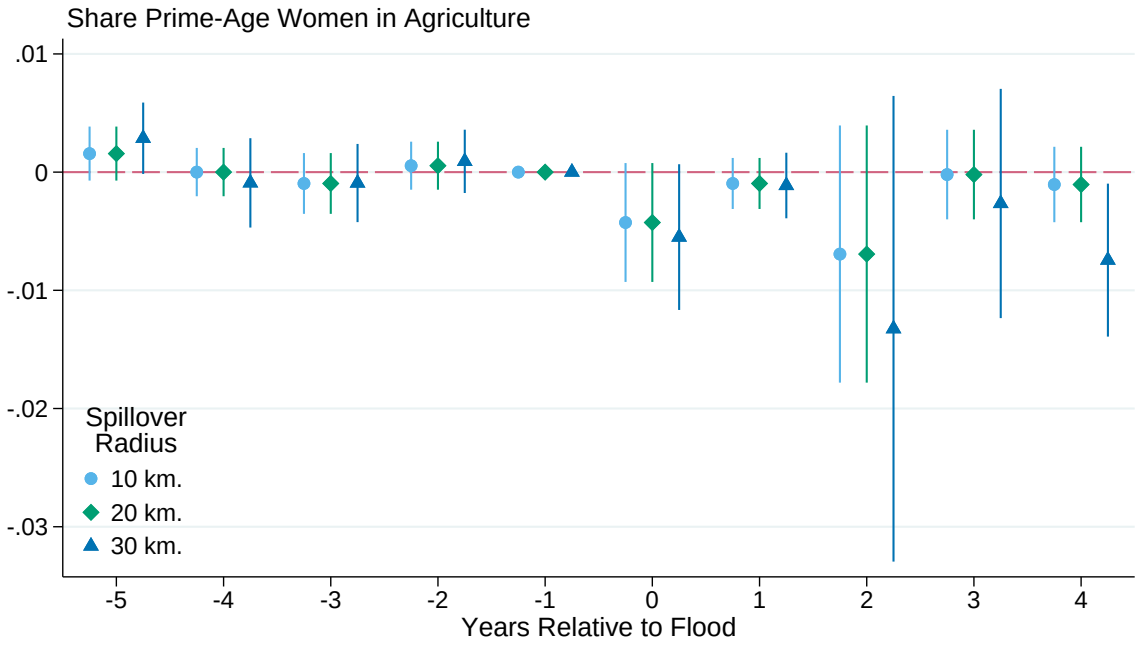


Figure 9: The Marginal Impact of Floods on Agricultural Employment—Women



Note: This figure plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. All regressions include age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.

Figure 10: The Marginal Impact of Floods on Business Employment—Men

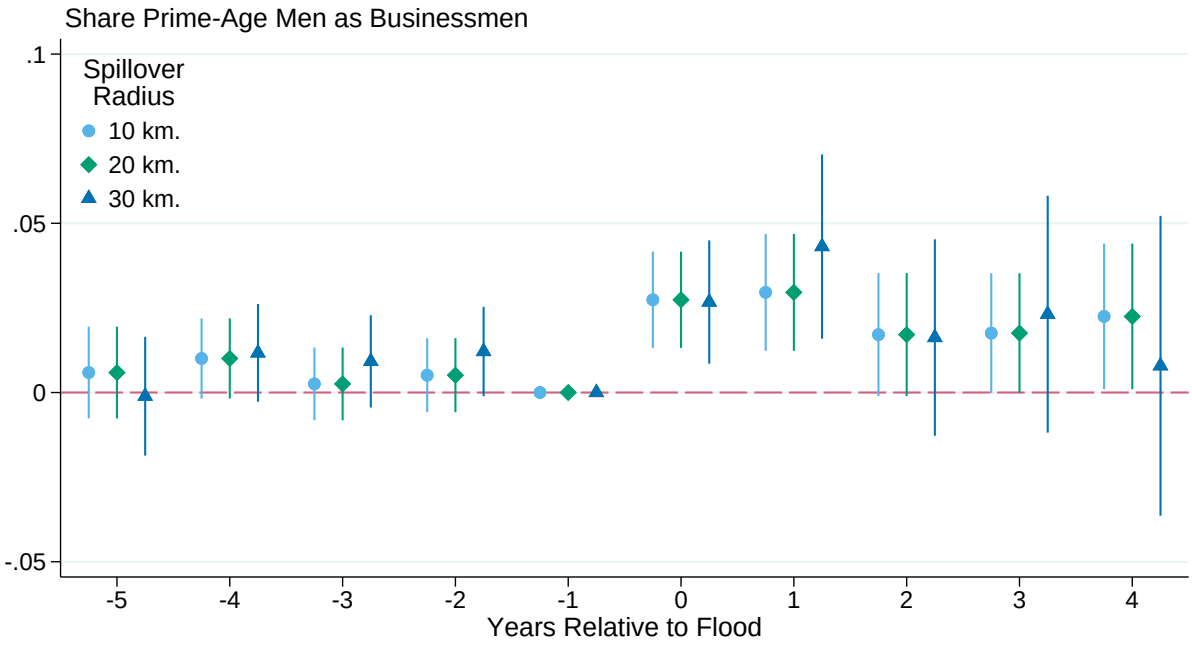
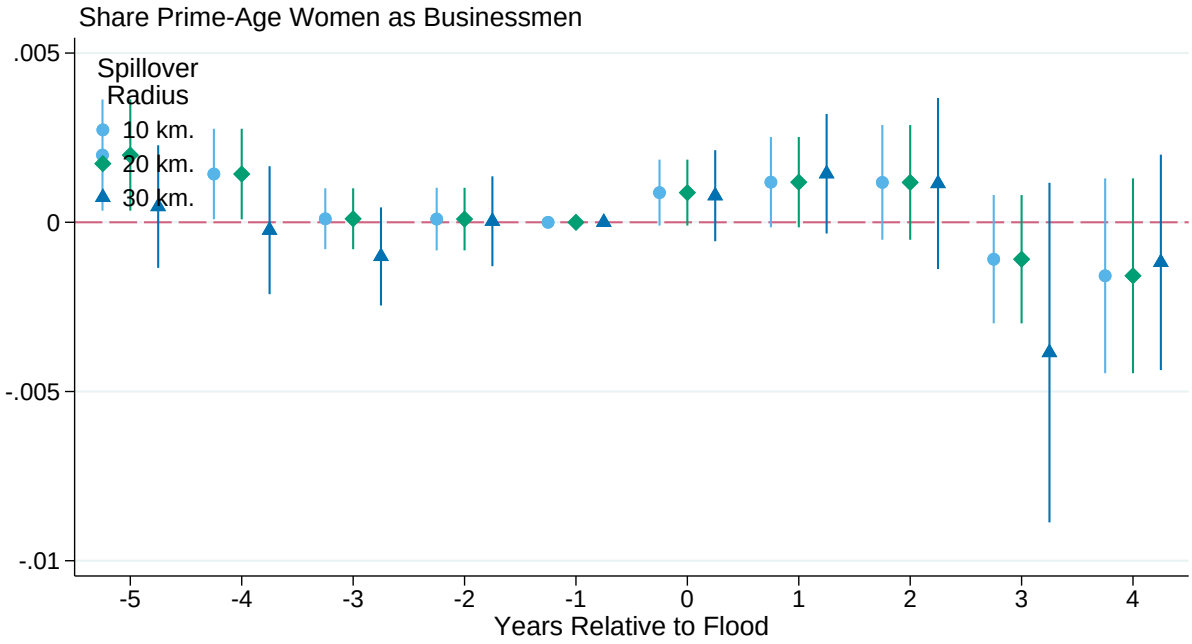


Figure 11: The Marginal Impact of Floods on Business Employment—Women



Note: This figure plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. All regressions include age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.

Figure 12: The Marginal Impact of Floods on Office Employment—Men

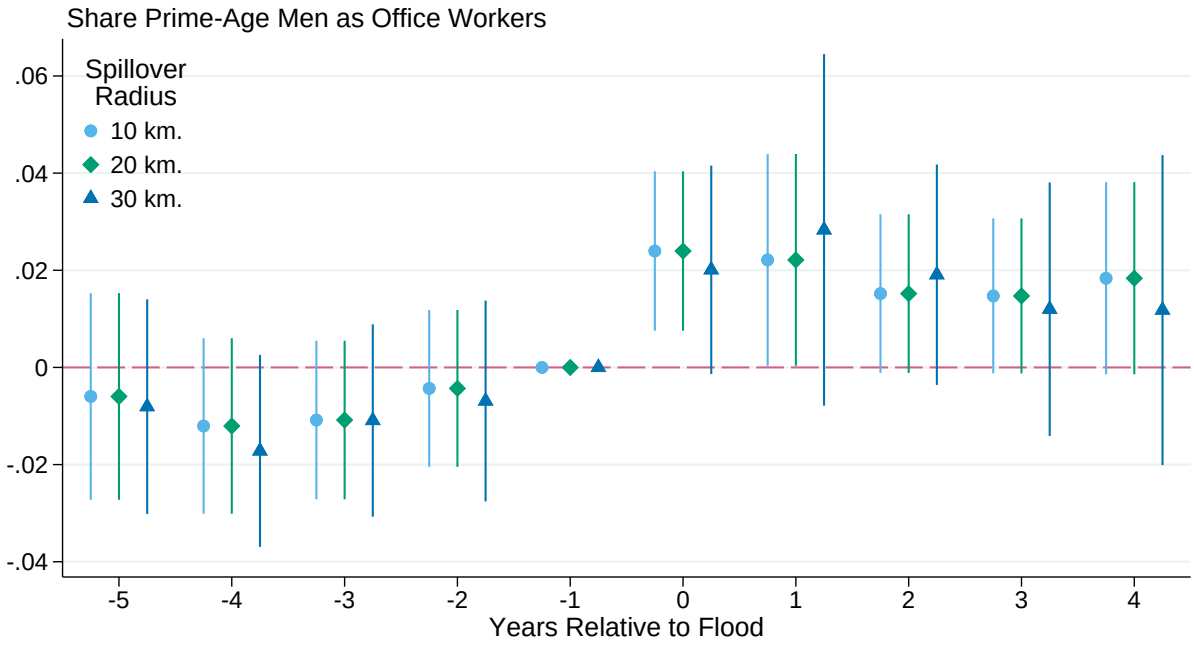
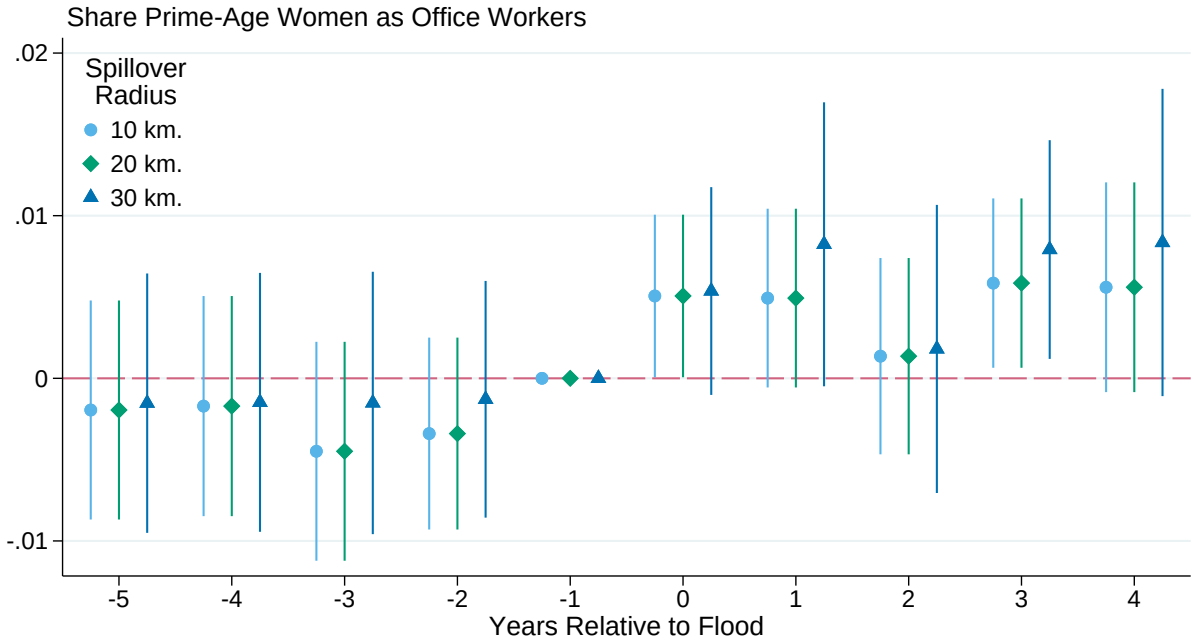


Figure 13: The Marginal Impact of Floods on Office Employment—Women



Note: This figure plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. All regressions include age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.

Figure 14: The Marginal Impact of Floods on Child Schooling Status—Boys

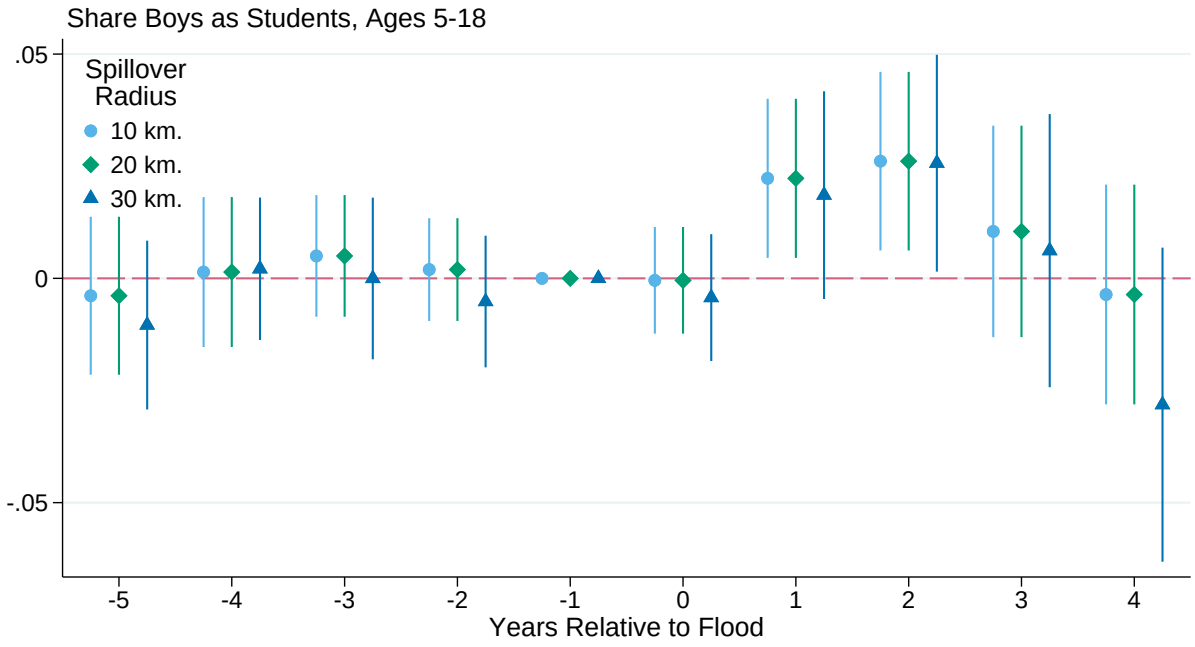
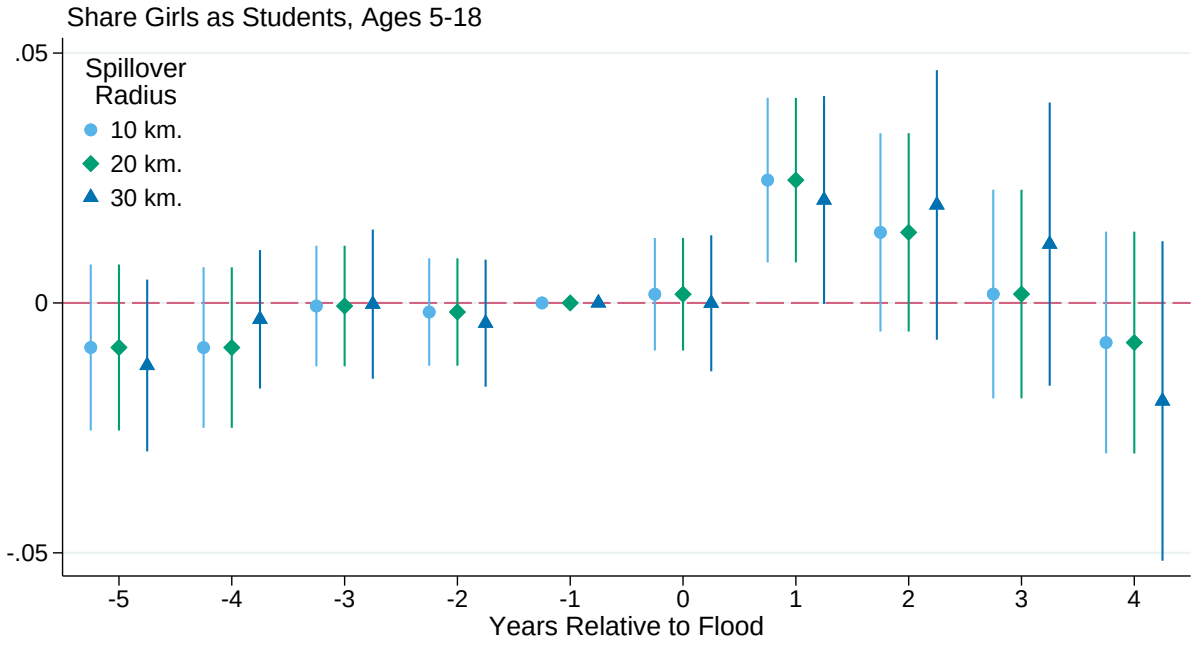


Figure 15: The Marginal Impact of Floods on Child Schooling Status—Girls



Note: This figure plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. All regressions include age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.

Table 5: Impact on Agricultural Employment

	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	-0.0168***	-0.0168***	-0.0168***	-0.0168***	-0.0171***	-0.0178***
	(0.00473)	(0.00452)	(0.00473)	(0.00452)	(0.00565)	(0.00552)
Sample	All	All	All	All	All	All
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.163	.163	.163	.163	.171	.171
Observations	9,683,696	9,683,696	9,683,696	9,683,696	5,338,387	5,338,387
Clusters	1,044	1,044	1,044	1,044	783	783
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	-0.0211***	-0.0213***	-0.0211***	-0.0213***	-0.0215***	-0.0238***
	(0.00579)	(0.00555)	(0.00579)	(0.00555)	(0.00709)	(0.00700)
Sample	First	First	First	First	First	First
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.16	.16	.16	.16	.169	.169
Observations	7,533,172	7,533,172	7,533,172	7,533,172	3,926,172	3,926,172
Clusters	883	883	883	883	643	643

Note: Table 5 presents stacked difference-in-differences estimates of the impact of flooding on whether prime-age individuals are working in agriculture. All specifications include union fixed effects, stratum by relative time fixed effects, age fixed effects, and gender. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 6: Impact on Office Employment

	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.0135***	0.0126***	0.0135***	0.0126***	0.00918**	0.00855**
	(0.00406)	(0.00388)	(0.00406)	(0.00388)	(0.00384)	(0.00361)
Sample	All	All	All	All	All	All
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.084	.084	.084	.084	.081	.081
Observations	9,683,696	9,683,696	9,683,696	9,683,696	5,338,387	5,338,387
Clusters	1,044	1,044	1,044	1,044	783	783
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.0158***	0.0152***	0.0158***	0.0152***	0.0115***	0.0112***
	(0.00529)	(0.00495)	(0.00529)	(0.00495)	(0.00445)	(0.00424)
Sample	First	First	First	First	First	First
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.087	.087	.087	.087	.082	.082
Observations	7,533,172	7,533,172	7,533,172	7,533,172	3,926,172	3,926,172
Clusters	883	883	883	883	643	643

Note: Table 6 presents stacked difference-in-differences estimates of the impact of flooding on whether prime-age individuals are working in an office. All specifications include union fixed effects, stratum by relative time fixed effects, age fixed effects, and gender. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 7: Impact on Businessperson Employment

	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.0119***	0.0119***	0.0119***	0.0119***	0.0103***	0.00939**
	(0.00332)	(0.00332)	(0.00332)	(0.00332)	(0.00371)	(0.00371)
Sample	All	All	All	All	All	All
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.121	.121	.121	.121	.112	.112
Observations	9,683,696	9,683,696	9,683,696	9,683,696	5,338,387	5,338,387
Clusters	1,044	1,044	1,044	1,044	783	783
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.0145***	0.0148***	0.0145***	0.0148***	0.0111**	0.0113**
	(0.00411)	(0.00405)	(0.00411)	(0.00405)	(0.00447)	(0.00447)
Sample	First	First	First	First	First	First
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.124	.124	.124	.124	.114	.114
Observations	7,533,172	7,533,172	7,533,172	7,533,172	3,926,172	3,926,172
Clusters	883	883	883	883	643	643

Note: Table 7 presents stacked difference-in-differences estimates of the impact of flooding on whether prime-age individuals are working as businesspersons. All specifications include union fixed effects, stratum by relative time fixed effects, age fixed effects, and gender. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 8: Impact on Children's Schooling

	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.00743	0.00764	0.00743	0.00764	0.000763	0.00142
	(0.00502)	(0.00503)	(0.00502)	(0.00503)	(0.00553)	(0.00562)
Sample	All	All	All	All	All	All
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.691	.691	.691	.691	.691	.691
Observations	7,667,367	7,667,367	7,667,367	7,667,367	4,307,470	4,307,470
Clusters	1,044	1,044	1,044	1,044	783	783
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.00245	0.00419	0.00245	0.00419	-0.00000215	0.000886
	(0.00680)	(0.00668)	(0.00680)	(0.00668)	(0.00691)	(0.00717)
Sample	First	First	First	First	First	First
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.693	.693	.693	.693	.694	.694
Observations	5,816,317	5,816,317	5,816,317	5,816,317	3,123,769	3,123,769
Clusters	883	883	883	883	643	643

Note: Table 9 presents stacked difference-in-differences estimates of the impact of flooding on whether children are in school. All specifications include union fixed effects, stratum by relative time fixed effects, age fixed effects, and gender. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 9: Impact on Net Migration

	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.00179	0.00176	0.00179	0.00176	0.00213	0.00175
	(0.00146)	(0.00145)	(0.00146)	(0.00145)	(0.00179)	(0.00190)
Sample	All	All	All	All	All	All
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.012	.012	.012	.012	.012	.012
Observations	24728482	24728482	24728482	24728482	13745606	13745606
Clusters	1,044	1,044	1,044	1,044	783	783
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	0.00364**	0.00330**	0.00364**	0.00330**	0.00455**	0.00389*
	(0.00171)	(0.00168)	(0.00171)	(0.00168)	(0.00199)	(0.00219)
Sample	First	First	First	First	First	First
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.013	.013	.013	.013	.013	.013
Observations	19041806	19041806	19041806	19041806	10058260	10058260
Clusters	885	885	885	885	645	645

Note: Table 9 presents stacked difference-in-differences estimates of the impact of flooding on net migration flows. All specifications include union fixed effects, stratum by relative time fixed effects, age fixed effects, and gender. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 10: Impact on Marriage—Women

	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	-0.0137**	-0.0129**	-0.0137**	-0.0129**	-0.0221***	-0.0202***
	(0.00577)	(0.00558)	(0.00577)	(0.00558)	(0.00611)	(0.00582)
Sample	All	All	All	All	All	All
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.692	.692	.692	.692	.711	.711
Observations	3,253,233	3,253,233	3,253,233	3,253,233	1,783,091	1,783,091
Clusters	1,044	1,044	1,044	1,044	783	783
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flood	-0.0165**	-0.0158**	-0.0165**	-0.0158**	-0.0246***	-0.0235***
	(0.00727)	(0.00684)	(0.00727)	(0.00684)	(0.00746)	(0.00684)
Sample	First	First	First	First	First	First
Weights	Propensity	None	Propensity	None	Propensity	None
Spillover Radius	10 km.	10 km.	20 km.	20 km.	30 km.	30 km.
Control Mean	.687	.687	.687	.687	.705	.705
Observations	2,525,621	2,525,621	2,525,621	2,525,621	1,315,257	1,315,257
Clusters	883	883	883	883	643	643

Note: Table 10 presents stacked difference-in-differences estimates on the ever married status of women. All specifications include union fixed effects, stratum by relative time fixed effects, and age fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table 11: Impact of Flooding on Flood Risk Perceptions

	(1) Expected Flood Days Next Year (P)	(2) Expected Flash Flood Crop Share Damage (M)	(3) Expected Flood Crop Damage Next Year (PM)
Flooded in 2022	3.106**	0.0322***	1.290*
	(1.286)	(0.0108)	(0.709)
Non-Flooded Mean	2.244	0.040	2.599
Observations	2036	2004	1962
Clusters	250	250	250

Note: Table 11 shows the impact of experiencing a flood on farmers' beliefs about flood risk. Regressions include fixed effects for flood risk ventile and survey round fixed effect for when the question was asked. Those who fail belief elicitation comprehension checks are excluded. All standard errors are clustered at the union level.

Table 12: Perceived Flood Risk and Expectations about Children's Future

	(1) Expected Years of School	(2) Expect To Hold Salaried Job	(3) Expect To Work As Farmer
Perceived Flood Hazard	0.663*	0.305	-0.0917
	(0.368)	(0.274)	(0.100)
Non-Flooded Mean	12.950	4.834	0.533
Observations	1643	1712	1712
Clusters	810	826	826

Note: Table 12 presents ordinary least squares regressions relating parents' perception of the likelihood of flooding with their expectations about their children's future educational attainment and occupation. All specifications control for objective flood risk ventiles, child's gender, child's current education, and child's current age. The regression is at the child level, where children are all family members aged 23 or younger at the time of the survey. These questions were asked only during the baseline survey to a random half of respondents.

Table 13: Migration Costs as Mediator for Exposure Effects

	(1)	(2)	(3)
Post × Flood	-0.0223***	-0.0145***	-0.0230***
	(0.00611)	(0.00229)	(0.00263)
Post × Flood × Past Exp.	0.00343	0.00614***	0.00921***
	(0.00508)	(0.00150)	(0.00181)
Post × Flood × Log(Dist.)	-0.00234	0.000627	0.00182***
	(0.00183)	(0.000522)	(0.000493)
Log(Dist.) × Past Exp.	-0.00315***	-0.00135**	0.000841
	(0.000972)	(0.000590)	(0.000677)
Post × Flood × Past Exp. × Log(Dist.)	0.00234**	-0.0000454	-0.000551**
	(0.00100)	(0.000228)	(0.000252)
Non-Flooded Mean	-0.140	-0.156	-0.156
Observations	Synth.	Prop.	None
Clusters	59040	117504	117504
	2738	3436	3436

Note: Table 13 presents regressions on the impact of floods on median nighttime luminosity, interacting treatment with past exposure to floods (defined as the number of floods experienced in the preceding five years), and distance to the straight line drawn on the Colonial transportation network. All specifications include union and stratum by relative time fixed effects.

Appendices

Table of Contents

A Flood Detection Details, Tables, and Figures

This section provides further details on the method I develop to measure flooding and the steps I take to analyze the satellite data.

A.1 Measuring Surface Water from Space using Radar-Based Satellite Data

I use data from a dual-polarization C-band Synthetic Aperture Radar instrument at 5.405GHz (C band) on board the pair of satellites constituting the Sentinel-1 mission carried out by the European Space Agency (ESA) which collects such imagery. I use the S1 Ground Range Detected (GRD) scenes provided by the ESA via Google Earth Engine that have been processed to backscatter coefficients in decibels (dB). This backscatter coefficient σ° captures target backscattering area (radar cross-section) per unit of ground area and has been converted to dB as $10 * \log_{10} \sigma^\circ$. Following the recommendation of the United Nations Office for Outer Space Affairs, I use the "VH" polarization band since it is more sensitive to changes on the land surface. The instruments send pulses of microwave radiation at the earth that are then differentially reflected back to the sensor depending on the physical characteristics of the terrain. Surface water—in contrast to land, vegetation, and buildings, for instance—scatters the radiation, thereby sending far less signal back to the orbiting sensor. With a

spatial resolution of 10 meters, these satellites can detect even small changes in surface water inundation. To ease the computational burden of processing the data, I conduct all analyses at a scale of 90 meters. Temporal coverage of these SAR measures beings October 3^{rd}, 2014. The frequency with which these satellites pass over any given point on the earth varies both across place and over time. I analyze data from every satellite pass for each polygon. I conduct all analysis separately for the ascending and descending directions of satellite orbits, only combining estimates at the very end of the process. Combining backscatter data from ascending and descending directions before processing presents challenges because the signal is not only modified by the topography of the terrain but also by different scattering processes of the surface due to varying acquisition angles, rendering the data not directly comparable (see, for example, (Garioud et al., 2020)).

I hold constant the exact position in space of each satellite over each Bangladeshi union to ensure consistency across time within each polygon. I calculate the orbit that maximizes coverage of the total polygon area, and restricting to only those passes of the satellite from that particular orbit pass. For example, for Holokhana union in the descending direction, the orbit 150 yields the best coverage. Given the small size of unions, almost all polygons are entirely covered by the relevant orbit, and I remove any observations that deviate more than one percent below the 95^{th} percentile coverage value. This process yields the raw data shown in Figures A.8a and A.8b. Before processing the data to identify surface water, I adjust for irregularities in the backscatter values that may distort the results. I address two primary issues. First, SAR data feature a "speckle" effect that degrades the content, similar to the grainy spots on ultrasounds in medicine. I filter out this noise, which stems from the interference of waves. To filter out this speckle and enhance the quality of the overall data, I implement the Lee Sigma algorithm (Lee et al., 2009), which rests on the premise that most pixels are distributed within two standard deviations of the mean. The algorithm classifies speckle as pixels falling outside of this range, where the mean is estimated at the center of a given scanning window based on a Gaussian distribution. The results of this processing can be seen in Figures A.8c and A.8d. Second, variation in the slope of the terrain can overpower differences in the land cover itself when impacting backscatter, requiring a slope correction. I implement the angular-based radiometric slope correction algorithm from (Vollrath et al., 2020) based on a global digital surface model. The resulting data can be seen in Figures A.8e and A.8f. In the case of Bangladesh, these elevation corrections have little consequence, though they impact results more in countries with less flat terrain.

To convert the radar backscatter into a measure of surface water, I use an unsupervised machine learning algorithm to analyze the differential return of the radiation and separate water from land. Given the bimodal distribution of backscatter for an area that includes both water and non-water, I classify surface water based on returned radiation falling below a specific threshold. An important challenge to this approach is that the *levels* of backscatter can fluctuate across space and over time due to, for instance, weather conditions, terrain, or the angle of the satellite. Critically, these effects do not differentially impact the return of radiation from surface water, allowing the *relative* comparison to remain highly accurate. Nevertheless, these effects necessitate endogenously estimating this threshold for each orbit instead of choosing a constant cutoff across areas and days. I use an algorithm that searches for a cutoff that minimizes the intra-class variance in flooded and non-flooded pixels. I implement a variation of the unsupervised algorithm known as Otsu's method in computer

vision and image processing to identify these thresholds. From a high-level, this approach minimizes the intra-class variance of water (denoted with subscript w) and non-water pixels (denoted with subscript n). For a cutoff c in the backscatter distribution, the objective function to be minimized is given by $\sigma^2(c) = \omega_w(c)\sigma_w^2(c) + \omega_n(c)\sigma_c^2(c)$, where σ_w and σ_n are the variances of the water and non-water classes respectively, and ω weights denote probabilities of the two classes computed from the histogram. In the fuzzy Otsu method I implement, I randomly select features after segmenting pixels according to an initial threshold in order to calculate a minimum and maximum threshold for each image and then calculate the midpoint using a fuzzy Gaussian approach to ultimately classify water. I set the initial threshold to -20. Example histograms of the processed values used as inputs to this algorithm can be seen in Figure A.10. The final output from this process is the share of surface water $S_{i,t}$ in each polygon i on each date t that the SAR satellite orbits. In the case of Bangladesh, this radar-based data yield 907,394 polygon-by-day observations of surface water inundation from ascending orbits and 1,013,675 observations from descending ones. The density of this measure varies considerably over time: figure A.6 plots the average monthly frequency per polygon from the beginning of the Sentinel-1 satellite coverage.

A.2 Satellite Data and Machine Learning for Predicting Surface Water

As the first data source for the components of the vector $\mathbf{X}_{i,t}$, I use data collected via the Moderate Resolution Imaging Spectroradiometer (MODIS) instruments onboard NASA's Aqua and Terra satellites. In the case of Bangladesh, Terra orbits such that it passes over the country from north to south in the morning, while Aqua does so in the opposite direction in the afternoon. Figure A.9 presents images from MODIS for same example polygon on the same days as does Figure A.8a for the SAR data. These images highlight both the difficulty in relying on optical sensors to measure floods as well as the types of signals that nevertheless can be extracted from these data. I use images from Google Earth Engine that have already been corrected for atmospheric scattering absorption. At a 250-meter resolution, these images have information on reflectance values for visible and near-infrared wavelengths. I first pan-sharpen all other bands down to the 250-meter level using a corrected reflectance algorithm. Then for each image and polygon, I obtain information on each of seven wavelengths, the ratio of red and near-infrared bands, cloud shadow, and the average of the cloud state assessment, which for each pixel is a categorical variable produced as part of the data product. For these 10 variables, I also calculate the 10^{th} through 90^{th} deciles within each polygon. Additionally, I construct the standard measure of surface water based on optical photos from the remote sensing literature (Tellman et al., 2021) that uses a specific functional form to combine information from different wavelengths, as well as versions of that same indicator where I relax that functional form assumption in different ways. This formula uses the reflectance values from band 1 (near infrared NIR), band 2 (red RED), and band 7 (short-wave infrared $SWIR$). This yields three conditions: (1) $RED < 2,027$, (2) $\frac{NIR+13.5}{RED+1,081.1} < 675$, and (3) $SWIR < 675$. To produce alternate versions of this indicator, I scale the key thresholds by .8 and 1.2. This ultimately yields 103 variables per image, or 206 variables per polygon per day. To address days with missing data, I create indicator

variables for missingness for each variable, and replace missing values with the variable's mean. I also include five-days-worth of lags and leads for each of these values.

For the second component of the vector $\mathbf{X}_{i,t}$, I use data from NASA's Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) version 6. For further details on the construction of these estimates, see (Huffman, 2019). Although the underlying model does include many radar based satellites, it does not incorporate the Sentinel-1 synthetic aperture radar images used to estimate $A_{i,t}$ above. I calculate the total amount of daily rainfall, the number of 30 minute blocks without precipitation, the maximum precipitation rate in any given 30 minute period, and the standard deviation of precipitation during the each day. For these four variables, I also create the 14-days-worth of leads and lags for each union and date.

From this set of 5,562 features, I take two steps to reduce the size of the vector to increase computational speed in the cross-validation. In practice, the results are quite similar when estimating the gradient boosted regression tree using the full set of possible predictors. First, I drop all predictors with variances below .05 within the training sample. Second, I calculate the cross-correlation across all features, and drop those with correlations above .95. After these steps, the final vector $\mathbf{X}_{i,t}$ in the case of Bangladesh, for instance, typically consists of approximately 700 features, depending on the specific division. Note that I do not include polygon fixed effects or any explicit indication of seasonality.

In the Bangladesh example, I estimate the prediction models separately within each of Bangladesh's seven divisions to ease the computational demands from these algorithms. Note that a pooled model with the proper inclusion of division identifiers must perform at least as well. I use the scikit-learn package to implement the algorithms (Pedregosa et al., 2011). As discussed above, my preferred approach predicts the continuous measure of surface water to allow for considerably more flexible seasonal adjustment of permanent water, but I also report the prediction qualities for the binary case (in which I convert the radar-based data to flooding prior to the supervised machine learning) below. I compare several different algorithms to find the best prediction function $g(\cdot)$. To evaluate the quality of each approach indexed by p, I split data into a training dataset and an evaluation dataset and compare each algorithm's performance $\widehat{A^p}_{i,t}$ to the truth $A_{i,t}$ on two metrics: mean squared error and R^2. In the binary prediction problem, I assess performance based on accuracy and area-under-the-curve statistics. In the case of Bangladesh, I find that a histogram gradient boosted regression tree performs best. I additionally estimate alternative algorithms, including support vector machines, a multi-layer preceptron neural network, and a stochastic gradient descent linear model for both the binary and continuous prediction problems. The gradient boosted regression tree substantially outperforms the others in the continuous task, while the out-sample quality remains roughly equivalent across algorithms in the binary case. At a high level, histogram gradient boosted regression tree that performs best begins with a decision tree: a model that iteratively partitions the data for prediction. In the boosted version, models sequentially learn from the mistakes of previous models by focusing on the remaining set of difficult-to-predict observations at each step (the residuals in the continuous case). The histogram-based version of this model significantly speeds up the training process by binning continuous features into fixed buckets. Using a gridded search with the squared-error as the loss function, the hyperparameters I tune include the maximum number of leaves for each tree ($\{63, 75, 100\}$), the number of bins for each feature ($\{50, 75, 100\}$), and

the L_2 regularization parameter ($\{.1, .2\}$). I do not constraint the depth of each tree and set the shrinkage rate to be .1.

In this setting, two sources of correlation generate near-duplicates in the data for which the procedure needs to account or else risk effective test-sample leakage. First, neighboring polygons exhibit substantial spatial covariance. Second, within polygon, neighboring days feature strong temporal correlation. I take two steps to account for these patterns in my estimation procedure. First, when tuning the parameters of the supervised machine learning algorithms, I cross-validate at the polygon-month level. Second, to evaluate the overall performance, I assess the metrics using three different hold-out samples. The first is a "pure" random sample at the polygon-day level. The second is a "polygon" sample in which I exclude every observation of randomly selected polygons. Finally, I construct a "month" sample that omits all observations from randomly selected months across all polygons. I select a random 20 percent of union-days for the pure sample, 10 percent of unions for the polygon sample, and 10 percent of months for the month sample. The training sample of the algorithm excludes the union of all three hold-out samples, and I report statistics for each.

The results from Bangladesh show the algorithm performs extremely well. Figure A.13 plots the R^2 and area-under-the-curve for each model at the division-by-orbit-direction level across each hold-out-sample. As expected, the models performs best in the pure polygon-day test set, among which the near-duplicates issue poses the greatest threat. Observations in the hold-out sample removing entire months tend to have lower accuracy, though even in these tests, the algorithms explain the vast majority of variation. Tables A.1 and A.2 presents the underlying statistics behind these figures. As a benchmark, I also include the equivalent performance metrics from ordinary-least-squares and logistic regressions for the continuous and binary cases, respectively. The histogram gradient boosted decision tree on average explains 89.86 percent of the variation in surface water in the random polygon-days test sample, 83.82 percent in the random polygon test sample, and 81.28 percent in the random month test sample. This signifies a substantial gain above linear regression, which explains 53.23 percent, 48.03 percent, and 45.58 percent in the corresponding cases. In the binary prediction problem the supervised machine learning algorithm achieves an area-under-the-curve score of about 0.65 in the random polygon-day and random polygon hold-out sets, and about .5 in the months hold-out. This pattern likely reflects that by converting the continuous measure to a binary flood indicator equipped only with the radar-based data to remove seasonal trends introduces a significant amount of noise that makes the prediction task more difficult.

Are these statistical performances "good?" Despite the accuracy of the Sentinel instruments, there nevertheless exists noise in the "ground-truth" surface water to be predicted by these models. This includes variation generated from the processing of data, such as the random draws of the Otsu method, as well as at the data collection stage itself. Formally, following equation 2, the radar-based surface water measures includes the term $\varepsilon_{i,t}$. Therefore, perfect explanatory power featuring an R^2 equal to 1 is an impossibly high bar. To benchmark these evaluation metrics against more plausible performance standards, I take two complementary approaches. First, I use the fact that I observe orbiting satellites collecting the radar-based instruments in both the ascending and descending directions. Although these measures cannot directly be compared in levels due to a host of differences in backscatter patterns depending on the orbit, the correlation of these two independent datasets never-

theless is illustrative of the amount of noise in any given polygon pass. Among all 57,629 overlapping polygon-days in Bangladesh, the correlation between the surface water measurements calculated from the ascending and descending orbits is 0.78. Using a simple linear regression, I find that measures of surface water collected by satellites orbiting in the ascending direction explain 61.40 percent of the variation in surface water collected by descending satellites. While merely suggestive, these statistics suggest an important baseline amount of noise in the radar-based surface water measurement. Second, I compare the performance of these models to similar papers in the literature that combine machine learning with remote sensing data or related inputs. My predictive power typically meets or exceeds what other researchers have found. For example, Hsiao (2023) uses similar machine learning methods to estimate flood risk in Jakarta, finding an R^2 of .471 in his best performing algorithm (which is also a histogram gradient boosted decision tree). In another case, Tellman et al. (2021) use optical images from the MODIS sensors to map the extent of 913 flooding events from the Dartmouth Global Flood Database. In a validation test against Landsat scenes for 123 days of maximum inundation, they find a mean accuracy of 83%. Outside of flooding, Rolf et al. (2021) combine machine learning with satellite imagery across many domains, finding an R^2 of .91 for forest cover, .68 for elevation, .72 for population density, .85 for nighttime lights, .45 for income, .53 for road length, and .52 for house prices.

When does the algorithm perform worst? To examine this in the case of Bangladesh, I pool the three hold-out samples and assess determinants of the prediction error. I run linear regressions with the outcome equal to the difference between the predicted value and radar-based measure as a share of the truth, restricting the sample to those polygon days with at least one percent surface water to exclude extreme outliers. Equation 14 presents this specification, where $D_{i,t}$ denotes the determinant of interest. I examine year, calendar-month fixed effects, and polygon size.

$$\frac{|g(\mathbf{X}_{i,t}) - A_{i,t}|}{A_{i,t}} = \alpha + \beta D_{i,t} + \epsilon_{i,t} \tag{14}$$

The results of this exercise show mixed evidence of systematic biases in prediction accuracy. Polygon size cannot predict the gap between predicted and true surface water (p-value $= 0.539$ for ascending orbits, p-value $= 0.522$ for descending ones). Observations in June seem to exhibit differentially more error than other months across orbit directions. Across years, accuracy exhibits no systematic bias, with the exception of 2017 which features a significantly higher gap. With more advanced supervised machine learning methods than those used here, these patterns may disappear.

At the conclusion of this process, I arrive at a prediction model for the radar-based surface water measurement $A_{i,t}$ using exclusively the high-frequency remote sensing data $\mathbf{X}_{i,t}$. Applying this same model out-of-sample on days without orbiting SAR instruments, I construct a daily panel of surface water at the polygon level. I repeat this entire procedure separately for the radar-based data collected during descending and ascending orbits.

A.3 Alternative Approaches to Removing Permanent Water

Isolating deviations in surface water is a key step in identifying floods. Here, I discuss how my method differs from previous attempts to remove this seasonality, which have been limited

by data. The dominant approach has been to use the Global Surface Water data (Pekel et al., 2016) on the location of permanent water. These measures use optical images from the Landsat series to detect water, thereby falling prey to the same issues with non-random cloud cover motivating the use of radar-based data collection above. Perhaps because of the sparsity of usable images during the rainiest parts of the year, researchers have often turned to averages of annual estimates, thereby ignoring any seasonal shifts in water location. An alternative strategy has been to calculate differences in surface water during a given flooding event of interest by netting out inundated areas from the previous calendar year, where ideally both measures are estimated from radar-based instruments. This relies on the key assumption however that the previous year featured no floods, which again forces researchers to know *ex ante* when and where floods have occurred. The richness of my daily surface water panel allows me to remove seasonality much more flexibly than these previous methods by using the granular historical data for each polygon to flexibly estimate natural shifts in bodies of water both across the calendar year and over time.

Up to this point, I conduct every step separately for radar-based data collected via satellites orbiting in the ascending and descending directions. At this stage, I finally combine these two series by taking the average of the two residuals at the polygon-date level before entering into step #4.

A.4 Validating the Flooding Database

This section describes details on the validation of my new flooding measure.

Assessing Flooding Summary Statistics This dataset allows me to establish facts about flooding in Bangladesh at the union level. Floods are common; 51.20 percent of the 5,158 unions in the country experience at least one flood from mid-2002 through the end of 2022. Conditional on undergoing a flood in a year, unions experienced an average of 3.74 days of inundation, highlighting the need for high-frequency observations to estimate flooding. In all, this period featured 11,251 union-years with at least one flood, 10.39 percent of the total.

As a first validation test, I visualize basic patterns in the flooding data and evaluate whether they pass simple intuitive checks. First, I simply calculate the average flood rate by year and plot the time series, shown in Figure A.15. The peaks in aggregate flooding correspond to precisely the years when major flooding events occurred: the 2007 South Asian Floods, the 2017 Brahmaputra River Floods, and the 2022 India-Bangladesh Floods. Second, I map the geographic distribution of flood rates in Figure A.16. Consistent with local descriptions, the most flooded areas over time tend to fall along the banks of major rivers that can overflow during major monsoon rains. For instance, one can see the outline of the Brahmaputra and Sylhet rivers simply based on the choropleth map. Third, I calculate the probability of a flood across all unions by calendar day. Figure A.17 plots a three-day moving average of this value, clearly showing the increase in flooding events during the monsoon season followed by a collapse during the dry season.

Newspaper Coverage of Flooding I next compare my estimates to floods reported in news articles. I build two databases for this purpose (one rural and one urban) based on online searches for relevant keywords. To build the rural database, I conduct searches for articles mentioning floods with a focus on rural areas in the Khulna division. I use keywords "flood", "tidal surge", "cyclone", "water logging", "storm", and the equivalent Bangla versions, searching both on Google and Google News. I combine these keywords with the English and Bangla version of each union name in the survey sample, and geocode floods at the lowest administrative level possible. The final database includes 277 articles, the causes of which are shown in Table A.3. Because of difficulties in identifying local inundation from news articles, I assign floods to all unions within one of these higher levels. At a union-by-flood level, this dataset includes 7,213 observations, of which 22 percent were originally mentioned at the upazila level in the articles, and 64 percent were originally at the district level. This generates a very liberal measure of flooding. In the urban sample, I focus on the city of Chittagong—the second largest city in Bangladesh behind Dhaka, and one for which online searches for flooding yield much more relevant articles than the country's capital. I conduct the same search parameters combined with "Chittagong", and restricting to articles from 2021 and 2022, track down 30 reports of flooding in the city.

The data exhibit a strong recency bias, as shown in Figure A.19, underscoring the difficult of using news articles to construct a consistent measure of flooding. To test the accuracy of my estimates, I examine how my measure of surface water varies around the dates of news articles. For each newspaper article reporting a flood in union i on date t, I construct a control group using that same union on the same calendar day as t but in a different year when no flooding was reported. I then estimate the event study specification in equation 15 using this pooled sample, where $Article_{i,t}$ captures treatment—whether an article indeed recorded a flood—and d_j denotes relative time to the publication of the article. I include union fixed effects λ_i and calendar day fixed effects ψ_t, and cluster standard errors at the union level.

$$\widehat{A}_{i,t} = \alpha + \sum_{j=-3}^{3} d_j \times Article_{i,t} + \lambda_i + \psi_t + \epsilon_{i,t} \tag{15}$$

Figure A.20 plots the coefficients from this regression for the rural and urban news databases separately. Both samples exhibit an increase in residualized surface water around the time of publication of a news article. In the urban sample, water coverage increases sharply and quickly returns to lower levels. In the rural sample, by contrast, inundation exhibits a more gradual increase and decrease. This pattern matches a scenario in which rural floods both accumulate and drain more slowly while urban areas more likely experience flash floods—consistent with local narratives.

Government Flooding Reports I next compare my measure to flooding events I extract from official reports produced by the Bangladeshi government. I code up records from annual reports of the Flood Forecasting and Warning Centre (FFWC) of the Bangladesh government's Water Development Board. I use all available reports from 2010 to 2020. Within these reports, I first categorize floods based on descriptions of when water levels across river stations clear danger levels. According to the FFWC, the danger level at a given location is the level above which it is likely that the flood may cause damages to nearby

crops and homes. One chapter of each report describes when water levels (WL) crossed these danger levels (DL), organized by basin. From the descriptions of river behavior, I code the flood start date, end date, cause, and the upazila of the river station. For stations within Khulna, additional online research of the flood is conducted in an attempt to identify the specific upazila, union, or village near that river station that experienced flooding, using keywords including "Flood", "Water logging", the year, and location, both in English and in Bangla targeting national and local newspapers. The final dataset includes 616 unique union-date observations. My estimated surface water systematically varies with this sample exactly as one might predict. Figure A.21 presents a binned scatter plot of surface water in a union by the relative day leading up to and following the date when the government declared a flood. Across both the ascending orbit and descending orbit series, surface water significantly increases in the week prior to when the government classifies a flood before leveling off during the inundation event itself.

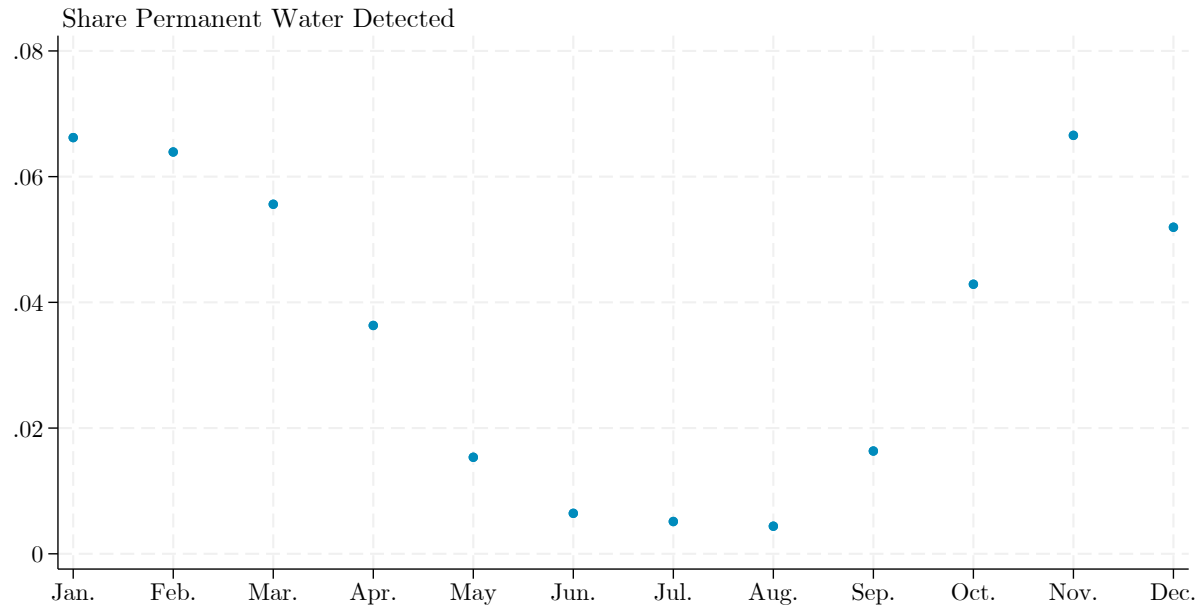
***in situ* Water Level Data** As a final test, I obtain 3-hourly water-level data from 48 river stations in Bangladesh, for a total of 700,813 measurements. Linking these river stations to their corresponding unions, I compare the average height of the river each day to my estimates of surface water in that union. Figure A.22 plots a binned scatter plot of this relationship for the descending orbit measure (the graph looks nearly identical for the ascending orbit data.) The pattern matches a story in which at lower heights, surface water at the union and water level follow each other more closely than at higher heights—exactly what one might expect given the issues with converting these local vertical measures to horizontal inundation coverage.

These data also allow me to examine the scope for another potential concern with my method: differential accuracy over time. The radar-based data suffers from two limitations in its temporal coverage: first, prior to the launch of the satellite at time $\underline{t}$, the data are missing for all periods. Second, even for days after the instrument is active $t \geq \underline{t}$, the orbits are sufficiently infrequent to leave significant gaps between observations. Expanding the frequency of this accurate measurement therefore consists of two problems related to the two eras when the radar-based data $A_{i,t}$ is missing: the *extrapolation* periods prior to the launch of the satellite, and the *interpolation* periods in between orbits. To test whether my predictions perform differentially well depending on the period, I assess whether the relationship between my predicted surface water and river height varies statistically significantly between different years. I estimate the regression from equation 16, where λ_i denotes river station fixed effects, ψ signifies year fixed effects, and γ captures the extent to which river height in the extrapolation period differentially relates to surface water depending on the number of years prior to the Sentinel-1 mission launch. The coefficient ρ serves as a placebo test: detecting a differential relationship after the onset of radar-based data collection. I obtained river height measurements from 2007 and 2012 for the pre-period and from 2017 through 2022 for the post. I cluster standard errors by date.

$$\widehat{A}_{i,t} = \alpha + \beta Height_{i,t} + \gamma Height_{i,t} \times Years\ Until\ SAR_t + \rho Height_{i,t} \times Years\ Since\ SAR_t + \psi_t + \lambda_i + \epsilon_{i,t} \tag{16}$$

The results show statistically significant but small in magnitude difference in the link between

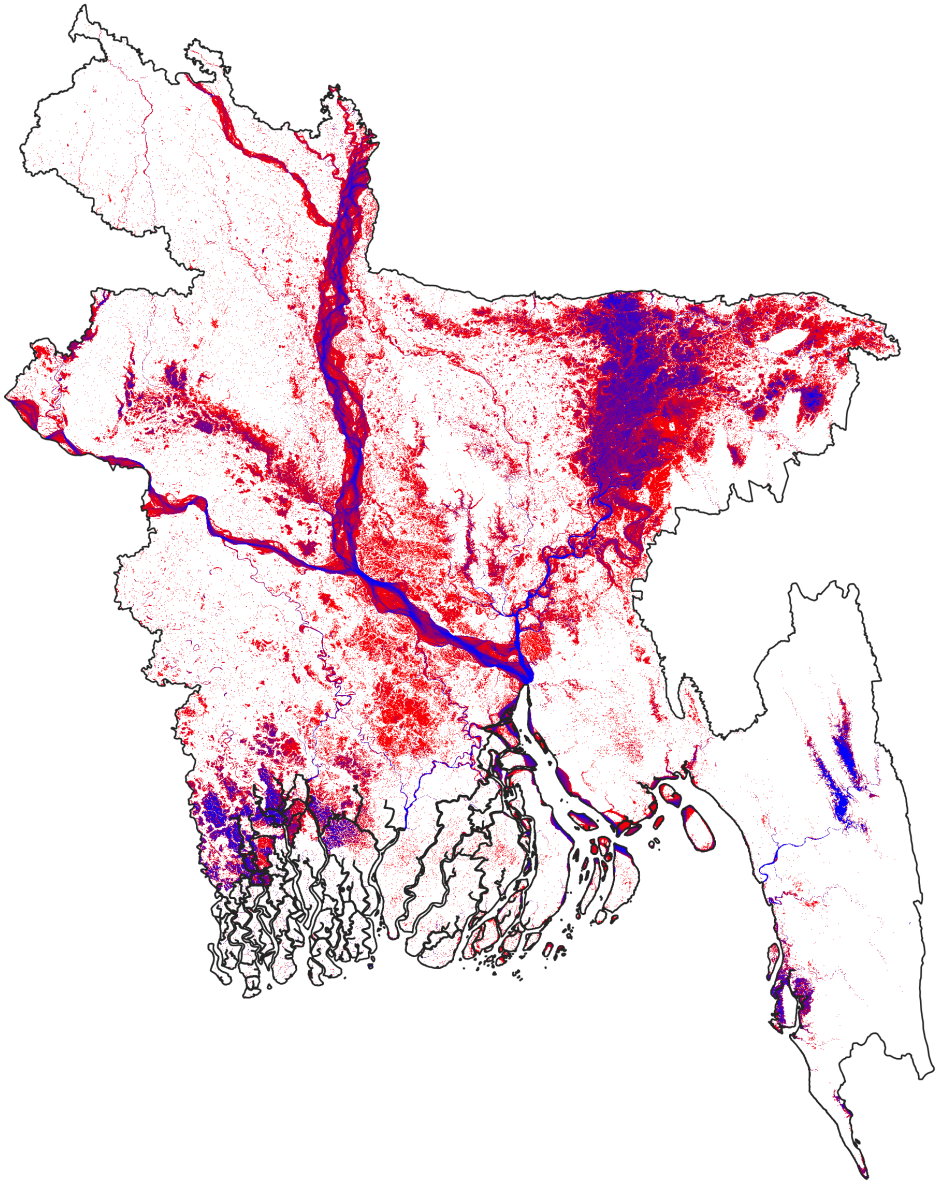
Figure A.1: Share of Permanent Water Bodies Detected using MODIS Images by Month



Note: Figure A.1 plots the average share of permanent water pixels across unions that are accurately detected as surface water using the methods to analyze MODIS data from Tellman et al. (2021), separately by calendar month. I use daily images from the Moderate Resolution Imaging Spectroradiometer onboard both the Aqua and Terra satellites that have been corrected for atmospheric scattering absorption. I pan-sharpen the short-wave infrared band to 250 meters in line with the resolution of the visible and near-infrared wavelengths using a corrected reflectance algorithm. I classify a pixel as water if three conditions are met using the near-infrared band (NIR), red band, and short-wave infrared band (SWIR):$RED < 2,027$, $(NIR + 13.5)/(RED + 1081.1) < 675$, and $SWIR < 675$. To combat the risk of false positives, I construct three-day periods and classify a pixel as water if in half or more images in that set, the three reflection conditions are met. I also exclude pixels that have elevation greater than five degrees for similar false positive reasons.

surface water and river height depending on the years until Sentinel-1's launch. As a share of the baseline linear relationship, one year further away from the SAR instrument data changes the association by 0.092 percent (p-value = 0.000). Although the "placebo" check in years after the launch statistically differs from zero, the effect is considerably smaller (0.013 percent, p-value = 0.054).

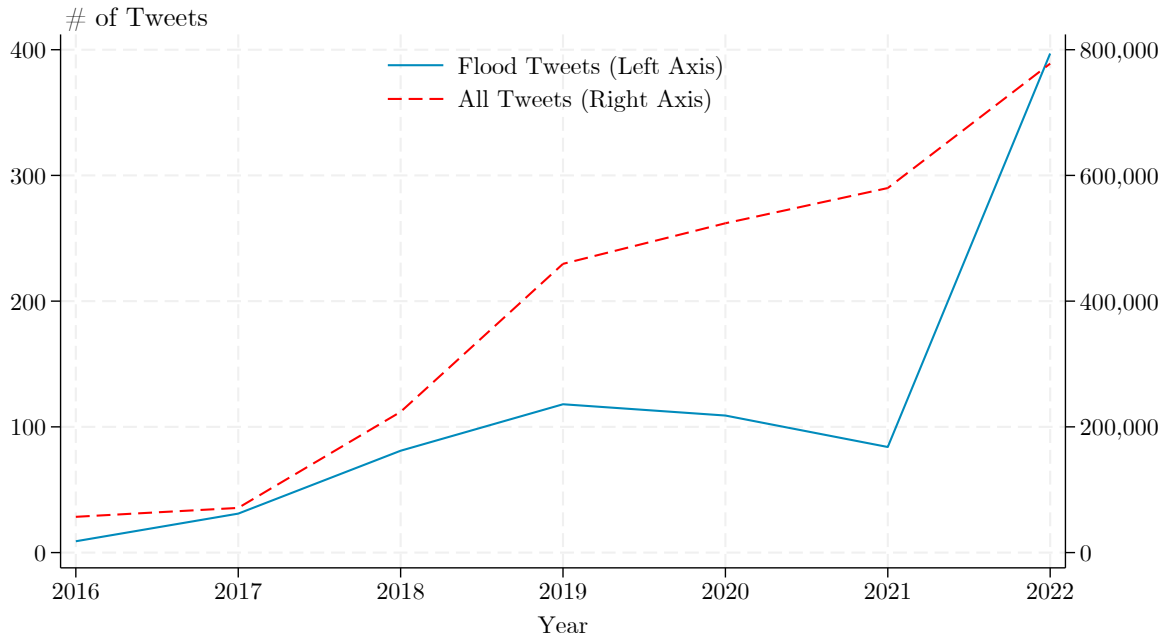
Figure A.2: Seasonal Variation in Surface Water in Bangladesh



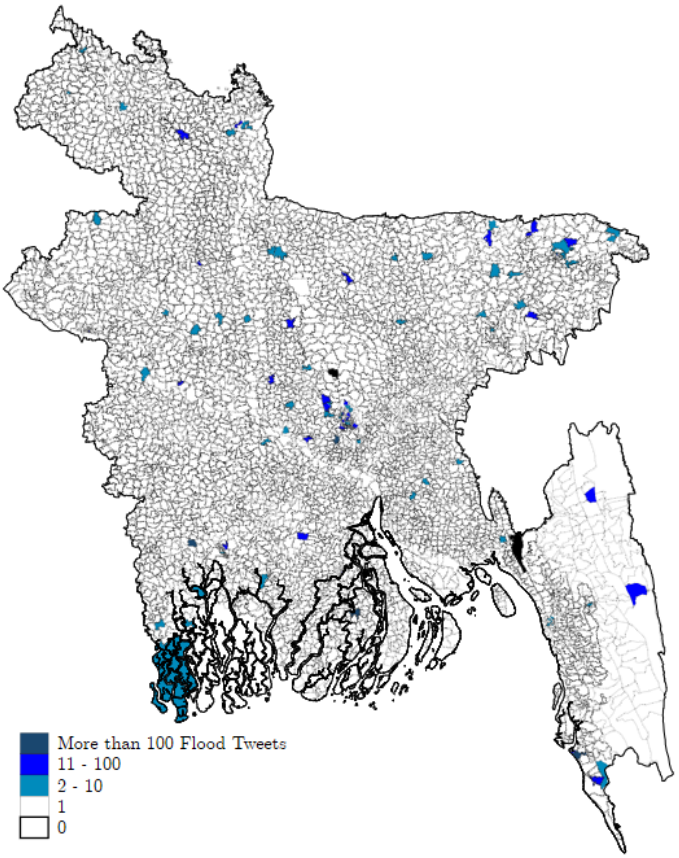
Note: Figure A.2 shows surface water that exists year round in blue against water that only appears seasonally in red for Bangladesh based on the cumulative data from Pekel et al. (2016), in which they analyze yearly LANDSAT composites since 1984.

Figure A.3: Distribution of Public, Geo-Coded Tweets Mentioning "Flood"

(a) Flood Tweets Across Time



(b) Flood Tweets Across Space



Note: These figures are based off of all publicly available, geo-coded tweets in Bangladesh between 2016 and 2022. I define a flood tweet as one that mentions either "flood" or "bonna" (the corresponding Bangla word). Figure A.3a presents the total number of flood tweets in each year. Figure A.3b maps the total number of tweets by union using the coordinates of the tweet.

Figure A.4: Farmers' Perceptions of Flood Risk by Length

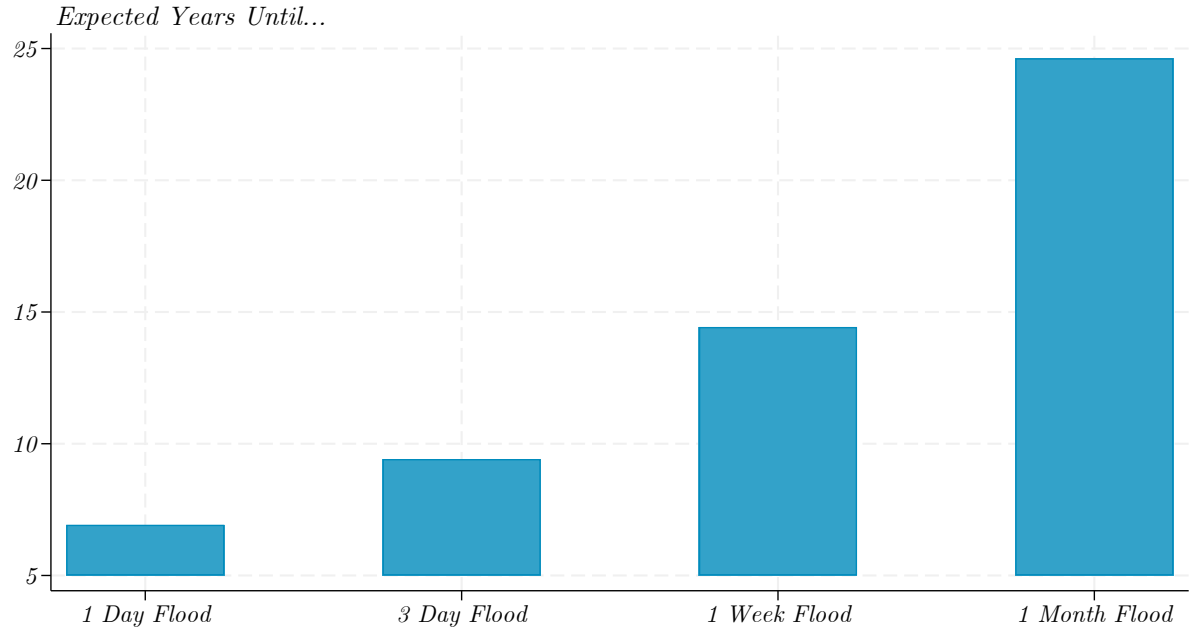
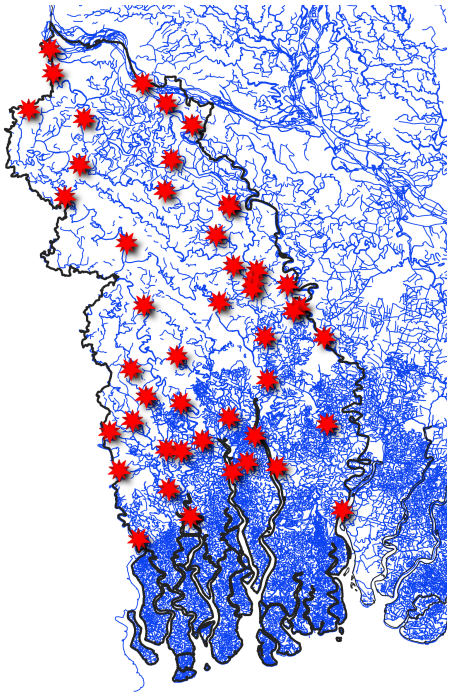


Figure A.5: River Stations in Khulna Division



Note: Figure A.5 maps the rivers and water level stations maintained by the Bangladesh Water Development Board in the Khulna division of Bangladesh. Figure A.4 plots the distribution of farmers' responses to the question, "How many years do you think it would take for a [TIME] long flood to happen in this village?" where [TIME] is one-day, three-day, week, and month.

Figure A.6: Average Ground Truth $A_{i,t}$ Observations per Polygon per Month

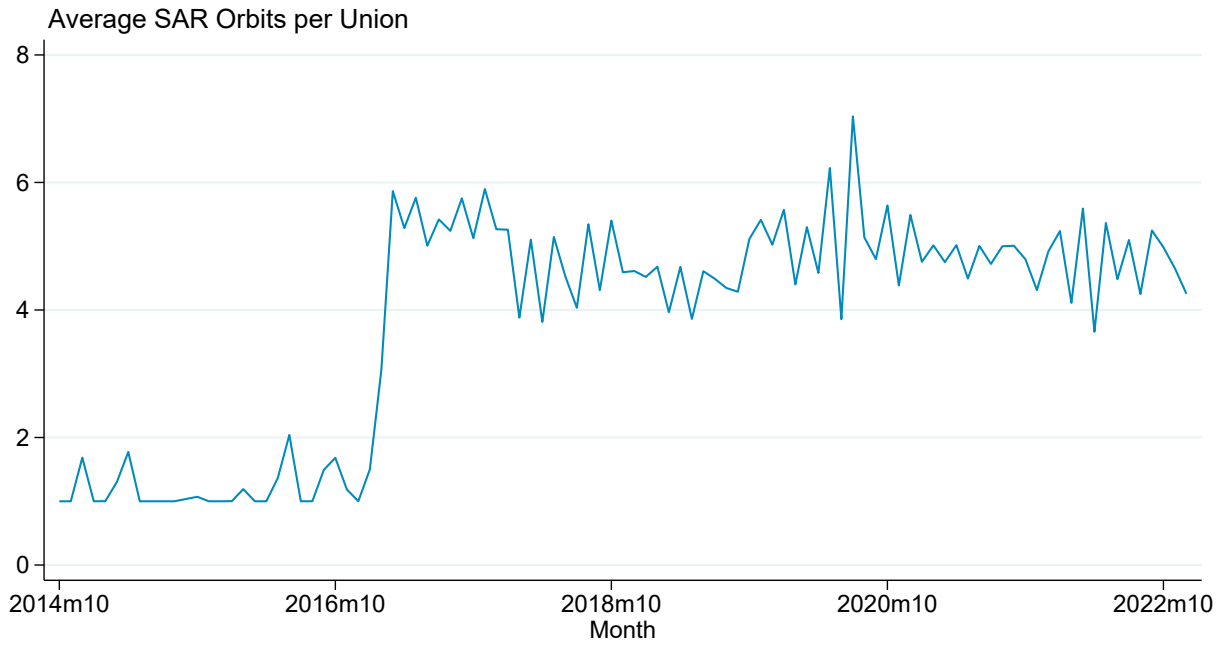
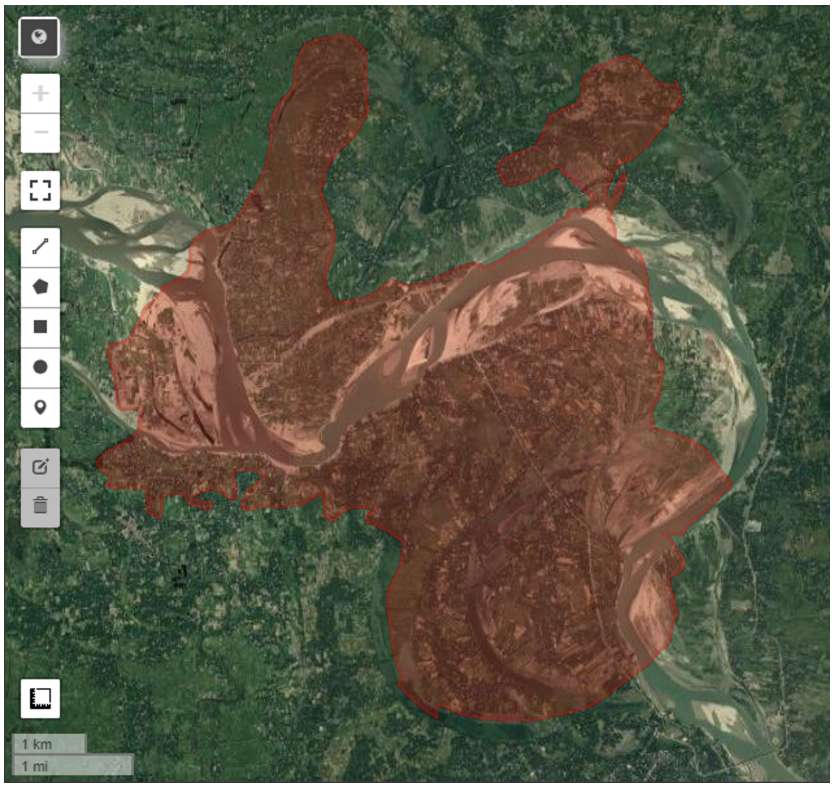


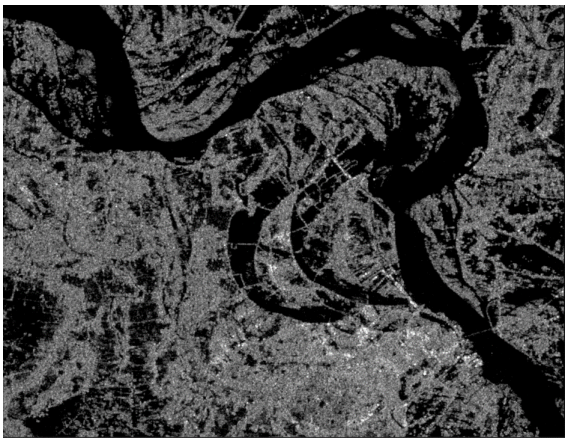
Figure A.7: Sample Union Outline



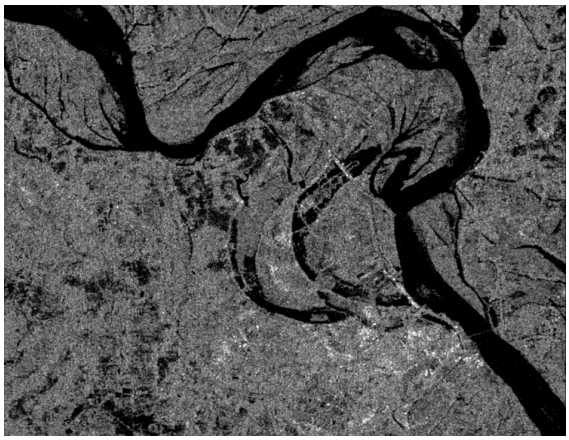
Note: Figure A.6 plots the number of surface water inundation observations by month from the Sentinel-1 satellite series in Bangladesh, taking the average across all polygons. Figure A.7 presents a map of the Holohkana union, which I use as an example in the visualization of the remote sensing data processing.

Figure A.8: Visualization of Example Sentinel-1 VH Polarization Backscatter Processing

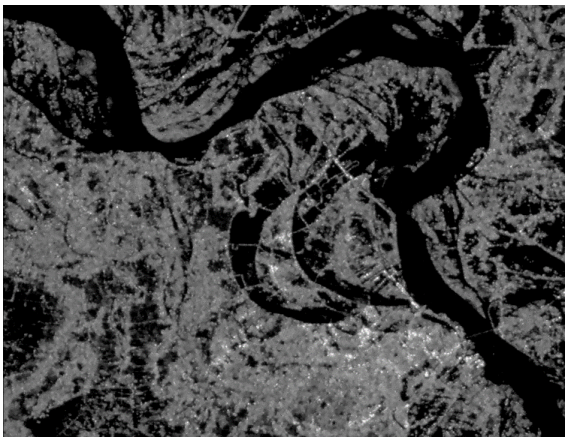
(a) Raw Data—July 27, 2020 (Flood)



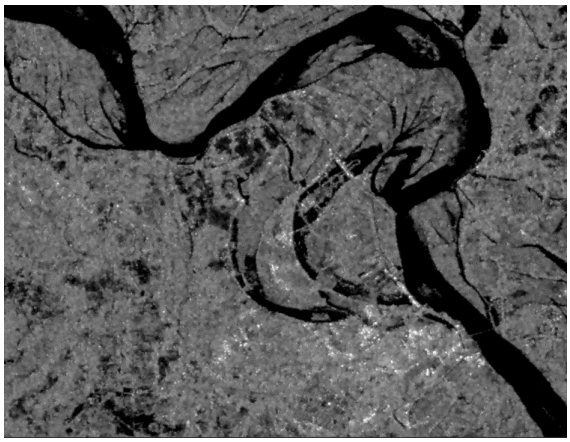
(b) Raw Data—July 22, 2021 (No Flood)



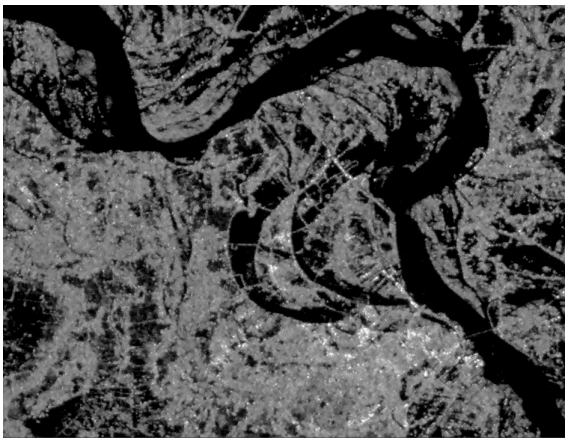
(c) Speckle-Filtration—July 27, 2020



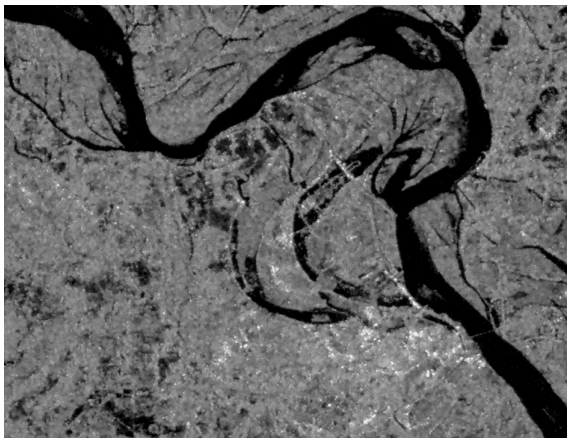
(d) Speckle-Filtration—July 22, 2021



(e) Elevation Adjustment—July 27, 2020



(f) Elevation Adjustment—July 22, 2021



Note: Figure A.8 presents the evolution of the processing of the Sentinel-1 data from the raw "VH" polarization band to the backscatter data ready to be analyzed using the fuzzy Otsu algorithm to identify water. The images are shown for the Holokhana union on two dates: July 27, 2020 during which the area experienced intense flooding, and July 22, 2021 during which it did not for comparison. Figures A.8a and A.8b present the raw values for 2020 and 2021, respectively. Figures A.8c and A.8d present the data after filtering out speckle using the Lee Sigma algorithm for 2020 and 2021, respectively. Figures A.8e and A.8f visualize the data after implementing the angular-based radiometric slope correction algorithm for 2020 and 2021, respectively.

Figure A.9: Visualization of Example Optical Image Data

(a) Red Band from MODIS Aqua Satellite—July 27, 2020 (Flood)

(b) Red Band from MODIS Aqua Satellite—July 22, 2021 (No Flood)

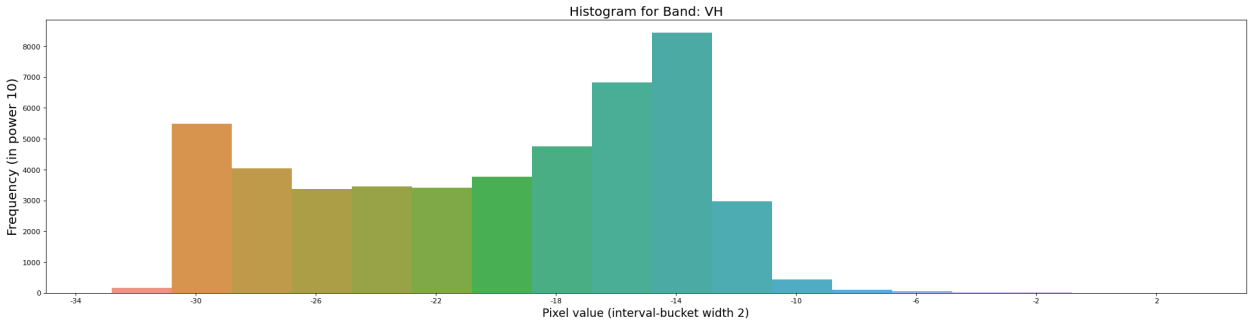
(c) Red Band from MODIS Terra Satellite—July 27, 2020 (Flood)

(d) Red Band from MODIS Terra Satellite—July 22, 2021 (No Flood)

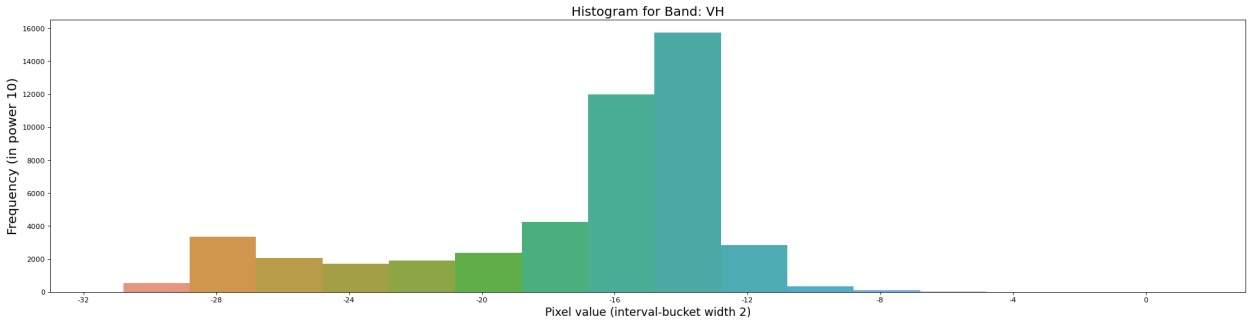
Note: Figure A.9 presents visualizations of the red band from MODIS instruments passing over the Holokhana union on two dates: July 27, 2020 during which the area experienced intense flooding, and July 22, 2021 during which it did not for comparison. Figures A.9a and A.9b show data from the Aqua satellite, which passes over Bangladesh from south to North in the afternoon, while figures A.9c and A.9c stem from the Terra satellite, which orbits in the opposite direction in the morning. Note that figure A.9c appears empty because all data is missing. The pixel resolution is 250 meters, and the data have been corrected for atmospheric scattering, but otherwise these images show the raw reflectance values.

Figure A.10: Example Histograms of Backscatter Values

(a) Histogram of Backscatter Values—July 27, 2020 (Flood)



(b) Histogram of Backscatter Values—July 22, 2021 (No Flood)



Note: Figure A.10 presents histograms of the backscatter values after undergoing speckle filtration and angular-based radiometric slope correction for Holokhana union on two example dates: July 27, 2020 during which the area experienced intense flooding as shown in Figure A.10a, and July 22, 2021 during which it did not for comparison as shown in Figure A.10b.

Figure A.11: Sample River Station Water Level Seasonality Pattern

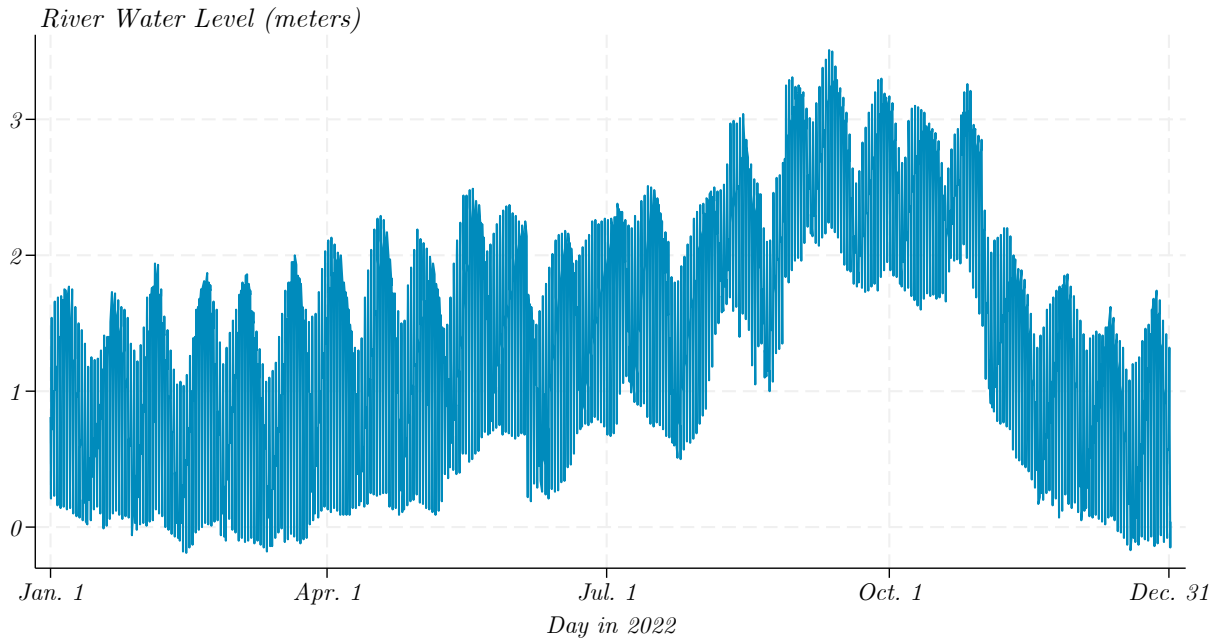
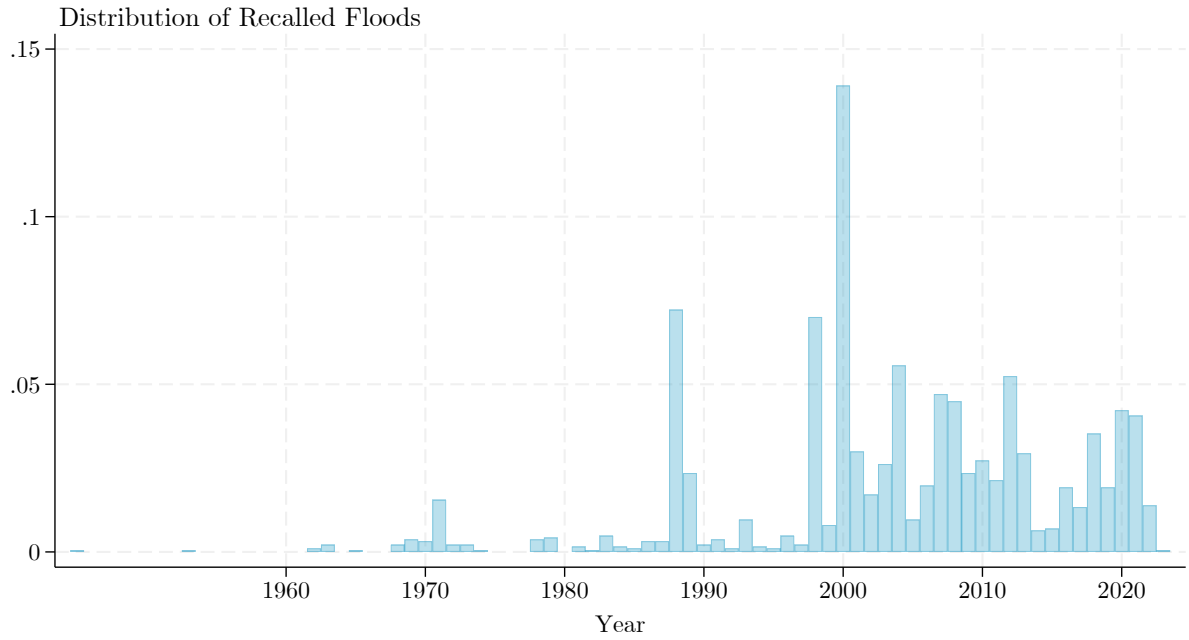


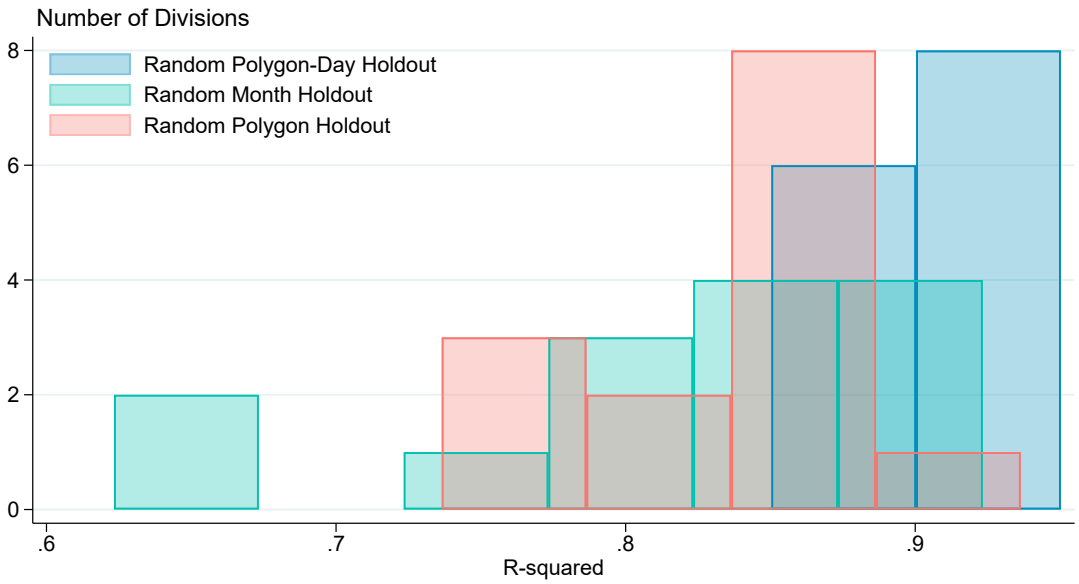
Figure A.12: Distribution of Recalled Floods by Year



Note: Figure A.11 plots the water level for a single river station in Bangladesh over the course of 2022 based on data obtained from the Bangladesh Water Development Board. Figure A.12 plots the distribution of recalled floods across all surveyed farmers by year from the endline elicitation.

Figure A.13: Prediction Algorithm Performance

(a) Surface Water (Continuous)



(b) Flooding (Binary)

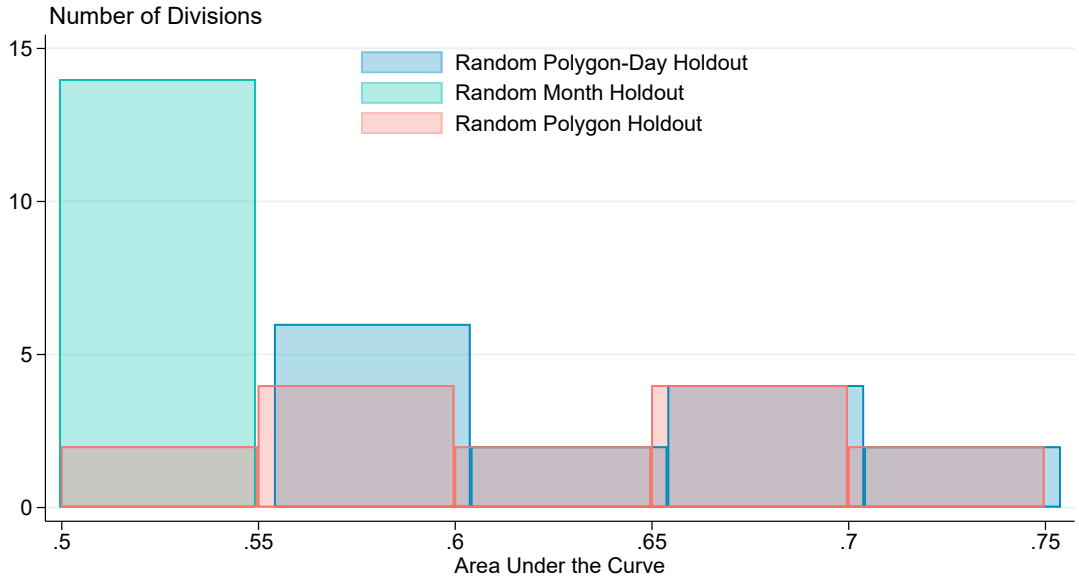


Figure A.14: Figure A.13 plots the distribution of prediction algorithm performances at the division-by-orbit-direction level for Bangladesh. Figure A.13a shows the R^2 of the histogram gradient boosted decision tree using the share of surface water in each union as the predicted outcome; Figure A.13b does the same for the area-under-the-curve in the binary case.

Table A.1: Prediction Algorithm Performance—Continuous Prediction

Division	Hold-Out Sample	Ascending Orbits OLS R^2	Ascending Orbits OLS MSE	Ascending Orbits HGBDT R^2	Ascending Orbits HGBDT MSE	Descending Orbits OLS R^2	Descending Orbits OLS MSE	Descending Orbits HGBDT R^2	Descending Orbits HGBDT MSE
Barisal	Polygon-Days	0.471	0.005	0.906	0.001	0.517	0.005	0.911	0.001
Chittagong	Polygon-Days	0.379	0.007	0.853	0.002	0.360	0.007	0.915	0.001
Dhaka	Polygon-Days	0.539	0.009	0.899	0.002	0.569	0.008	0.923	0.001
Khulna	Polygon-Days	0.600	0.012	0.898	0.003	0.638	0.009	0.928	0.002
Rajshahi	Polygon-Days	0.477	0.012	0.850	0.004	0.590	0.010	0.914	0.002
Rangpur	Polygon-Days	0.502	0.010	0.874	0.002	0.547	0.009	0.874	0.002
Sylhet	Polygon-Days	0.585	0.021	0.902	0.005	0.679	0.013	0.933	0.003
Average	Polygon-Days	0.507	0.011	0.883	0.003	0.557	0.009	0.914	0.002
Barisal	Polygons	0.351	0.005	0.782	0.002	0.446	0.004	0.786	0.002
Chittagong	Polygons	0.306	0.005	0.736	0.002	0.344	0.003	0.858	0.001
Dhaka	Polygons	0.526	0.013	0.844	0.004	0.566	0.012	0.880	0.003
Khulna	Polygons	0.491	0.009	0.869	0.002	0.452	0.007	0.852	0.002
Rajshahi	Polygons	0.432	0.012	0.824	0.004	0.520	0.010	0.866	0.003
Rangpur	Polygons	0.530	0.009	0.847	0.003	0.576	0.009	0.819	0.004
Sylhet	Polygons	0.535	0.021	0.876	0.006	0.647	0.012	0.894	0.004
Average	Polygons	0.535	0.021	0.876	0.006	0.647	0.012	0.894	0.004
Barisal	Months	0.427	0.007	0.853	0.002	0.483	0.006	0.886	0.001
Chittagong	Months	0.363	0.010	0.775	0.003	0.373	0.009	0.894	0.002
Dhaka	Months	0.495	0.011	0.829	0.004	0.441	0.011	0.893	0.002
Khulna	Months	0.271	0.016	0.653	0.008	0.458	0.013	0.867	0.003
Rajshahi	Months	0.394	0.016	0.623	0.010	0.544	0.011	0.871	0.003
Rangpur	Months	0.494	0.011	0.760	0.005	0.452	0.011	0.774	0.004
Sylhet	Months	0.497	0.027	0.794	0.011	0.689	0.018	0.907	0.006
Average	Months	0.420	0.014	0.755	0.006	0.491	0.011	0.870	0.003

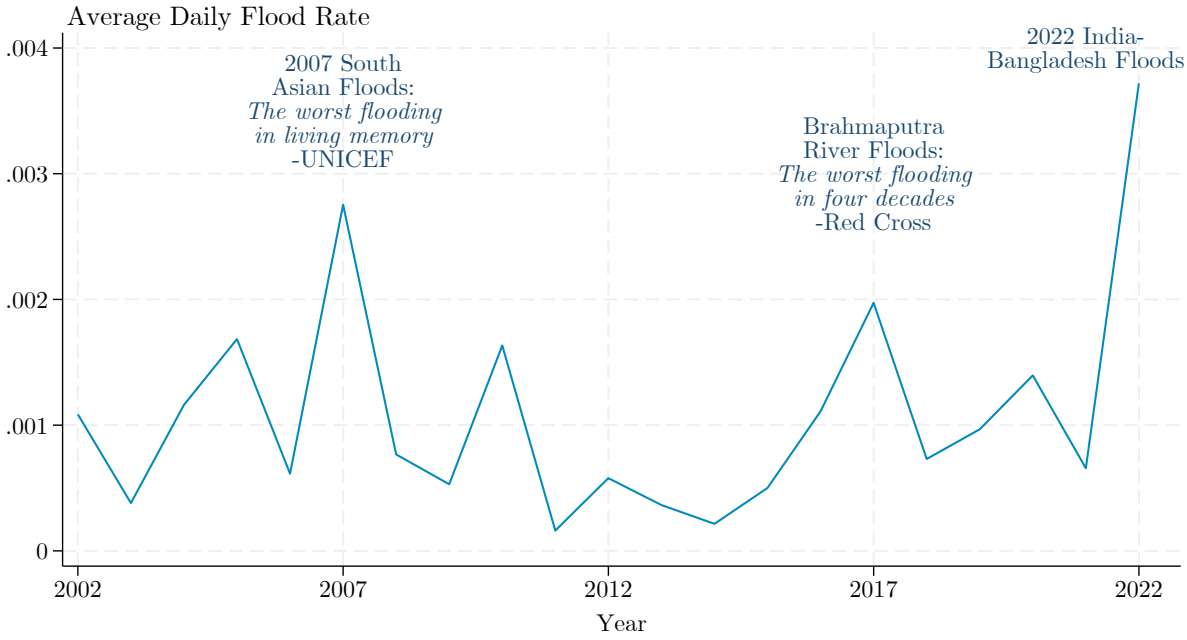
Note: Table A.1 presents the mean squared error and R^2 for each division, orbit direction, test sample, and algorithm (either ordinary least squares or histogram gradient boosted decision tree).

Table A.2: Prediction Algorithm Performance—Binary Prediction

Division	Hold-Out Sample	Ascending Orbits				Descending Orbits			
		Logistic		HGBDT		Logistic		HGBDT	
		Acc.	AUC	Acc.	AUC	Acc.	AUC	Acc.	AUC
Barisal	Polygon-Days	0.997	0.538	0.997	0.557	0.997	0.540	0.998	0.554
Chittagong	Polygon-Days	0.995	0.557	0.995	0.578	0.997	0.526	0.997	0.589
Dhaka	Polygon-Days	0.991	0.623	0.993	0.721	0.992	0.549	0.993	0.650
Khulna	Polygon-Days	0.993	0.509	0.994	0.603	0.997	0.536	0.998	0.582
Rajshahi	Polygon-Days	0.984	0.615	0.988	0.703	0.986	0.582	0.989	0.684
Rangpur	Polygon-Days	0.979	0.640	0.983	0.716	0.986	0.637	0.987	0.698
Sylhet	Polygon-Days	0.975	0.571	0.982	0.649	0.975	0.567	0.981	0.665
Average	Polygon-Days	0.988	0.579	0.990	0.647	0.990	0.563	0.992	0.632
Barisal	Polygons	0.997	0.499	0.998	0.500	0.998	0.633	0.998	0.567
Chittagong	Polygons	0.996	0.548	0.996	0.595	0.997	0.500	0.998	0.560
Dhaka	Polygons	0.991	0.629	0.992	0.682	0.992	0.545	0.992	0.650
Khulna	Polygons	0.995	0.510	0.996	0.588	0.998	0.522	0.998	0.522
Rajshahi	Polygons	0.983	0.598	0.987	0.716	0.985	0.601	0.988	0.696
Rangpur	Polygons	0.986	0.653	0.990	0.738	0.985	0.582	0.988	0.670
Sylhet	Polygons	0.976	0.573	0.983	0.670	0.976	0.581	0.981	0.648
Average	Polygons	0.976	0.573	0.983	0.670	0.976	0.581	0.981	0.648
Barisal	Months	0.992	0.498	0.995	0.499	0.996	0.499	0.999	0.500
Chittagong	Months	0.990	0.518	0.988	0.504	0.993	0.505	0.993	0.500
Dhaka	Months	0.980	0.504	0.986	0.503	0.985	0.510	0.985	0.513
Khulna	Months	0.996	0.500	0.996	0.500	0.996	0.520	0.998	0.500
Rajshahi	Months	0.969	0.505	0.973	0.515	0.974	0.534	0.980	0.531
Rangpur	Months	0.968	0.563	0.975	0.502	0.981	0.516	0.982	0.503
Sylhet	Months	0.942	0.535	0.958	0.524	0.944	0.531	0.947	0.504
Average	Months	0.977	0.517	0.982	0.507	0.981	0.516	0.984	0.507

Note: Table A.2 presents the accuracy rates and area-under-the-curve statistics for each each division, orbit direction, test sample, and algorithm (either logistic regression or histogram gradient boosted decision tree).

Figure A.15: Flood Rate Over Time



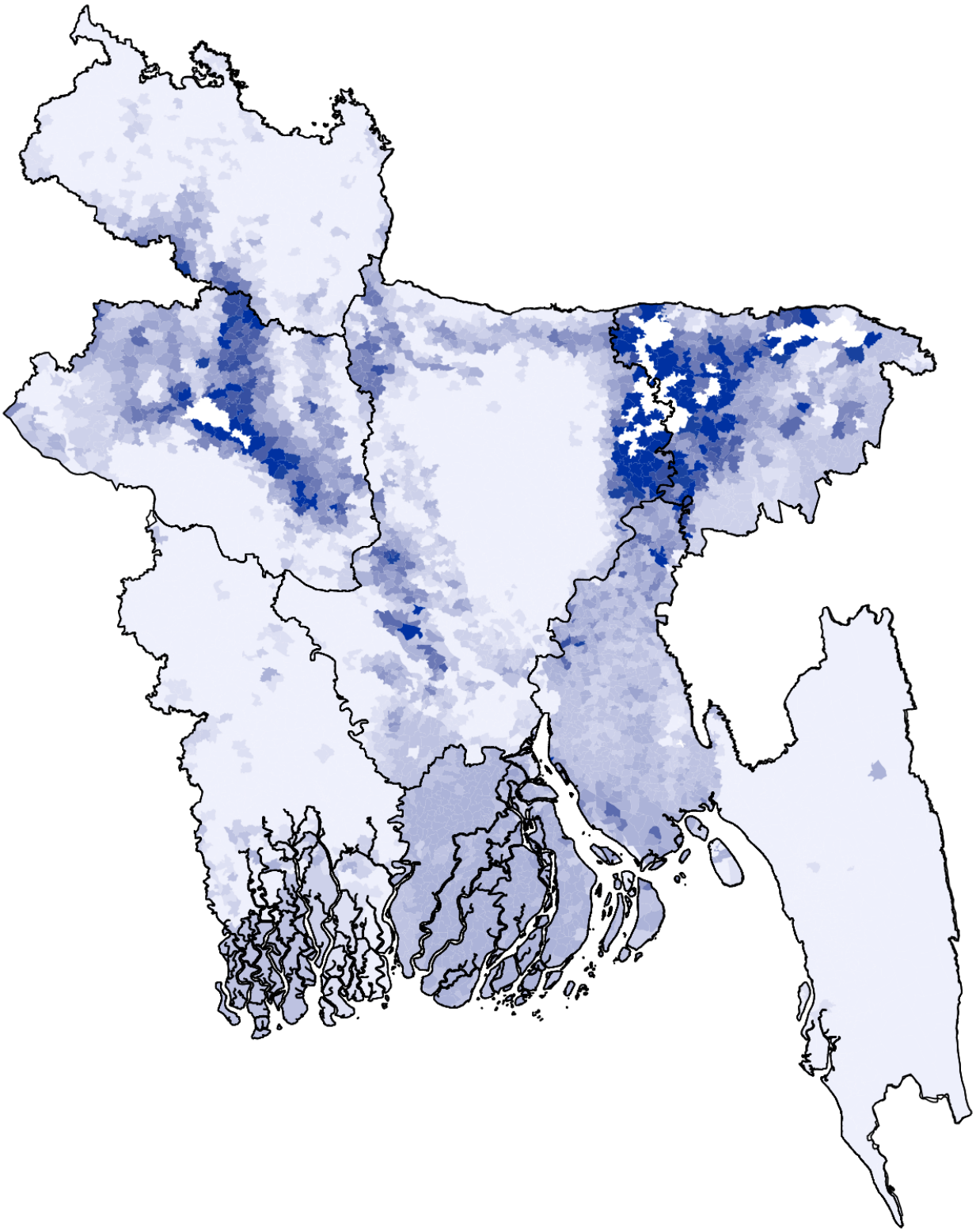
Note: Figure A.15 plots the evolution of average flooding across Bangladeshi unions over time.

Table A.3: Flood Causes — Rural News Article Database

	Share of Floods by Cause
Tidal Surge	0.447
Cyclone	0.076
Heavy Rainfall	0.187
River Erosion/Embankment Failure	0.168
Dam Failure	0.046
Water Logging	0.011
N	262

Note: Table A.3 shows the distribution of causes of each flooding event from the database of articles mentioning floods in rural areas that I compile.

Figure A.16: Flood Incidence Across Bangladesh



Note: Figure A.16 plots the distribution of average flood exposure across unions in Bangladesh, using my method to estimate flooding inundation.

Figure A.17: Flooding by Calendar Day

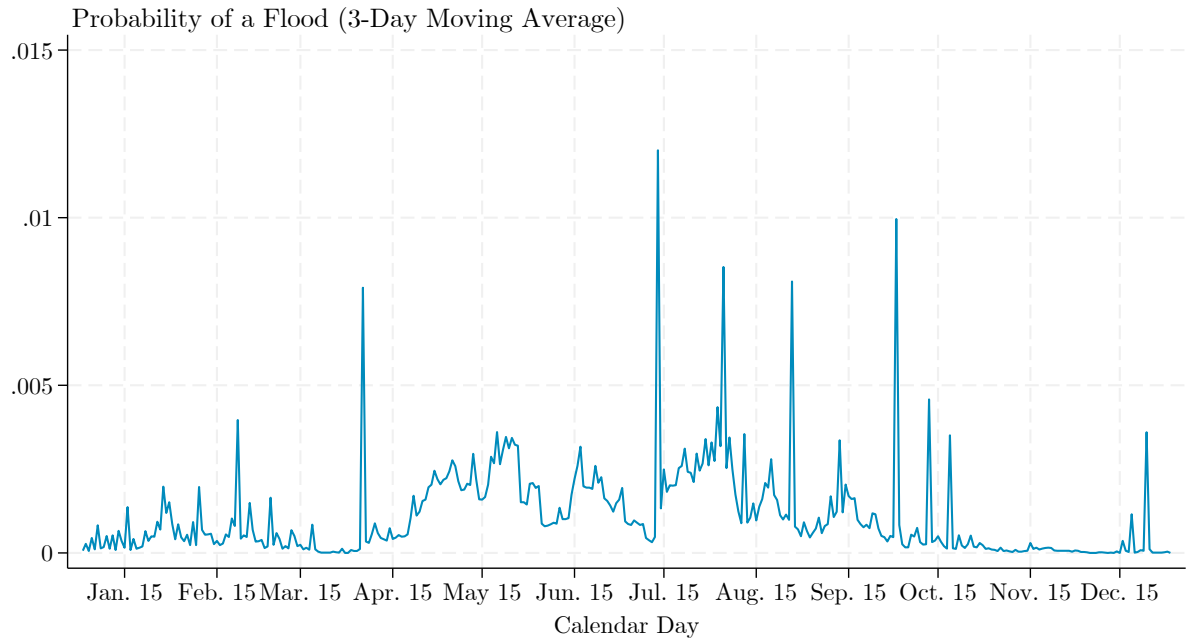
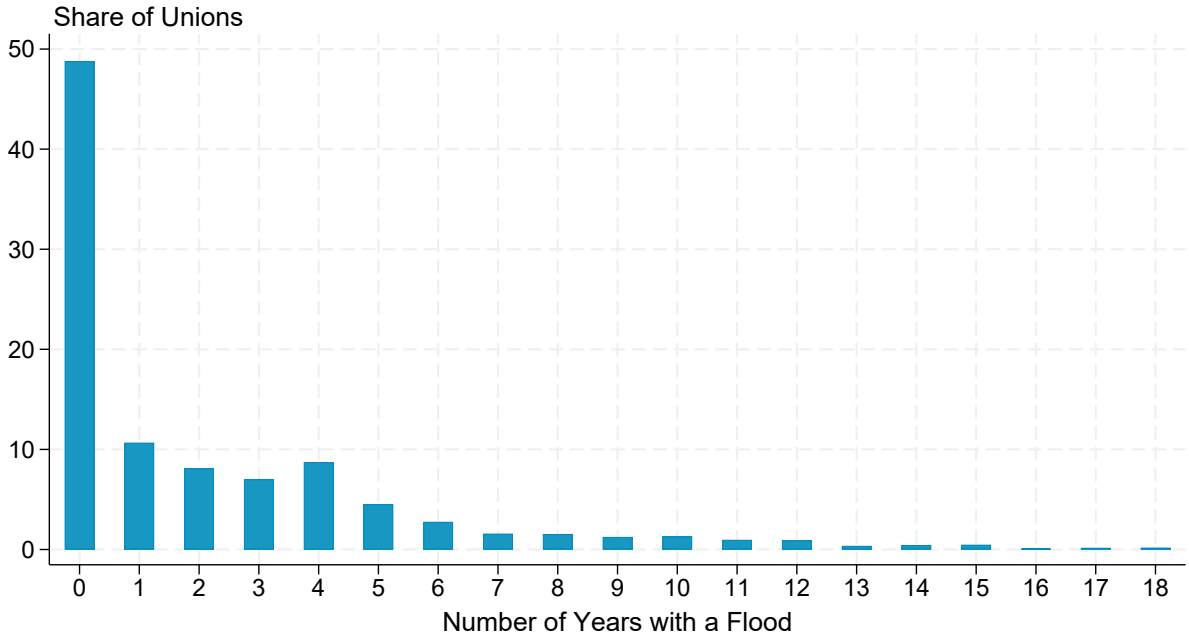
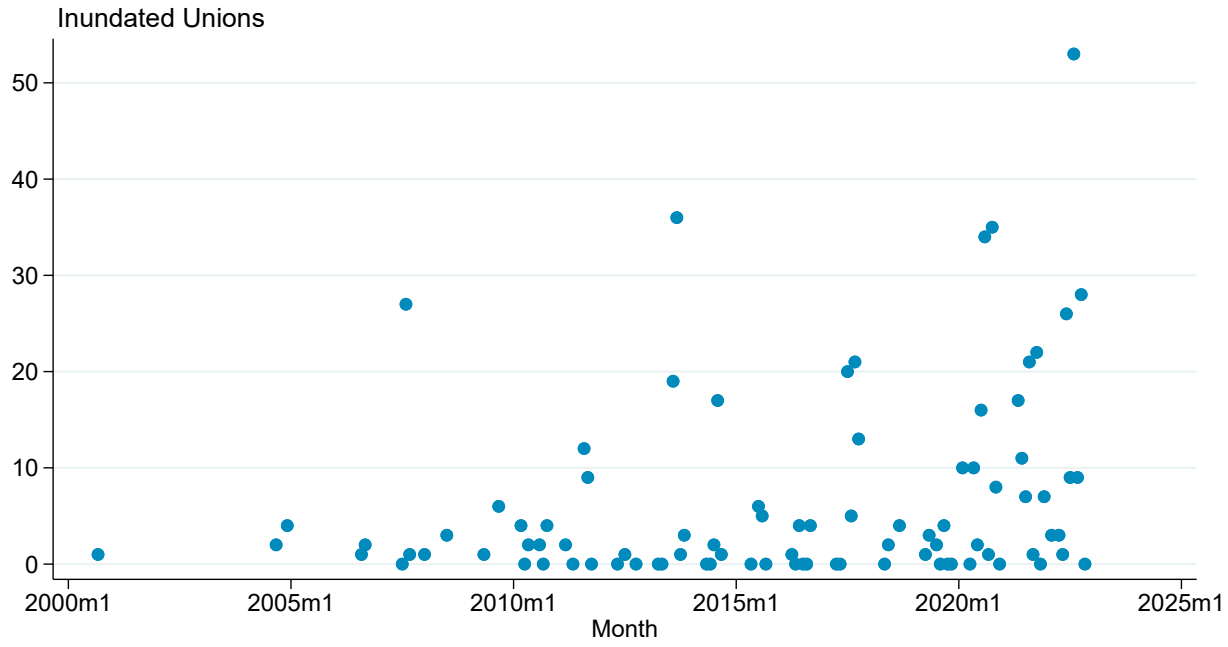


Figure A.18: Distribution of Flood Years by Union



Note: Figure A.17 plots a three-day moving average of the probability of a flood across all unions at the calendar-day level. Figure A.18 plots the distribution of total years experiencing a flood between mid-2002 and the end of 2022 at the union level.

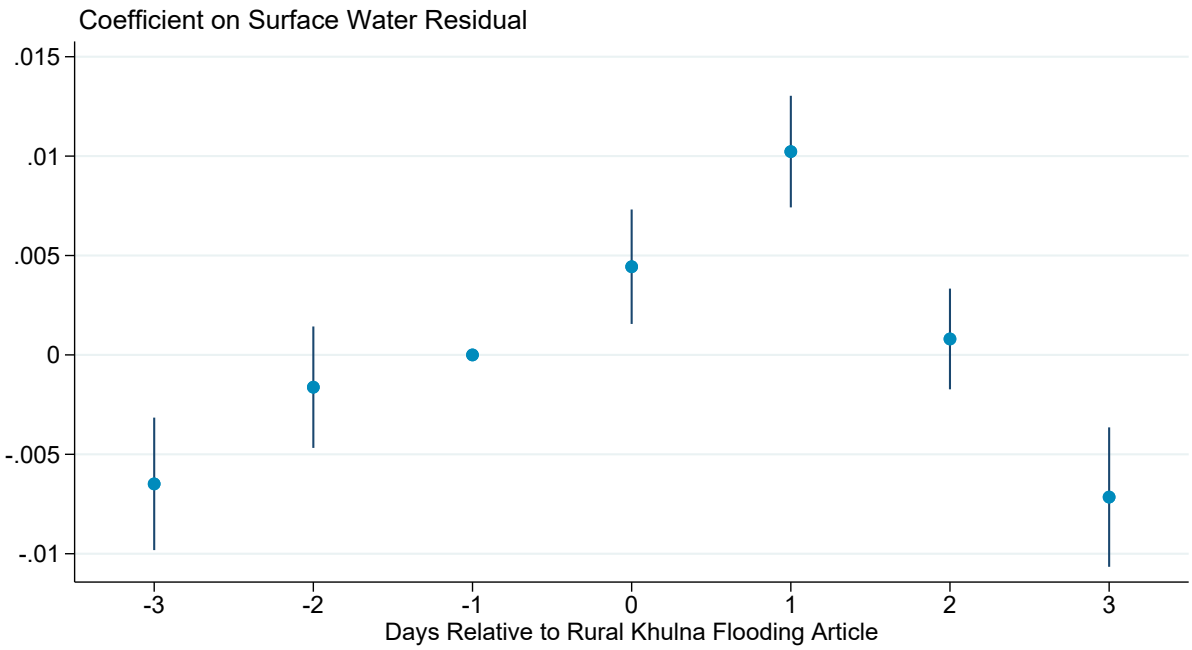
Figure A.19: Union-Level Flood Inundation Frequency — Rural News Article Database



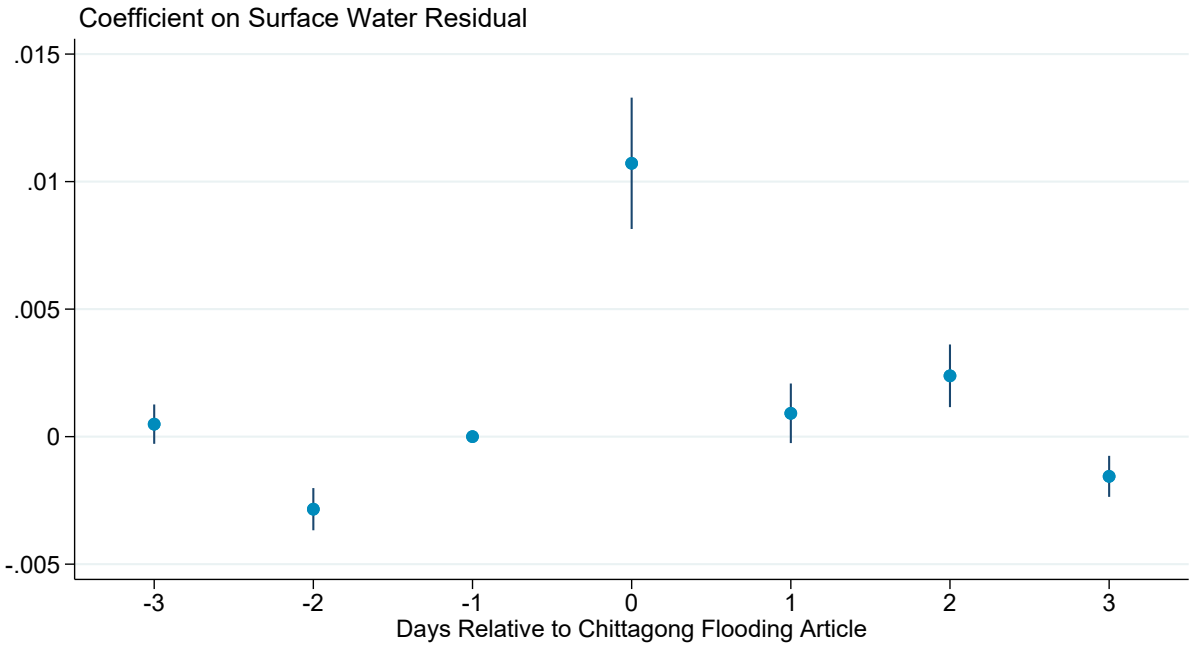
Note: Figure A.19 plots the number of unions inundated by floods per month based on the database constructed from news articles.

Figure A.20: Surface Water and News Reports of Flooding

(a) Rural Sample



(b) Urban Sample



Note: Figure A.20 presents event study plots of how my estimated measure of surface water varies at the union-level relative to the date of the publication of a news article on flooding. In the urban sample, the pool of articles spans 2021 and 2022. The control group is defined at the calendar-day level in the other year. In the rural sample, I include all years since 2010. All regressions include union and calendar day fixed effects, and standard errors are clustered at the union level.

Figure A.21: Government Flood Report Test

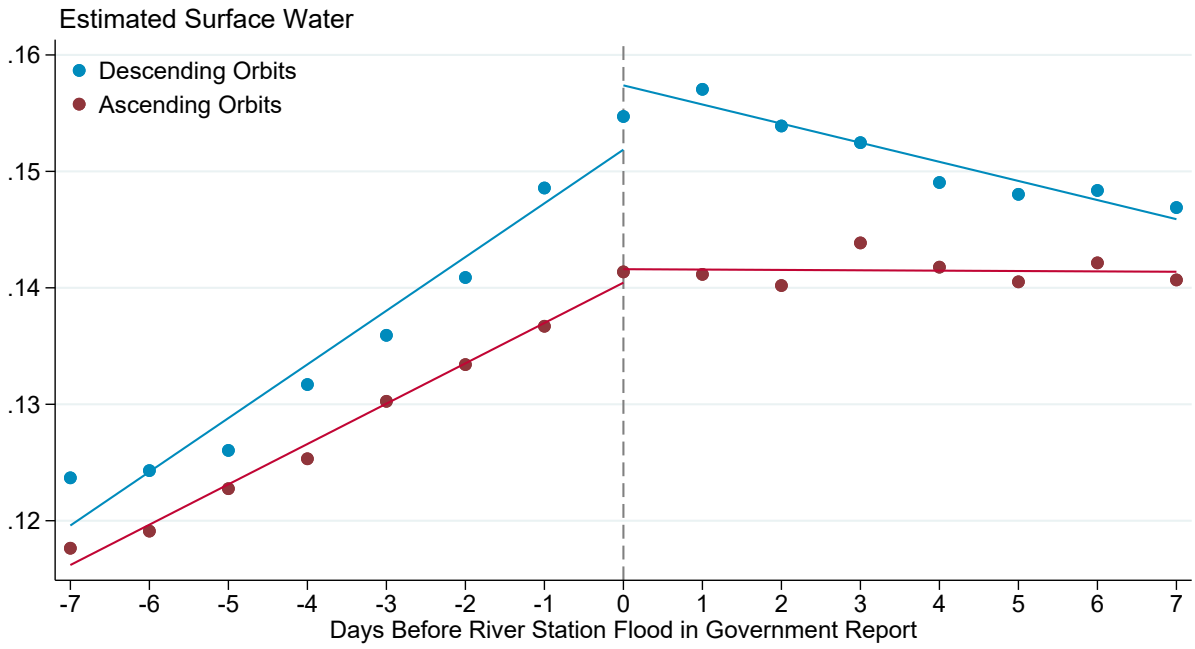
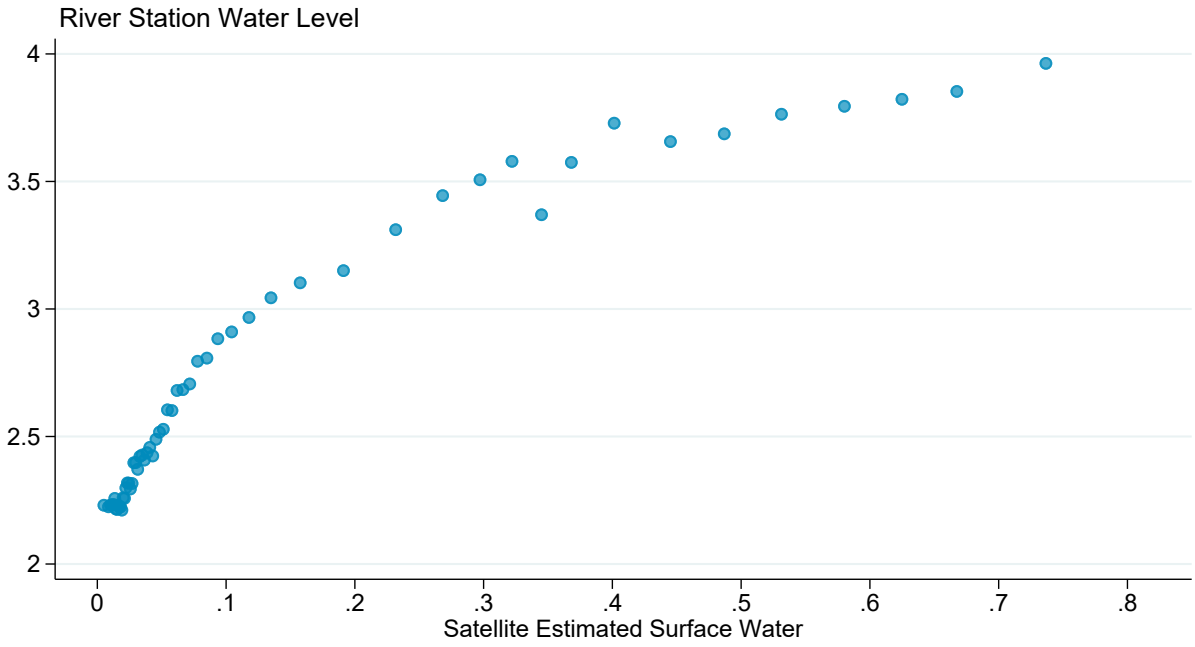


Figure A.22: *in situ* Water Level Measures vs. Estimated Surface Water



Note: Figure A.21 plots a binned scatter plot of how my estimate of surface water in each union varies in the days before and after the government reports the water level from a river in that union exceeded the danger level. Figure A.22 plots a binned scatter plot showing how the water height measured at river stations varies with my estimates of surface water at the union-by-day level, after absorbing union fixed effects.

Figure A.23: Distribution of Maximum Daily Surface Water Residual Each Year by Flood Status

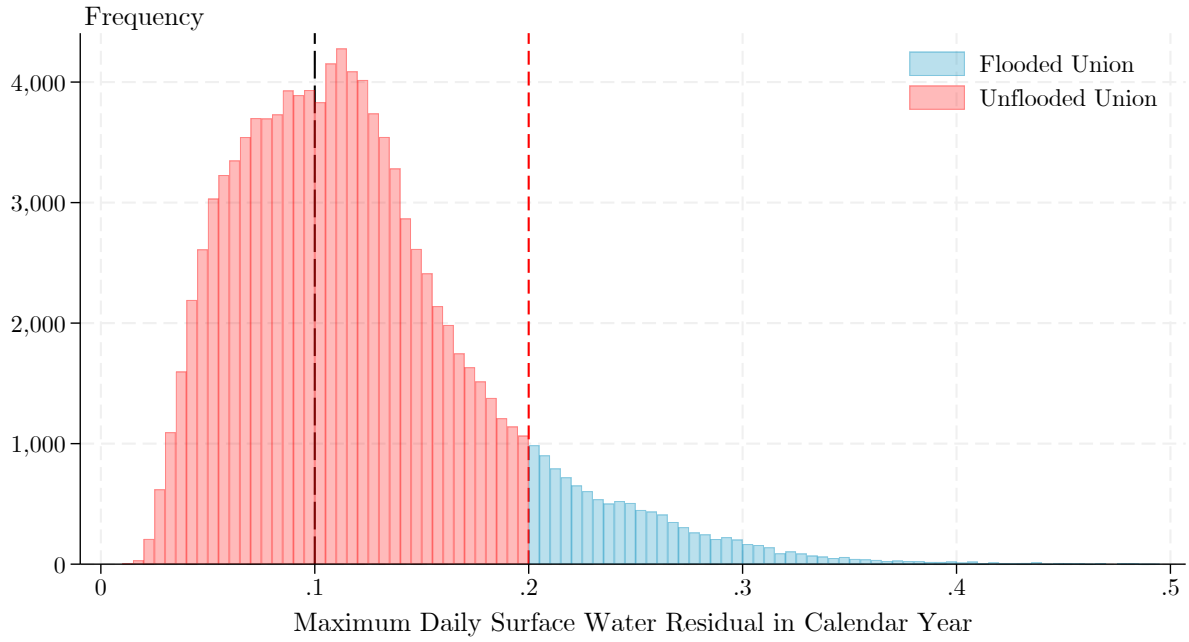
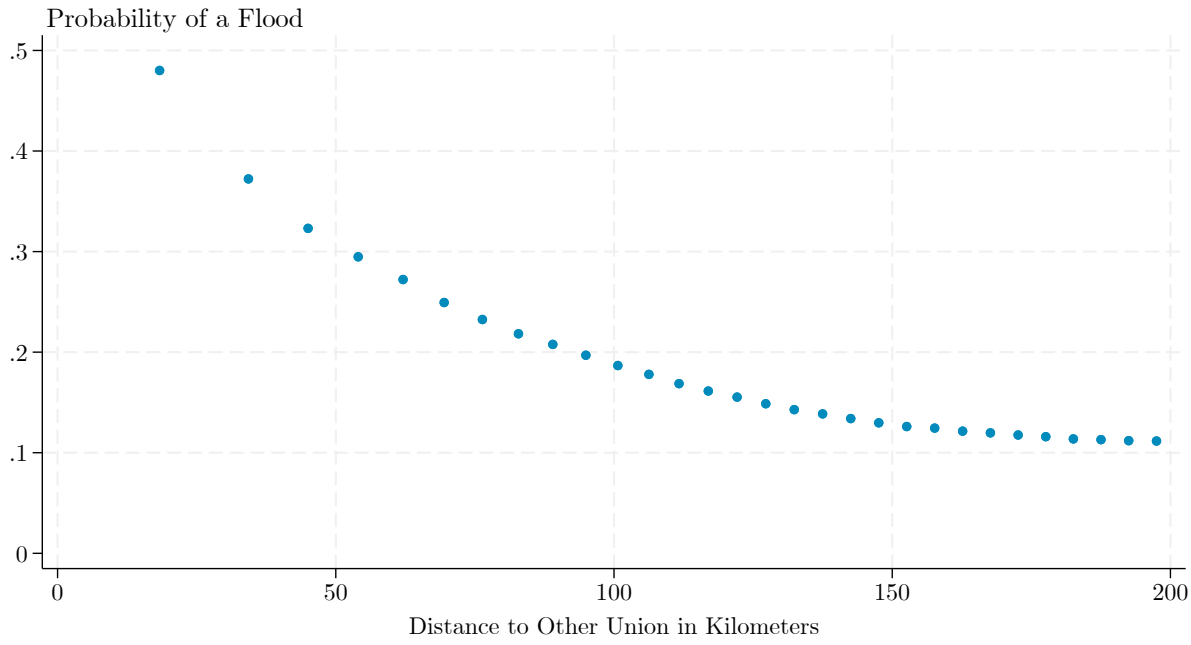


Figure A.24: Spatial Correlation of Flooding



Note: Figure A.23 plots the maximum daily surface water for each union in each year by flood status. Figure A.24 plots a binned scatter plot. The sample begins with all unions that experience a flood in every year from 2002 through 2022. Then, the graph plots for every other union in the country, how the probability of a *different* union within 200 kilometers also experiencing a flood varies with distance.

B Estimating Flood Risk

The primary empirical challenge with estimating flood risk is that idiosyncratic factors may have caused a flood in one area and not the other in recent years, yet both places could still have the same fundamental flood risk. The goal therefore is to use existing data to identify sets of locations where floods *might* happen using an objective scale. I use a data-driven approach to construct this measure using methods from supervised machine learning. This approach has the advantage of not relying on hydrological models, which can often be quite sensitive in their predictions, while also allowing me to let the data tell me which unions are comparable with one another. To flexibly estimate predicted flood, I train an algorithm to predict true flood experience based on geographic characteristics. First, I calculate the daily flood hazard rate based on the full panel of flooding experience. To account for the long tail of the distribution, I assign each union its percentile rank in this distribution of hazard rates, though results are broadly similar using the raw value. This variable constitutes the main flood risk outcome that I predict in the algorithm. As inputs, I calculate the mean and standard deviation of elevation in each union, binned latitude and longitude at the tenth of a degree level, the length of major rivers through that union, average drainage characteristics and flooding depth from the government's hydrological models produced by the Bangladesh Agricultural Research Council (BARC). I then randomly split the sample of 5,158 unions into a training dataset and a testing dataset, with 20 percent reserved for the latter. Finally, I train a random forest algorithm to predict flood rank using the full set of inputs. Applying this model out of sample to the hold-out unions, I can explain 0.71 percent of the variation in true historical incidence ranking using my predicted measure. The fact that this R^2 is less than one can be viewed as an advantage in this case because I can always control for the true past experience. This measure, by contrast, captures a data-driven notion of similarity.

C Data Descriptions and Variable Construction Details

C.1 Nighttime Lights

To measure nighttime light intensity, I use remote sensing data collected by the Visible and Infrared Imaging Suite (VIIRS) Day Night Band on board the JPSS satellites. This instrument provides a significant improvement over the DMSP-OLS series used previously. I use data adjusted following the methods of Elvidge et al. (2021), which removes sunlit, moonlit, and cloudy pixels as well as radiance outliers and fires.[57] Annual composite images are constructed by combining monthly composites with weights proportional to the number of observations that constitute each month's data. I calculate five variables within each union from the annual composite images.[58] My preferred measure is the average masked median luminosity, where masked indicates that any potential issues with contaminated

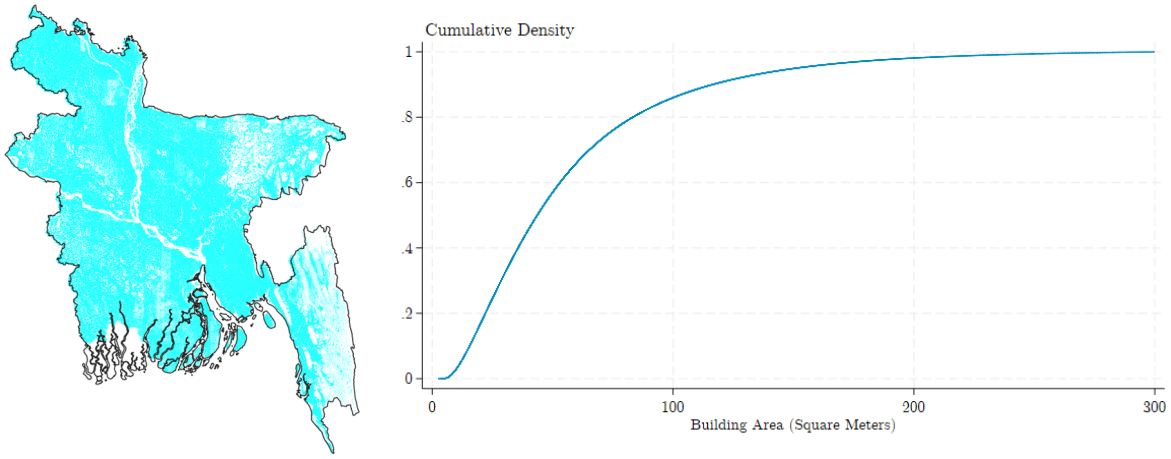
[57] On average for each union in my sample, there are 206.52 VIIRS observations per year free of sunlight and moonlight.

[58] Although monthly composite images are available and have been used in other settings, cloud cover severely reduces the quality of these higher-frequency images for Bangladesh, which is why I rely on the yearly images. At the annual level, the composite images are constructed off of an average of 106.05 cloud-free satellite observations per union per year in my sample.

Figure C.1: Google Open Buildings Data

(a) Spatial Distribution of Buildings

(b) CDF of Building Size



Note: Figure C.1a maps each observation from the Google open building dataset, and Figure C.1b plots the cumulative distribution function of the building footprint area in square meters, top-coded at 300 square meters (above the 99th percentile).

pixels are removed (see Elvidge et al. (2021)). I also report average masked mean luminosity, average mean luminosity, average median luminosity, and average maximum luminosity. I standardize each of these outcomes to the mean of the neverflooded units separately by urban and rural status, and winsorize at the 99^{th} percentile. I also calculate annual growth rates for these same variables, winsorized at the 5^{th} and 95^{th} percentiles.

C.2 Building Footprints

I use the Google Open Buildings dataset of building footprints (Sirko et al., 2021). These data come from a deep learning model trained on high resolution satellite imagery taken at the end of 2020 in the case of Bangladesh. I use the latitude, longitude, and area in square meters for all buildings in Bangladesh. The final sample size includes 3.75e+07 distinct buildings. Figure C.1b plots the cumulative distribution function of their size.

C.3 Bangladesh Sample Vital Registration System Data

Sample The Sample Vital Registration System (SVRS) constitutes a continuous data collection conducted by the Bangladesh Bureau of Statistics to gather demographic information for the country.[59] Notably, the data include information on births, deaths, marriages, and migration, among other key indicators. This is the primary data source upon which the government relies to estimate population shifts in between census rounds. Since 1995, the survey collects data from enough primary sampling units to reliably provide district-level

[59]I am grateful to Md. Karamat Ali of the Bangladesh Bureau of Statistics for his assistance with obtaining and understanding these data.

estimates.

I obtain the SVRS microdata from the Bangladesh Bureau of Statistics for all years between 2003 and 2020. Because the geographic identifiers are not consistent across years or between these data and the unions used by Global Administrative Areas (2018) upon which I base my flooding estimates, I have to construct a crosswalk between the two, much of which had to be done by hand given extensive coding errors in the SVRS data. Nevertheless, I am able to successfully match 99.01 percent of the 4,207,790 households to a corresponding union identifier used in the rest of my analysis. Of the 18,987,203 individual records, I succeed in linking 98.98 percent.

Variable construction The following details describe how I construct the key variables from the SVRS data.

Insolvent: For all years prior to 2016, I define the household-level measure of insolvency as respondents answering "temporary insolvent" in response to the question about their economic condition over the past year, as opposed to answering "permanently insolvent", "balanced income expenditure", "solvent", or "rich with savings". In years after 2016, this fifth option is not offered, but I keep the coding the same. I choose temporary insolvent as opposed to permanently insolvent because that is the most marginal group, and likely households would not even consider themselves to be permanently insolvent in the immediate aftermath of a negative shock like flooding, even if their steady state condition has changed. The pattern of results nevertheless is similar when using this pooled temporary and permanent insolvency condition.

Highest education: At the individual level, I define four categories for the highest educational level achieved with a focus on the ages for which marginal schooling decisions may impact each outcome: no schooling (ages four through eight), fifth grade or more (ages nine through 15), eighth grade or more (ages 12 through 17), and secondary or more (ages 14 through 19). From the individual primary activity module, I also construct an indicator variable equal to 1 if the child age 5 to 18 is reported to be a student.

Literacy: This outcome captures whether the individual can write, which I include for those with ages between 5 and 13.

Marriage: I define an individual as ever-married if they report being married, separated, divorced, widowed, or living separately with their spouse, focusing on those aged between 17 and 30. Separately, I calculate the average age separately by gender among those respondents married in a given year and union.

Births: I collect information on all births in the household—the date of birth, sex of the child, and whether the child is still alive.

Deaths: I collect information on deaths in the household, including sex, age, date of death, and cause. **Marriage:** I define an individual as ever-married if they report being married, separated, divorced, widowed, or living separately with their spouse, focusing on those aged between 17 and 30. Separately, I calculate the average age separately by gender among those respondents married in a given year and union.

In-/Out-Migration: For each household, I use the month and year of out-migration and in-migration for all who left the sample area for at least six months in the past year. Similarly, in-migration is defined in the reverse. For more than 99 percent of respondents,

Table C.1: Crop Calendar for Major Rice Seasons in Bangladesh

Season	Sowing	Harvest	Vegetation Period
Aus	Middle of March	End of August	June 2^{nd} — Aug. 15^{th}
Aman	End of June	Beginning of December	Aug. 18^{th} — Nov. 1^{st}
Boro	Beginning of December	Middle of May	Feb. 15^{th} — May 1^{st}

Note: Table C.1 shows the dates I use to link the major agricultural seasons to the observations from the remote sensing data based off of the Yearbooks of Agricultural Statistics.

the year of migration and the year of survey are the same, so in the count specification, I scale the count of migrants by the number of households interviewed in each union in that survey year. I also calculate the average age, share female, and share of migrants by destination type (rural, urban, abroad) at the union by year level. Among the out-migrants, I have the district of out-migration for those who left to a different district but remained in Bangladesh—approximately half of the sample. Similarly, in-migration is defined in the reverse.

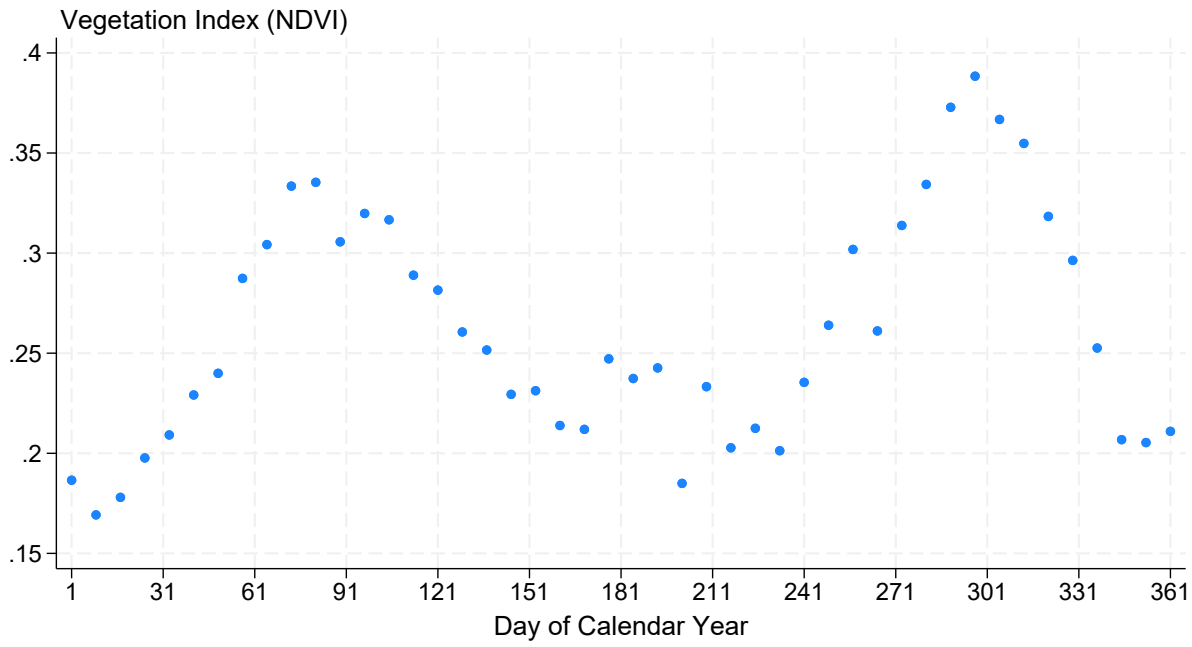
Primary economic activity: I construct six codes to classify the primary economic activity of members of the household. I restrict the sample to prime age workers between 24 and 54 years old. The precise codes vary from year to year. I include anyone from the following sub-categories under each umbrella:

- *Agriculture:* Land owner, Farmer and owner, Family farmer, Contractual agricultural workers, Sharecropper who owns land, Agricultural worker who does not own any land, Other non-agricultural worker
- *Student:* Student
- *Household management/helper:* Household management, Household helper
- *Businessperson:* Businessperson
- *Non-Agricultural Worker:* Other non-agricultural worker, Factor worker/Manufacturer, Transport/commute worker, Domestic worker
- *Office Worker:* Professional officer, Executive officer, Occupational officer, Other office worker, Teacher

C.4 Agricultural Statistics

I digitize records from the Bangladeshi government's Yearbooks of Agricultural Statistics, produced by the Bangladesh Bureau of Statistics. These publications include estimates of crop production at the district level, and I digitize the data for the three major rice seasons: Aus, Aman, and Boro for all available years. I construct crop calendars based off these same publications as shown in Table C.1, where I attempt to find a common period across different kinds of seeds and methods.

Figure C.2: Seasonal Variation in the Vegetation Index



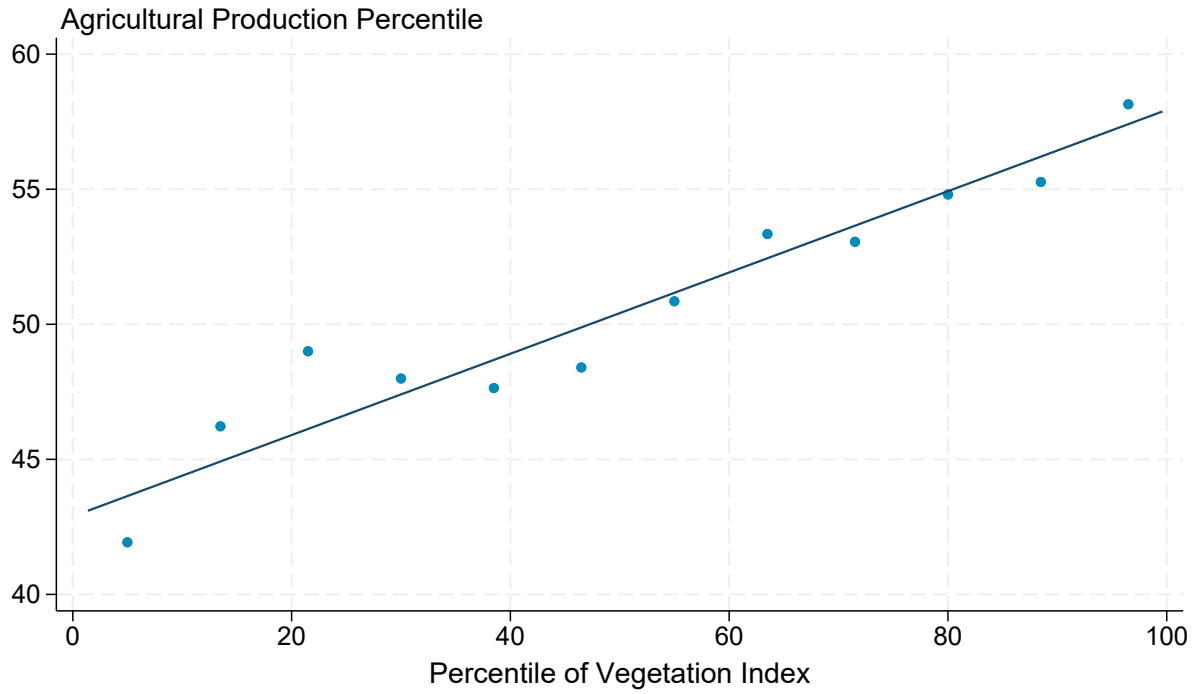
Note: Figure C.2 plots the average value of the NDVI vegetation index across all unions by calendar date of the first observation of the 8-day composite image.

C.5 Vegetation Greenness

I base my measure of greenness on the normalized difference vegetation index (NDVI) calculated from images collected by the Landsat 7 satellite. Note that I do not use information from this satellite when estimating flooding inundation. I calculate the 8-day average within each polygon using composite images made available through Google Earth Engine that have been constructed from Tier 1 orthorectified scenes using the computed top-of-atmosphere reflectance. The index uses the near infra-red (NIR) and red bands of each image, calculated according to $\frac{NIR-Red}{NIR+Red}$. I estimate the average vegetation within polygon and rice growing season by calculating the mean value for all 8-day composites that fall within the last 75 days of the growing season, the dates of which I denote in Table C.1. Figure C.2 plots the average vegetation value across all unions by calendar date of the first observation of the 8-day composite image.

To benchmark this index in meaningful economic terms, I calculate the values at the district level and link them to the statistics I digitize from the Yearbooks of Agricultural Statistics. I focus on the total production and total acres cultivated in each citizen, pooling across different seed varieties.

Figure C.3: Agricultural Production and Vegetation Index Correlation



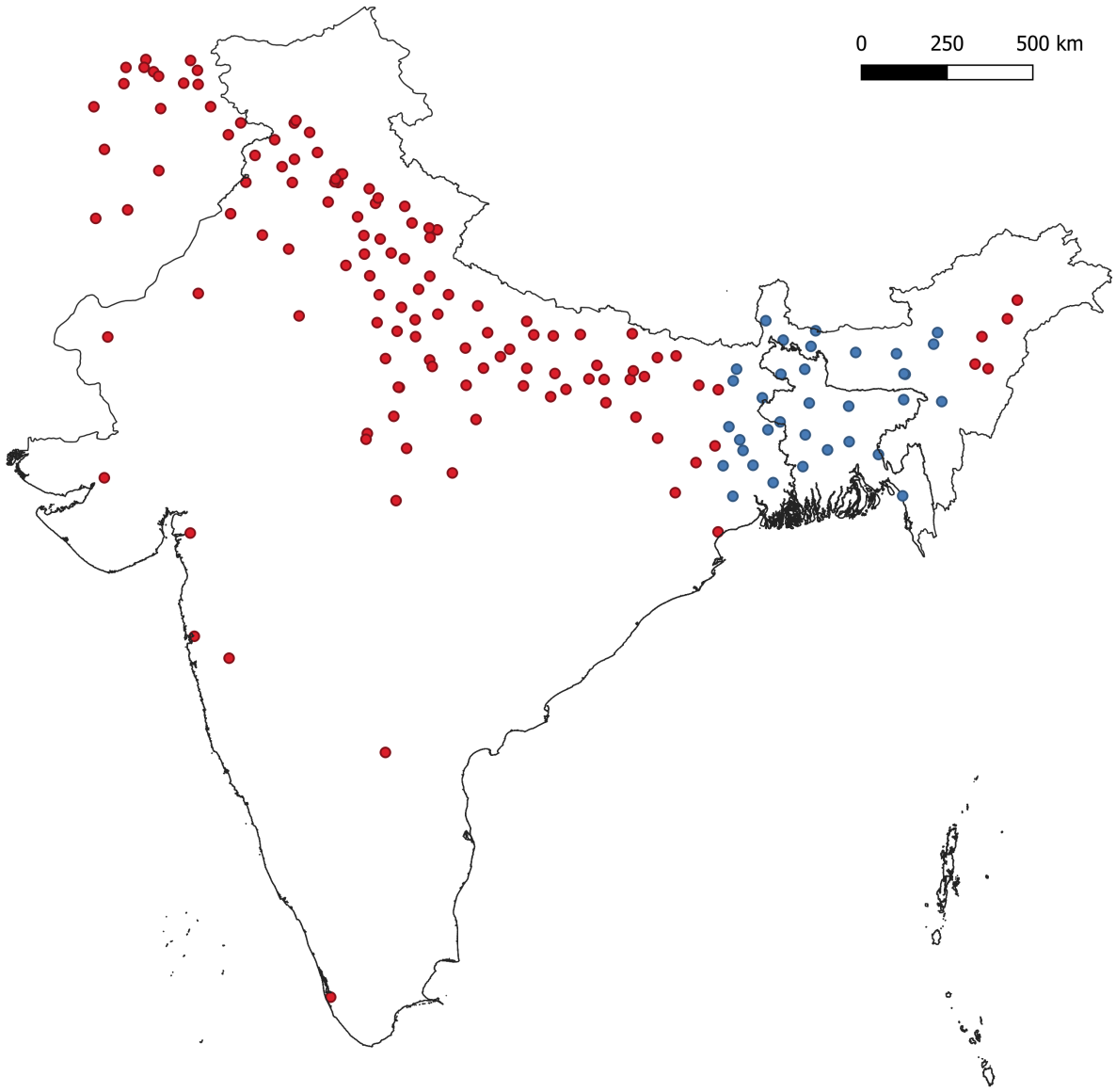
C.6 Colonial Road Networks

This section describes the digitization and construction of the colonial routes dataset from Dutton (1877).

I digitize all cities with longitudes between 87 degrees and 93 degrees.

Figure C.4 maps the full set of nodes superimposed on the modern-day boundaries of India and Bangladesh, with those places included in my network analysis in blue and the others in red.

Figure C.4: Nodes in the Historical Transportation Network



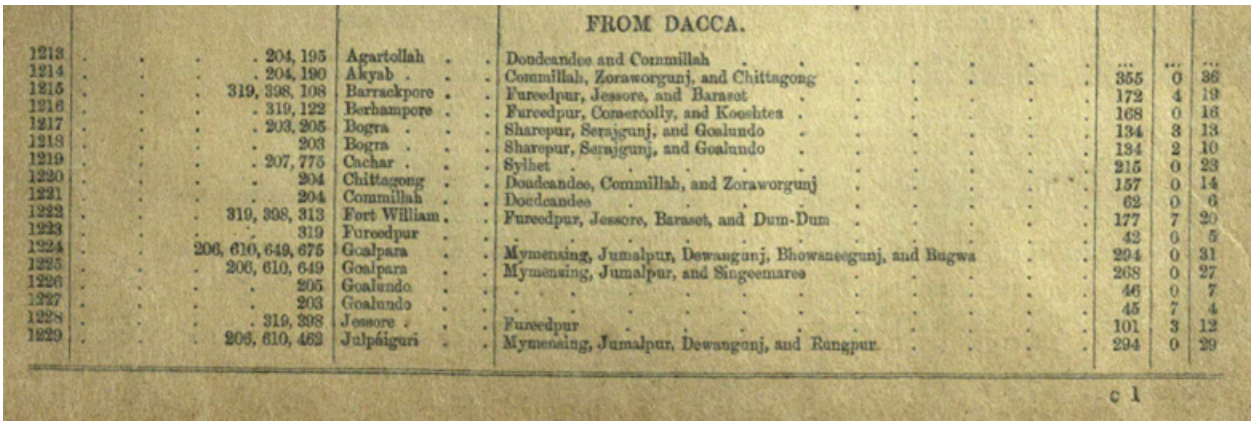
Note: In Figure C.4, I plot each node from Dutton (1877). The blue dots denote the target notes constituting my main sample.

Table C.2: Benchmarking the Vegetation Index using Agricultural Production

	(1) Production Percentile	(2) Total Production	(3) Log Production	(4) Production per Acre Percentile	(5) Production per Acre	(6) Log Production per Acre
NDVI Percentile	0.0172***			0.0393***		
	(0.00427)			(0.0113)		
NDVI		55051.1**	0.134*		0.157**	0.160***
		(26489.7)	(0.0730)		(0.0627)	(0.0563)
Observations	3262	3262	3239	3238	3238	3238

Note: Figure C.3 plots a binned scatter plot at the district-year-season level of production percentile against vegetation index percentile. The production value comes from the Yearbooks of Agricultural Statistics. The vegetation index is the average normalized difference vegetation index from 8-day composite images collected by the Landsat 7 satellite. This graph adjusts for district, year, and season fixed effects. Table C.2 presents ordinary least squares regressions of district-season-level production statistics on the vegetation index (NDVI) with heteroskedasticity-robust standard errors. All specifications include district, year, and season fixed effects, and the production specifications in columns (1) through (3) additionally include fixed effects for the percentile of total acres cultivated.

Figure C.5: Sample



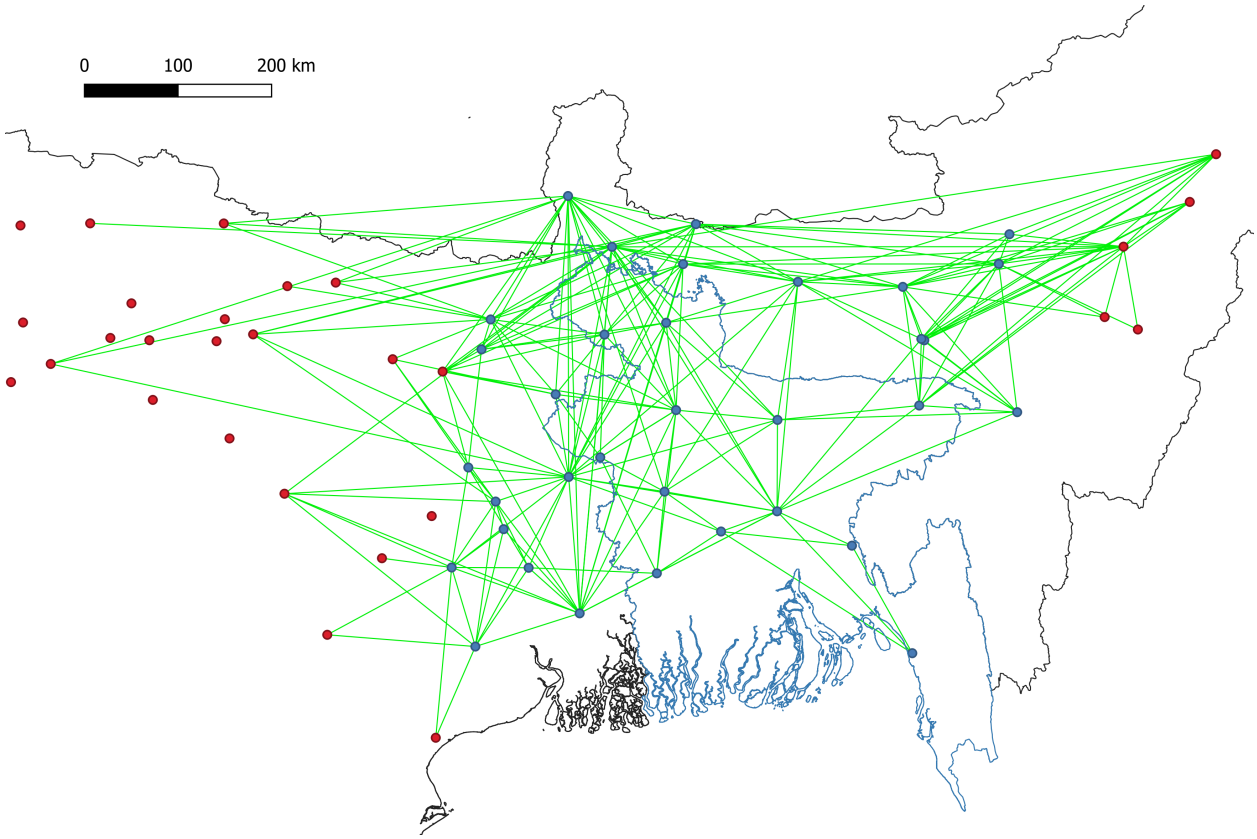
FROM DACCA.

1213	204, 195	Agartollah	Doudcandee and Commillah	...	...	...
1214	204, 190	Akyab	Commillah, Zoraworgunj, and Chittagong	355	0	36
1215	319, 398, 108	Barrackpore	Fureedpur, Jessore, and Baraset	172	4	19
1216	319, 122	Berhampore	Fureedpur, Comercolly, and Kooshtea	168	0	16
1217	203, 205	Bogra	Sharepur, Serajgunj, and Goalundo	134	3	13
1218	203	Bogra	Sharepur, Serajgunj, and Goalundo	134	2	10
1219	207, 775	Cachar	Sylhet	215	0	23
1220	204	Chittagong	Doudcandee, Commillah, and Zoraworgunj	157	0	14
1221	204	Commillah	Doudcandee	62	0	6
1222	319, 398, 313	Fort William	Fureedpur, Jessore, Baraset, and Dum-Dum	177	7	20
1223	319	Fureedpur		42	0	5
1224	206, 610, 649, 675	Goalpara	Mymensing, Jumalpur, Dewangunj, Bhowaneegunj, and Bugwa	294	0	31
1225	206, 610, 649	Goalpara	Mymensing, Jumalpur, and Singeemaree	268	0	27
1226	205	Goalundo		46	0	7
1227	203	Goalundo		45	7	4
1228	319, 398	Jessore	Fureedpur	101	3	12
1229	206, 610, 462	Julpáiguri	Mymensing, Jumalpur, Dewangunj, and Rangpur	294	0	29

c 1

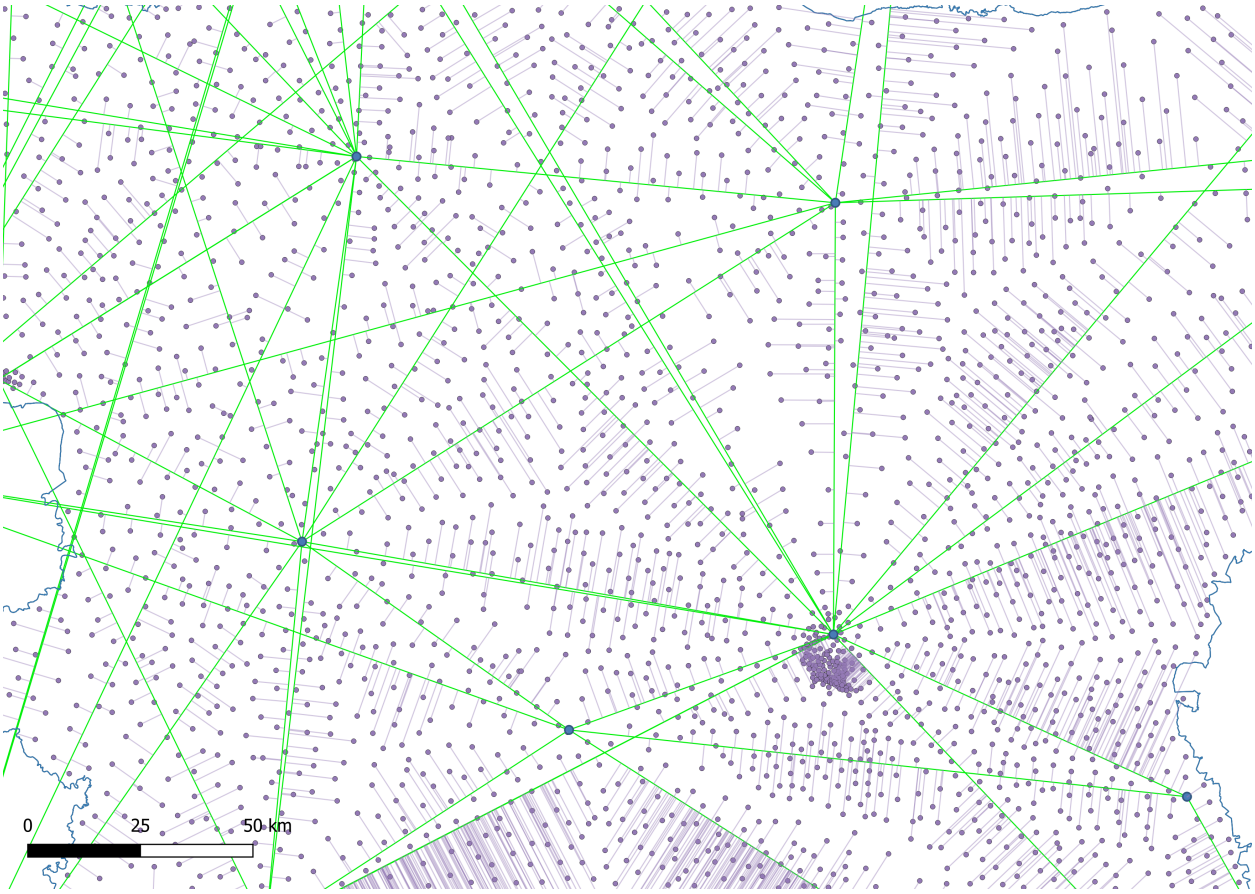
Note: Figure C.5 presents a sample entry from Dutton (1877).

Figure C.6



Note: Figure C.6 shows the straight-line version of the historical transportation network.

Figure C.7: Sample



Note: Figure C.7 presents a sample of the distance between the centroid of each union (shown in purple) to the straight-line version of the historical transportation network (shown in green).

D Adaptation Model Derivation

This section derives the result from the simple climate adaptation model. The steady state optimal investment level is given by equation 17, where δ is the discount factor, PM is the probability of a disaster times the maximum damage from that event, τ is the duration of adaptation investment before it completely depreciates, and $g(\cdot)$ is the adaptation production function.

$$1 = -\delta PM g'(\tau x) \tag{17}$$

Setting $g(\cdot)$ to $1/(1+\tau x)$, I consider the impact of increasing the efficiency of adaptation investment by moving to $g^*(\cdot) = 1/(1+\beta\tau x)$. Let α denote the relative size of equilibrium investment, such that $x^* = \alpha x$. Plugging in the associated derivatives to equation 17 gives:

$$\frac{\tau\delta PM}{(\tau x+1)^2} = \frac{\beta\tau\delta PM}{(\beta\tau\alpha x+1)^2}$$

The discount factor, disaster likelihood, and maximum damage terms drop out, yielding:

$$(\beta\tau\alpha x+1)^2 = \beta(\tau x+1)^2$$

Expanding the left-hand side and rearranging gives:

$$(\beta\tau x)^2\alpha + 2\beta\tau\alpha x + 1 - \beta(\tau x+1)^2 = 0$$

Applying the quadratic formula to solve for α:

$$\frac{-2\beta\tau x \pm \sqrt{4(\beta\tau x)^2 - 4(\beta\tau x)^2 * (1-\beta(\tau x+1)^2)}}{2(\beta\tau x)^2} = \alpha$$

Factoring out $2\beta\tau x$ simplifies this expression to become:

$$\frac{-1 \pm \sqrt{\beta(\tau x+1)^2}}{\beta\tau x} = \alpha$$

Focusing on the positive solution, α equals:

$$\alpha = \frac{\sqrt{\beta}\tau x + \sqrt{\beta} - 1}{\beta\tau x}$$

Plugging this expression into the denominator of the new production function gives:

$$\sqrt{\beta}(\tau x+1)$$

The larger this expression, the greater the climate damage mitigation. The new equilibrium will be higher if the following expression is positive, simply subtracting the denominator

from the previous level of adaptation:

$$\sqrt{\beta}(\tau x + 1) - (\tau x + 1)$$

Since $\beta > 0$ by assumption, this expression is always positive, and therefore equilibrium adaptation increases, even if $\alpha < 1$ $\square$

E Robustness Checks

Figure E.1: The Impact of Floods on Nighttime Luminosity—Alternative Spillovers

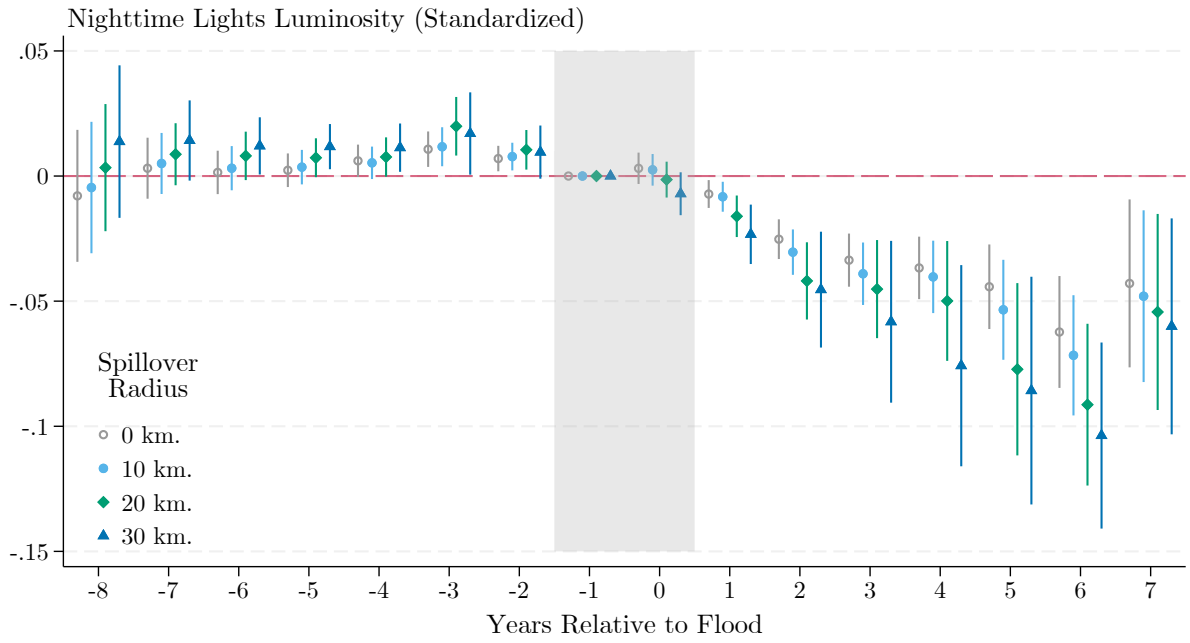
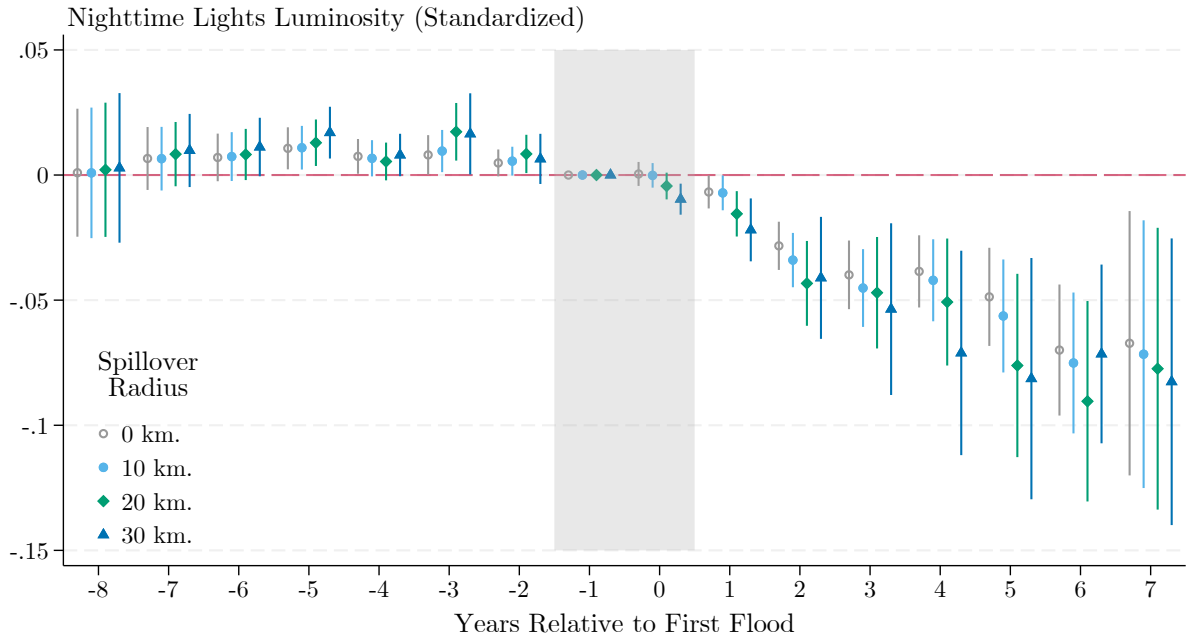


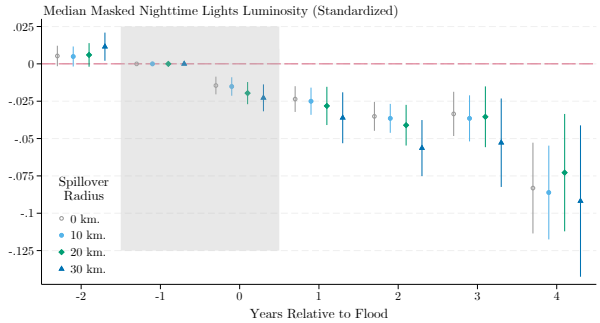
Figure E.2: The Impact of a First Flood on Nighttime Luminosity—Alternative Spillovers



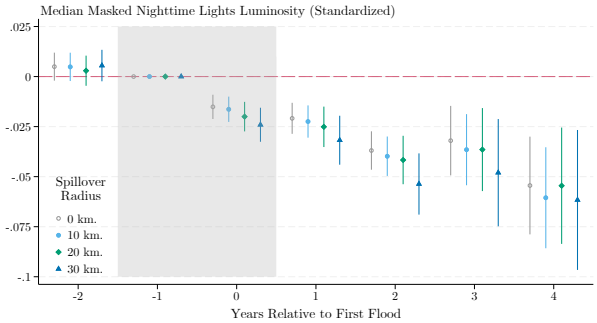
Note: Figure E.1 plots the coefficients from an event study of flooding using the sample to estimate the effect of an additional flood. Figure E.2 plots the coefficients from an event study of flooding using only the first-flood comparisons to isolate the impulse response function. The outcome in both is the average value within each polygon of the median luminosity after masking out problematic observations in the satellite data, standardized to the never-flooded unit distribution. All regressions include union and strata by relative time fixed effects with standard errors clustered at the union level and union-by-cohort level synthetic control weights. The spillover radius denotes the distance at which I exclude control unions whose centroids fall within a given radius of any treated union centroid. See Appendix Section C.1 for details on the data construction.

Figure E.3: The Impact of Floods on Nighttime Luminosity—Fixed Sample Specifications

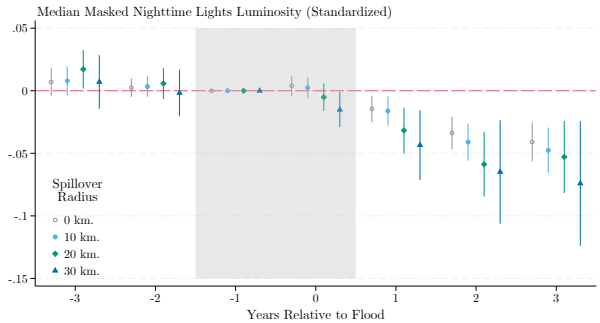
(a) $y = \{2015, 2016, 2017\}$ (All)



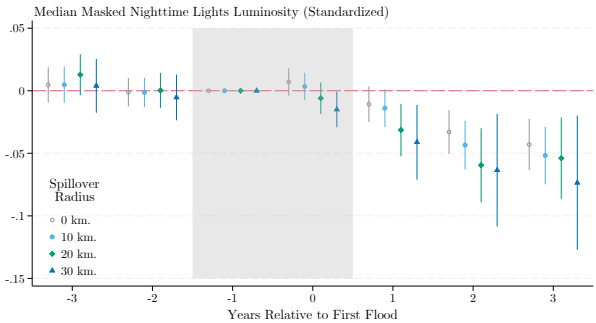
(b) $y = \{2015, 2016, 2017\}$ (First)



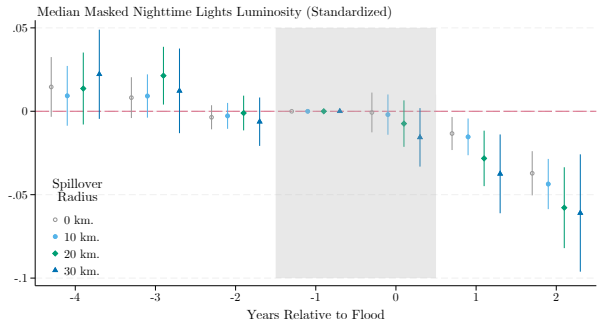
(c) $y = \{2016, 2017, 2018\}$ (All)



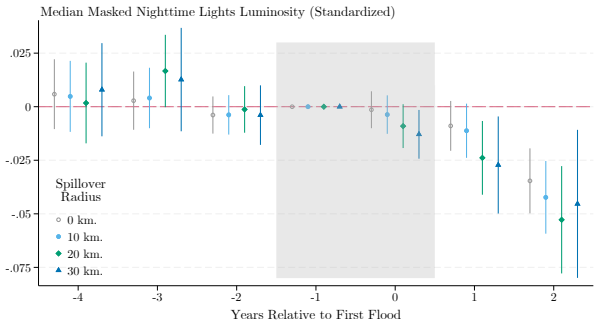
(d) $y = \{2016, 2017, 2018\}$ (First)



(e) $y = \{2017, 2018, 2019\}$ (All)



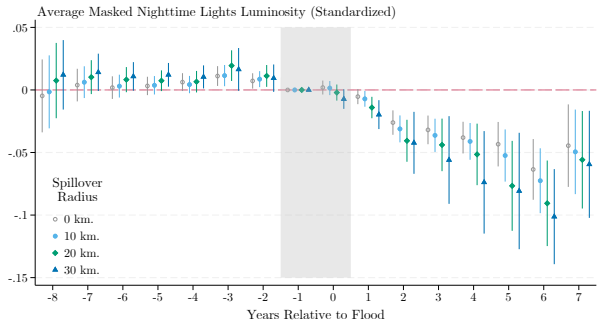
(f) $y = \{2017, 2018, 2019\}$ (First)



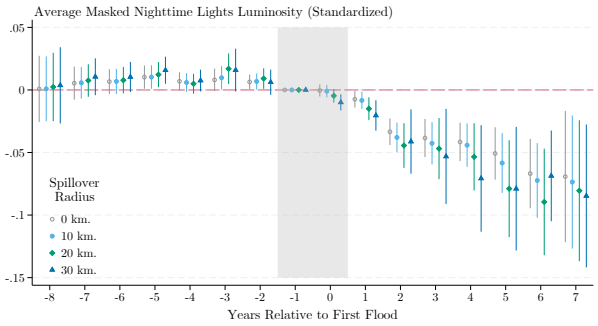
Note: Figure E.3 plots versions of the event studies of the impact of floods on average median masked nighttime luminosity standardized to the never-flooded distribution holding fixed the sample of unions across horizons. The set of calendar years in $y = \{\cdot\}$ denotes the year of flooding events that form the stacked regression, and all vs. first denotes the sample construction. All specifications include union and strata by relative time fixed effects, with standard errors clustered at the union level. See Appendix Section C.1 for details on the data construction.

Figure E.4: The Impact of Floods on Nighttime Luminosity—Alternative Outcomes

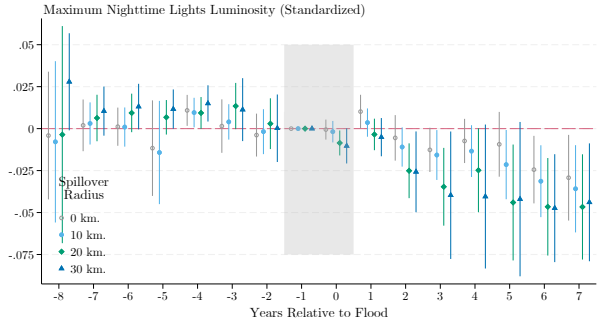
(a) Average Mean Luminosity (All)



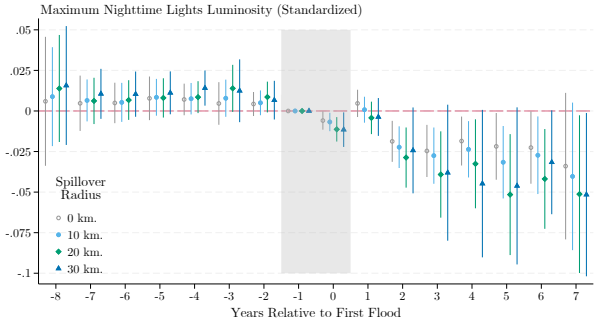
(b) Average Mean Luminosity (First)



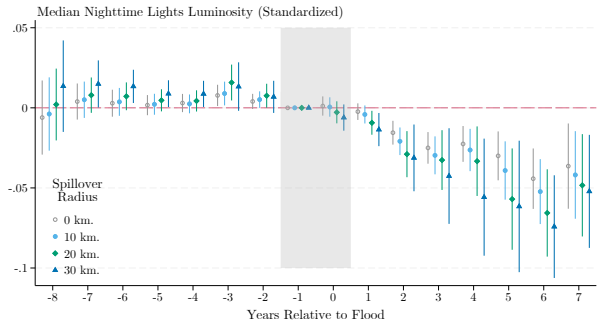
(c) Average Maximum Luminosity (All)



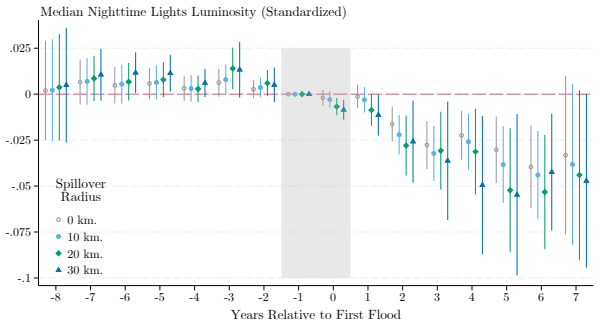
(d) Average Maximum Luminosity (First)



(e) Average Unmasked Median (All)



(f) Average Unmasked Median (First)



Note: Figure E.4 plots event study coefficients with different outcome variables, each of which has been standardized to the never-flooded distribution. All regressions include union and strata by relative time fixed effects, with standard errors clustered at the union level. See Appendix Section C.1 for details on the data construction.

Table E.1: The Impact of Floods on Maximum Nighttime Luminosity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.00786	-0.00725	-0.0184***	-0.0232***	-0.00992	-0.0274***	-0.0267**	-0.00922	-0.0330***
	(0.00586)	(0.00573)	(0.00431)	(0.00851)	(0.00700)	(0.00559)	(0.0128)	(0.00959)	(0.00695)
Sample	All	All	All	All	All	All	All	All	All
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.065	0.093	0.055	0.002	0.117	0.048	0.001	0.311	0.150
Control Mean	-.112	-.112	-.112	-.108	-.108	-.108	-.121	-.121	-.121
Observations	24867	24867	24867	20655	20655	20655	16884	16884	16884
Clusters	2,763	2,763	2,763	2,295	2,295	2,295	1,876	1,876	1,876
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0174***	-0.0179***	-0.0259***	-0.0321***	-0.0238***	-0.0347***	-0.0324**	-0.0272***	-0.0385***
	(0.00635)	(0.00434)	(0.00508)	(0.0100)	(0.00530)	(0.00649)	(0.0144)	(0.00671)	(0.00764)
Sample	First	First	First	First	First	First	First	First	First
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean	-.098	-.098	-.098	-.096	-.096	-.096	-.113	-.113	-.113
Observations	21789	21789	21789	17901	17901	17901	14769	14769	14769
Clusters	2,421	2,421	2,421	1,989	1,989	1,989	1,641	1,641	1,641

Note: Table E.1 presents stacked difference-in-differences estimates of the impact of flooding on average maximum luminosity, standardized to the never-flooded union distribution. All specifications include union fixed effects and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between synthetic control weights from Arkhangelsky et al. (2021), propensity score weights calculated from the strata, or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table E.2: The Impact of Floods on Average Nighttime Luminosity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0256***	-0.0167***	-0.0329***	-0.0372***	-0.0203***	-0.0412***	-0.0448***	-0.0215***	-0.0498***
	(0.00493)	(0.00375)	(0.00420)	(0.00809)	(0.00441)	(0.00562)	(0.0123)	(0.00569)	(0.00673)
Sample	All	All	All	All	All	All	All	All	All
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean	-.132	-.132	-.132	-.129	-.129	-.129	-.14	-.14	-.14
Observations	24543	24543	24543	20457	20457	20457	16749	16749	16749
Clusters	2,727	2,727	2,727	2,273	2,273	2,273	1,861	1,861	1,861
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0313***	-0.0222***	-0.0399***	-0.0429***	-0.0269***	-0.0488***	-0.0472***	-0.0289***	-0.0551***
	(0.00630)	(0.00428)	(0.00529)	(0.00996)	(0.00535)	(0.00678)	(0.0144)	(0.00676)	(0.00761)
Sample	First	First	First	First	First	First	First	First	First
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Control Mean	-.12	-.12	-.12	-.119	-.119	-.119	-.134	-.134	-.134
Observations	21447	21447	21447	17712	17712	17712	14634	14634	14634
Clusters	2,383	2,383	2,383	1,968	1,968	1,968	1,626	1,626	1,626

Note: Table E.2 presents stacked difference-in-differences estimates of the impact of flooding on average masked mean luminosity, standardized to the never-flooded union distribution. All specifications include union fixed effects and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between synthetic control weights from Arkhangelsky et al. (2021), propensity score weights calculated from the strata, or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Figure E.5: Placebo Tests for the Impact of Floods on Nighttime Luminosity: Synthetic Control Weights

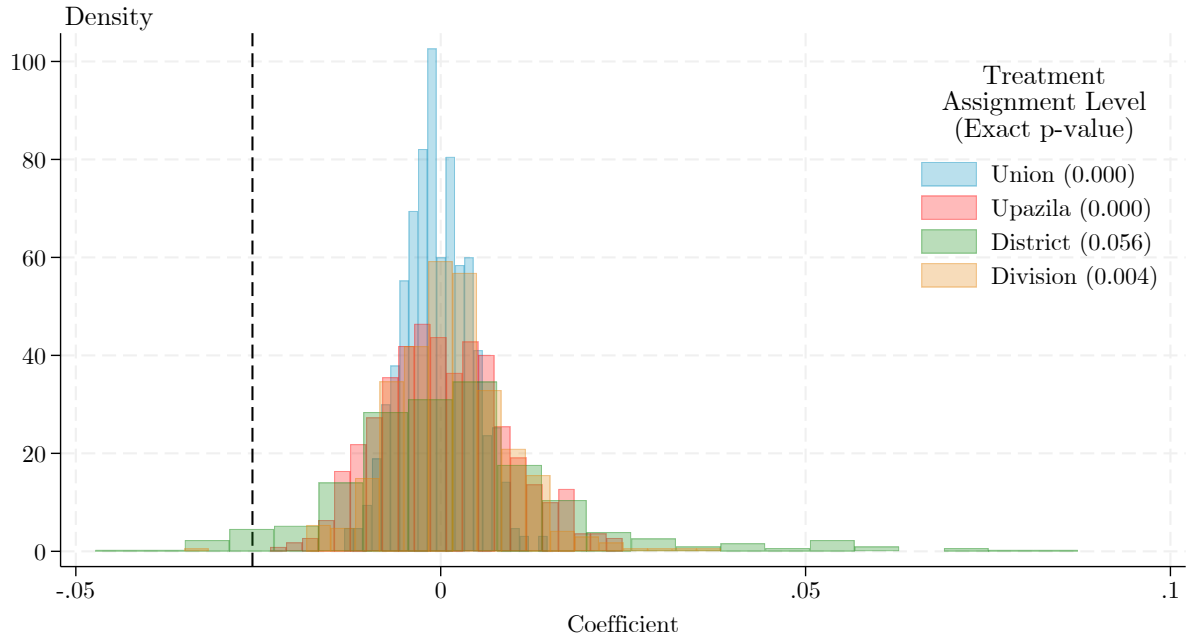
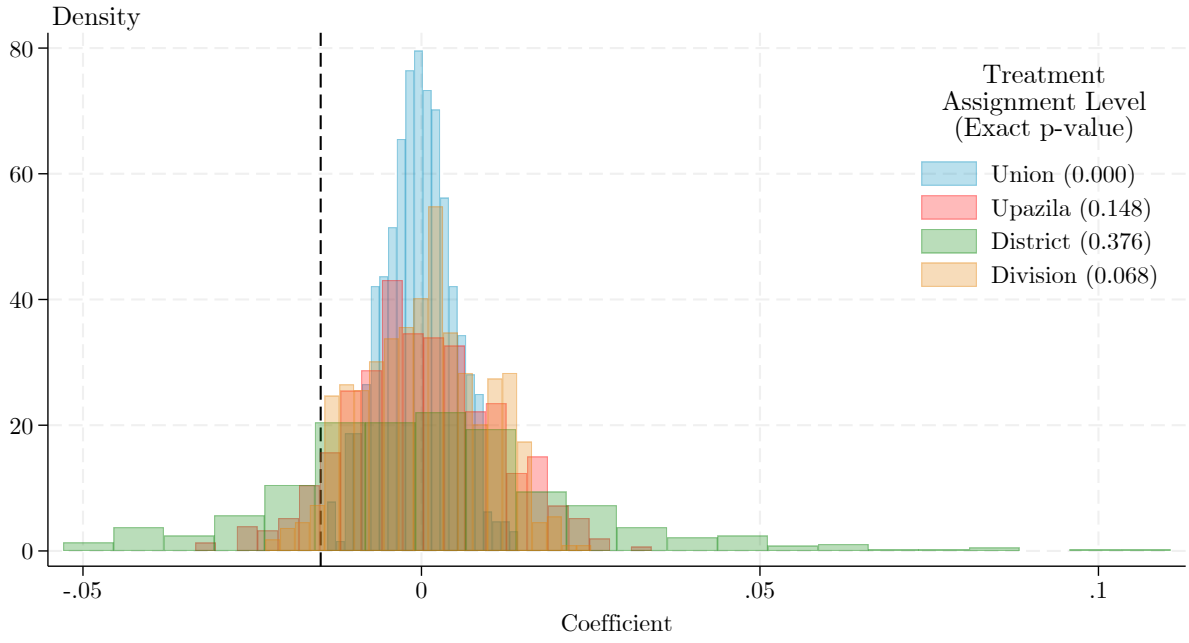


Figure E.6: Placebo Tests for the Impact of Floods on Nighttime Luminosity: Propensity Score Weights



Note: Figure E.5 plots histograms of placebo tests of the impact of flooding on average median nighttime luminosity using synthetic control weights, where the distribution of past flooding is drawn from different geographic levels. The dotted black line shows the truth. The legend reports exact two-sided p-values. Figure E.7 does the same with propensity score weights. The sample excludes control unions with centroids falling within 10 kilometers of the centroid of any treated union.

Figure E.7: Placebo Tests for the Impact of Floods on Nighttime Luminosity: No Weights

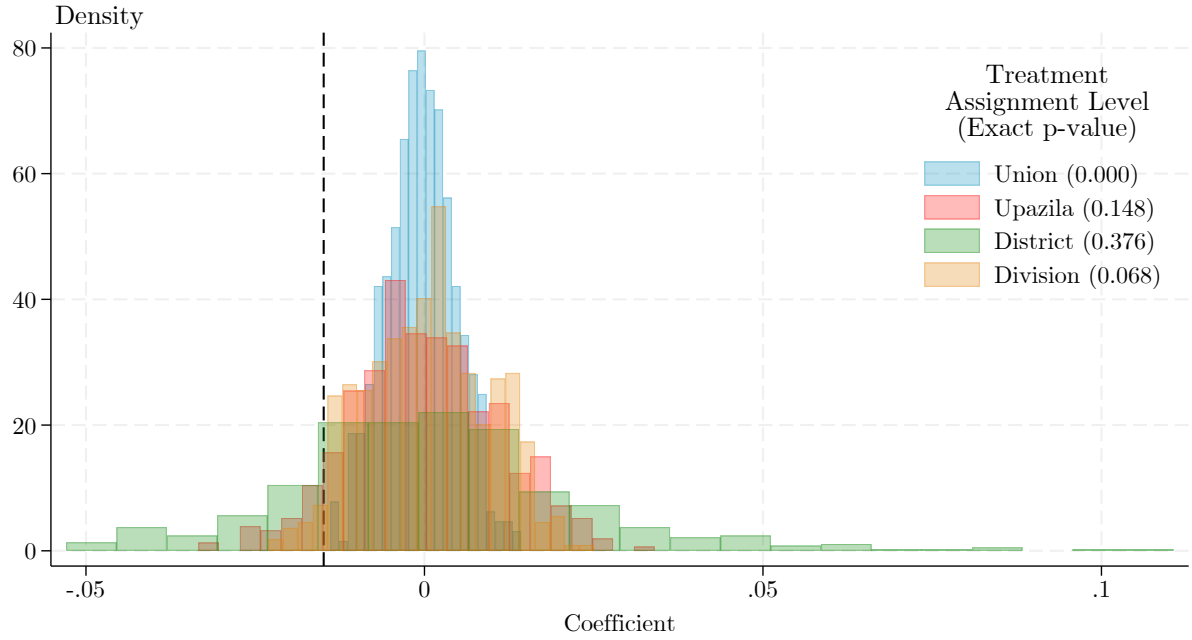
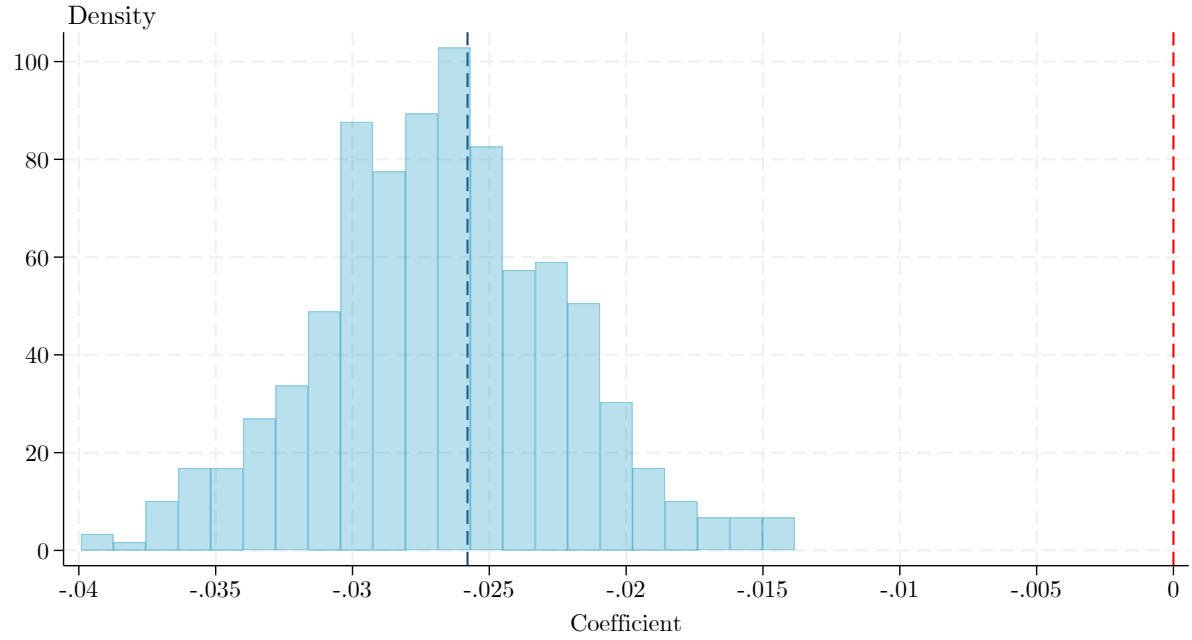


Figure E.8: The Impact of Floods on Nighttime Luminosity: Inference via Bayesian Bootstrap



Note: Figure E.5 plots histograms of placebo tests of the impact of flooding on average median nighttime luminosity using no weights, where the distribution of past flooding is drawn from different geographic levels. The dotted black line shows the truth. The legend reports exact two-sided p-values. Figure E.8 plots the distribution of coefficients from the synthetic control weight specification combined with a Bayesian bootstrap. The dotted blue line shows the true coefficient, and the dotted red line denotes 0. The sample excludes control unions with centroids falling within 10 kilometers of the centroid of any treated union.

Table E.3: The Heterogeneous Impact of Floods by Experience on Maximum Nighttime Luminosity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0161***	-0.0108***	-0.0112***	-0.0302***	-0.0132***	-0.0159***	-0.0337**	-0.0150***	-0.0194***
	(0.00577)	(0.00213)	(0.00201)	(0.00921)	(0.00265)	(0.00284)	(0.0137)	(0.00344)	(0.00362)
Post × Flood × Experience	0.0271**	0.00874**	0.00701***	0.0281**	0.0114**	0.0103***	0.0348	0.0154**	0.0129***
	(0.0121)	(0.00371)	(0.00206)	(0.0124)	(0.00457)	(0.00255)	(0.0220)	(0.00692)	(0.00367)
Sample	All	All	All	All	All	All	All	All	All
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.002	0.002	0.002	0.000	0.000	0.000	0.001	0.001	0.001
Conley Exp. p-value	0.055	0.055	0.055	0.006	0.006	0.006	0.005	0.005	0.005
Control Mean	-.112			-.108			-.121		
Observations	59,436	117,504	117,504	41,859	89,892	89,892	28,953	68,301	68,301
Clusters	2,763	3,436	3,436	2,295	3,168	3,168	1,876	2,909	2,909

Note: Table E.3 presents stacked difference-in-differences estimates of the impact of flooding on average maximum luminosity, standardized to the never-flooded union distribution. I define experience in the interaction term as the number of flooded years in that strata experienced in the five years leading up to treatment. All specifications include union fixed effects and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between synthetic control weights from Arkhangelsky et al. (2021), propensity score weights calculated from the strata, or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table E.4: The Heterogeneous Impact of Floods by Experience on Average Nighttime Luminosity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post × Flood	-0.0284***	-0.0128***	-0.0189***	-0.0405***	-0.0151***	-0.0250***	-0.0468***	-0.0167***	-0.0315***
	(0.00563)	(0.00203)	(0.00212)	(0.00911)	(0.00266)	(0.00303)	(0.0137)	(0.00343)	(0.00387)
Post × Flood × Experience	0.00930*	0.00610***	0.00816***	0.0135*	0.00775***	0.0132***	0.0101	0.00902***	0.0170***
	(0.00505)	(0.00145)	(0.00170)	(0.00731)	(0.00166)	(0.00208)	(0.0115)	(0.00225)	(0.00280)
Sample	All	All	All	All	All	All	All	All	All
Weights	Synthetic	Propensity	None	Synthetic	Propensity	None	Synthetic	Propensity	None
Spillover Radius	10 km.	10 km.	10 km.	20 km.	20 km.	20 km.	30 km.	30 km.	30 km.
Conley p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Conley Exp. p-value	0.002	0.002	0.002	0.001	0.001	0.001	0.000	0.000	0.000
Control Mean	-.132			-.129			-.14		
Observations	58,779	117,504	117,504	41,508	89,892	89,892	28,701	68,301	68,301
Clusters	2,727	3,436	3,436	2,273	3,168	3,168	1,861	2,909	2,909

Note: Table E.4 presents stacked difference-in-differences estimates of the impact of flooding on average masked mean luminosity, standardized to the never-flooded union distribution. I define experience in the interaction term as the number of flooded years in that strata experienced in the five years leading up to treatment. All specifications include union fixed effects and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between synthetic control weights from Arkhangelsky et al. (2021), propensity score weights calculated from the strata, or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Figure E.9: The Impact of Floods on Temporary Insolvency—Alternative Spillovers

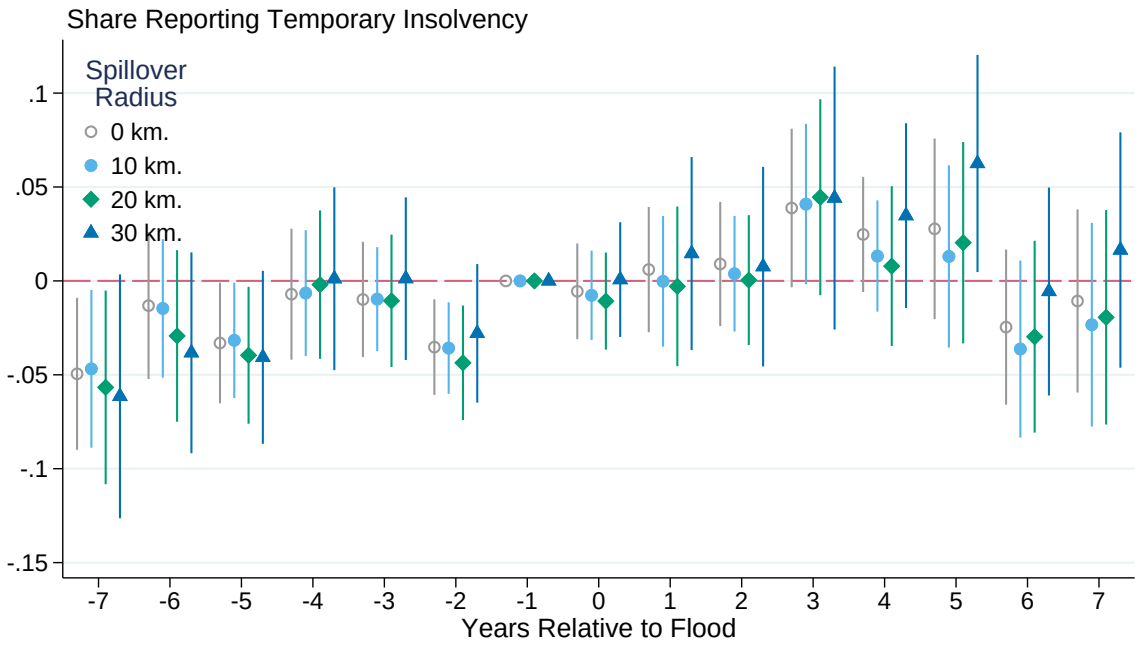
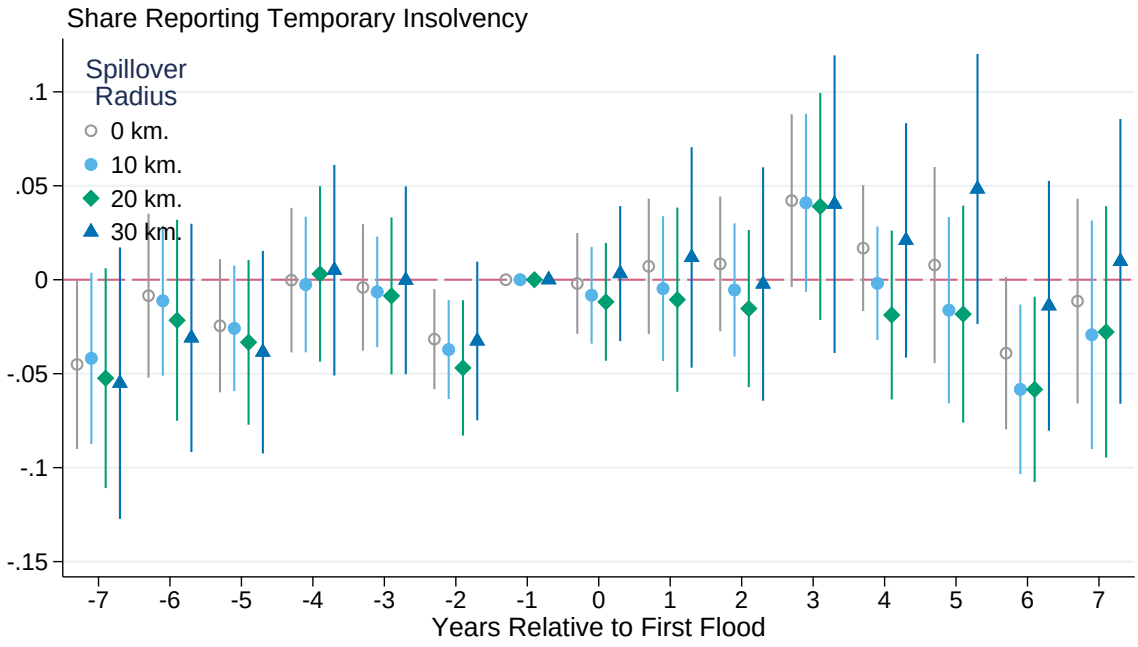


Figure E.10: The Impact of a First Flood on Temporary Insolvency—Alternative Spillovers



Note: Figure E.11 plots the coefficients from an event study of flooding using the sample to estimate the effect of an additional flood. Figure E.23 plots the coefficients from an event study of flooding using only the first-flood comparisons to isolate the impulse response function. The outcome in both is self-reported rates of temporary insolvency from the SVRS data—see Appendix Section C.3 for details. All regressions include union and strata by relative time fixed effects with standard errors clustered at the union level and machine learning generated propensity score weights. The spillover radius denotes the distance at which I exclude control unions whose centroids fall within a given radius of any treated union centroid.

Figure E.11: The Impact of Floods on Insolvency—Alternative Spillovers

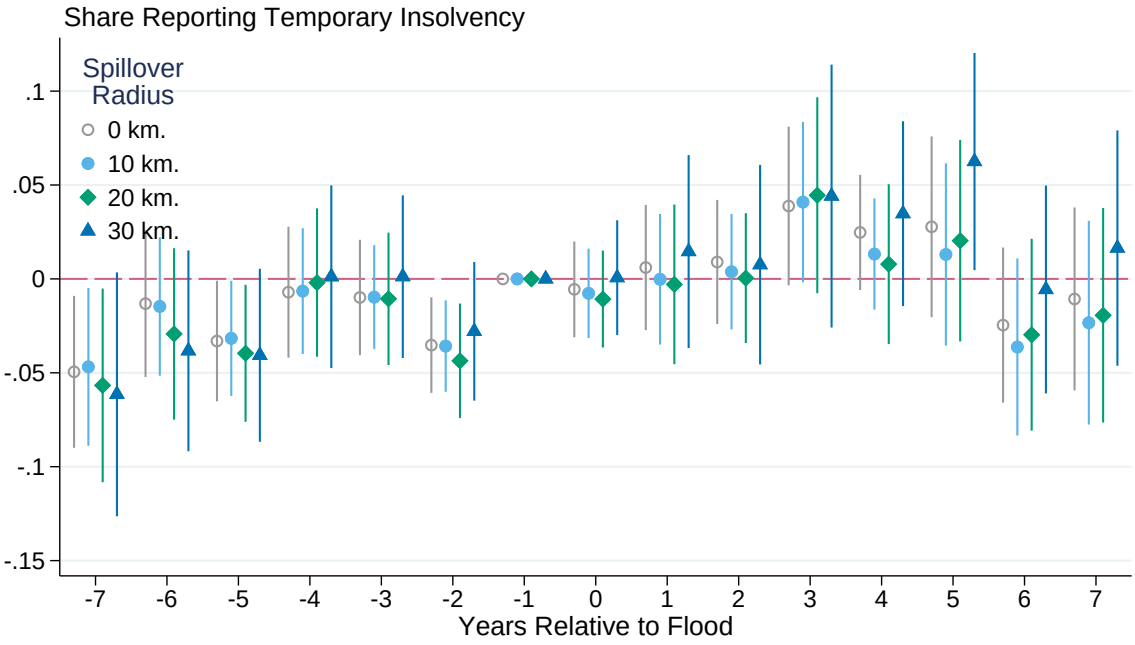
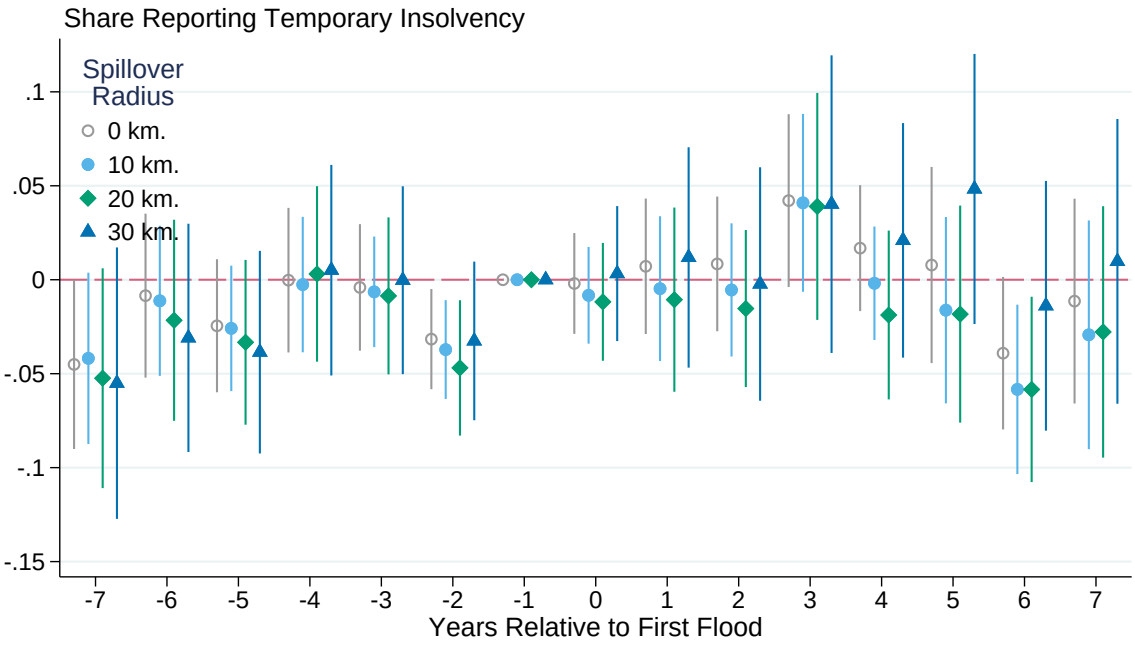


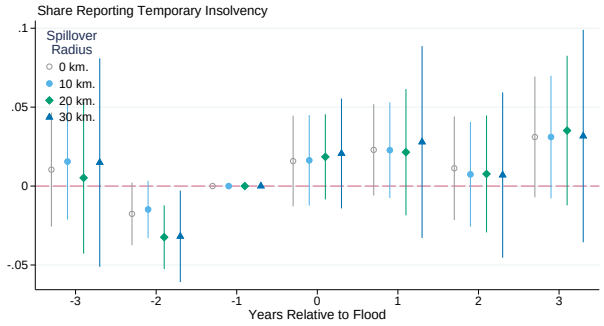
Figure E.12: The Impact of a First Flood on Insolvency—Alternative Spillovers



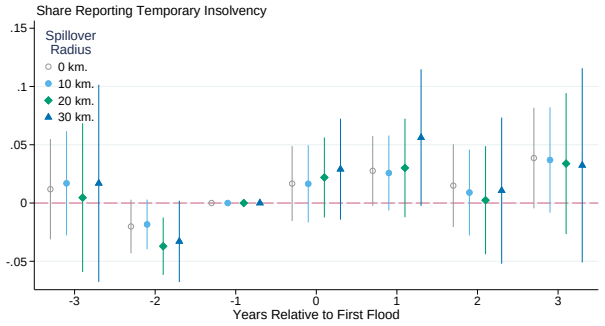
Note: Figure E.11 plots the coefficients from an event study of flooding using the sample to estimate the effect of an additional flood. Figure E.23 plots the coefficients from an event study of flooding using only the first-flood comparisons to isolate the impulse response function. The outcome in both is the combined self-reported rates of temporary and permanent insolvency from the SVRS data—see Appendix Section C.3 for details. All regressions include union and strata by relative time fixed effects with standard errors clustered at the union level and machine learning generated propensity score weights. The spillover radius denotes the distance at which I exclude control unions whose centroids fall within a given radius of any treated union centroid.

Figure E.13: The Impact of Floods on Temporary Insolvency—Fixed Sample

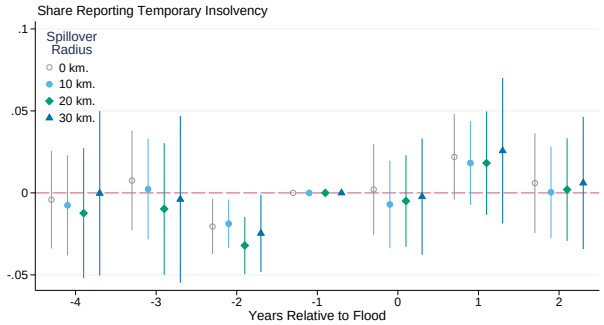
(a) All Floods—Sample #1



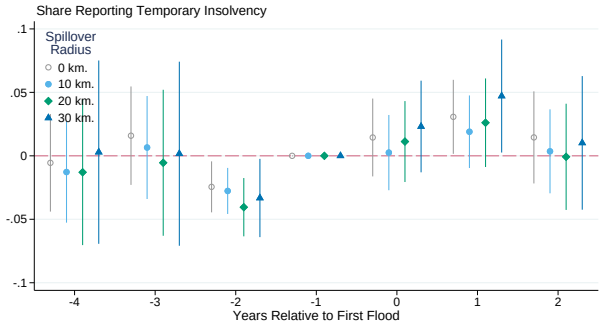
(b) First Floods—Sample #1



(c) All Floods—Sample #2



(d) First Floods—Sample #2



Note: Figure E.13 plots event study coefficients of flooding on households' self-reported economic condition holding fixed the sample of unions over each horizon. All regressions include union and strata by relative time fixed effects, with standard errors clustered at the union level and machine learning generated propensity score weights.

Figure E.14: The Marginal Impact of Floods on Non-Agricultural Labor—Men

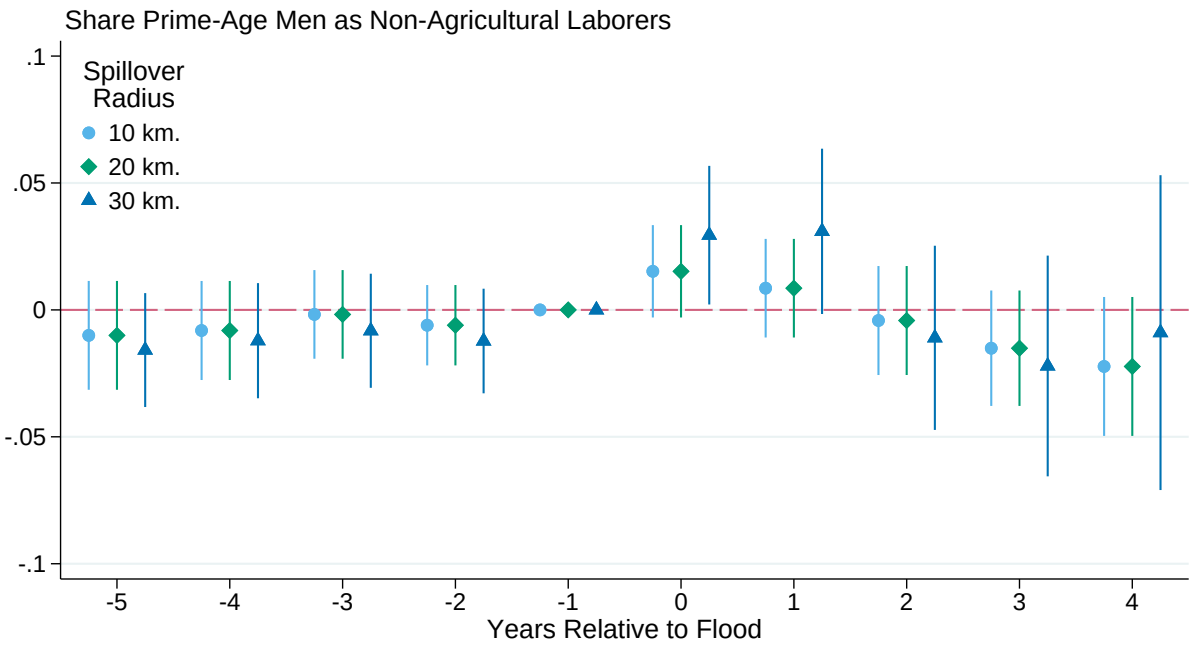
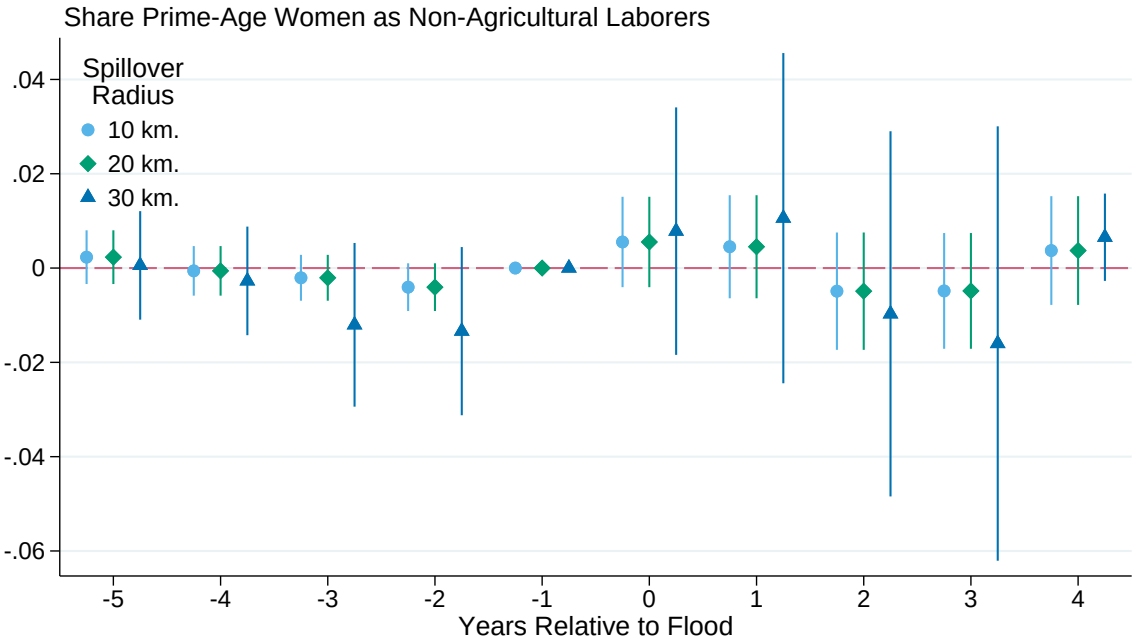


Figure E.15: The Marginal Impact of Floods on Non-Agricultural Labor—Women



Note: This figure plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. All regressions include age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.

Figure E.16: The Marginal Impact of Floods on Household Work—Men

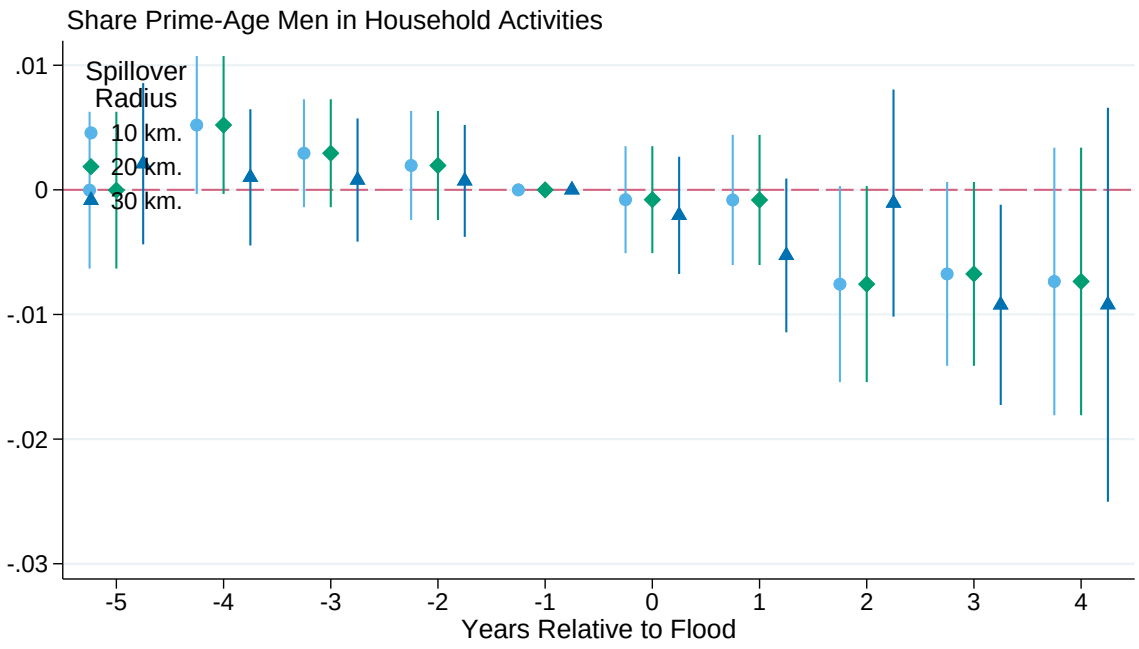
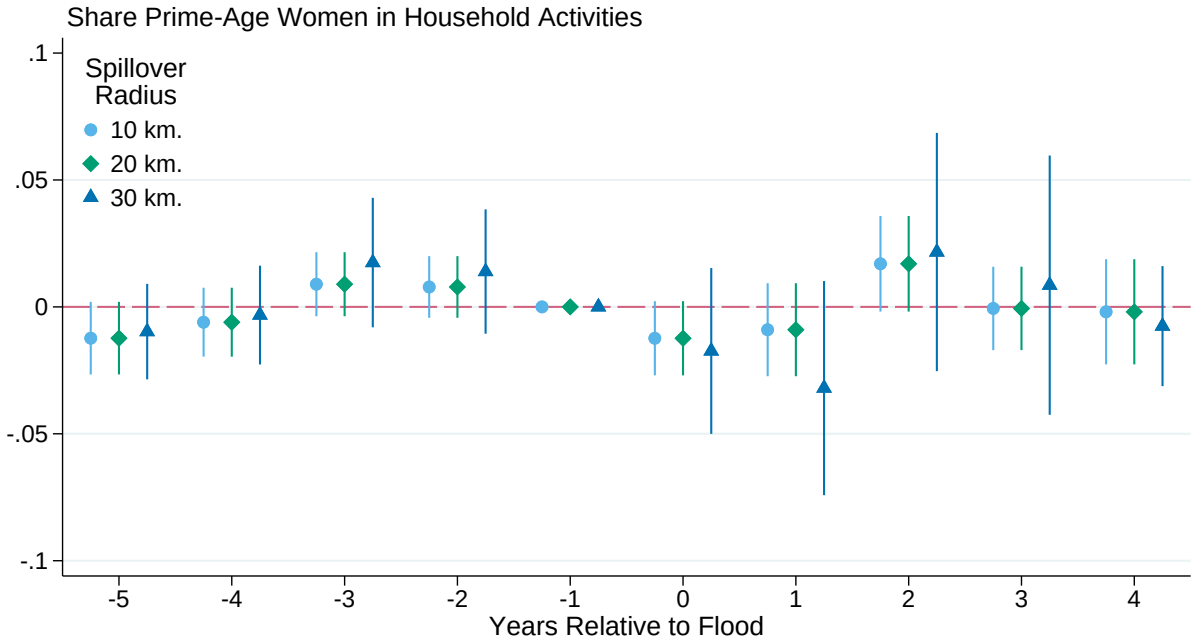


Figure E.17: The Marginal Impact of Floods on Household Work—Women



Note: This figure plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. All regressions include age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.

Figure E.18: The Marginal Impact of Floods on Children Without Schooling—Boys

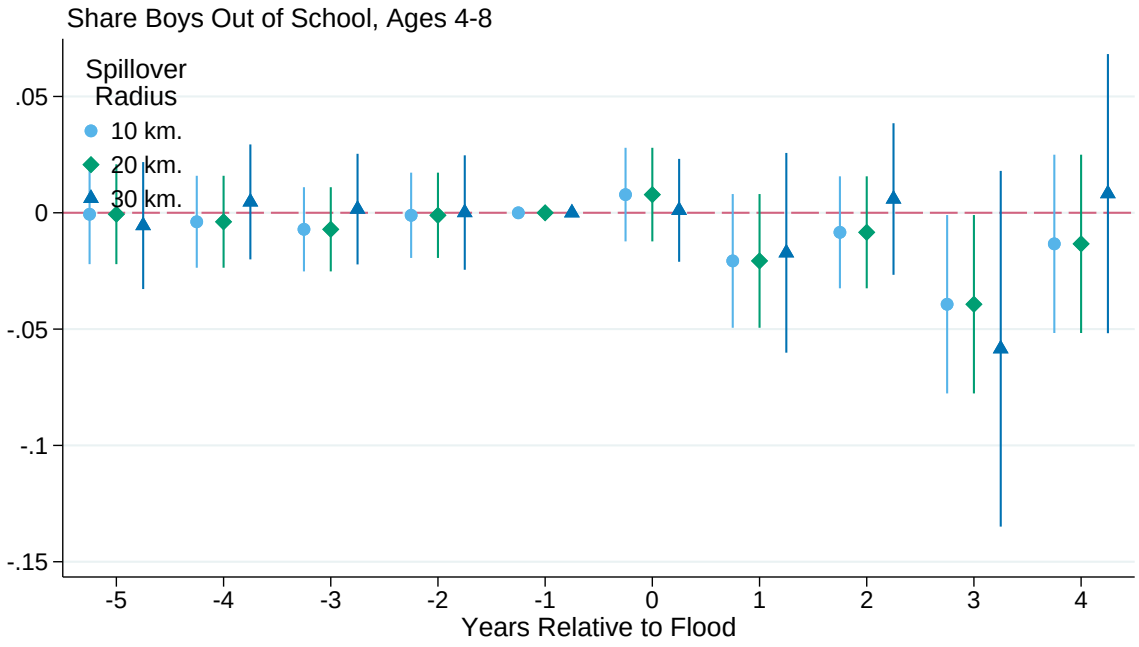
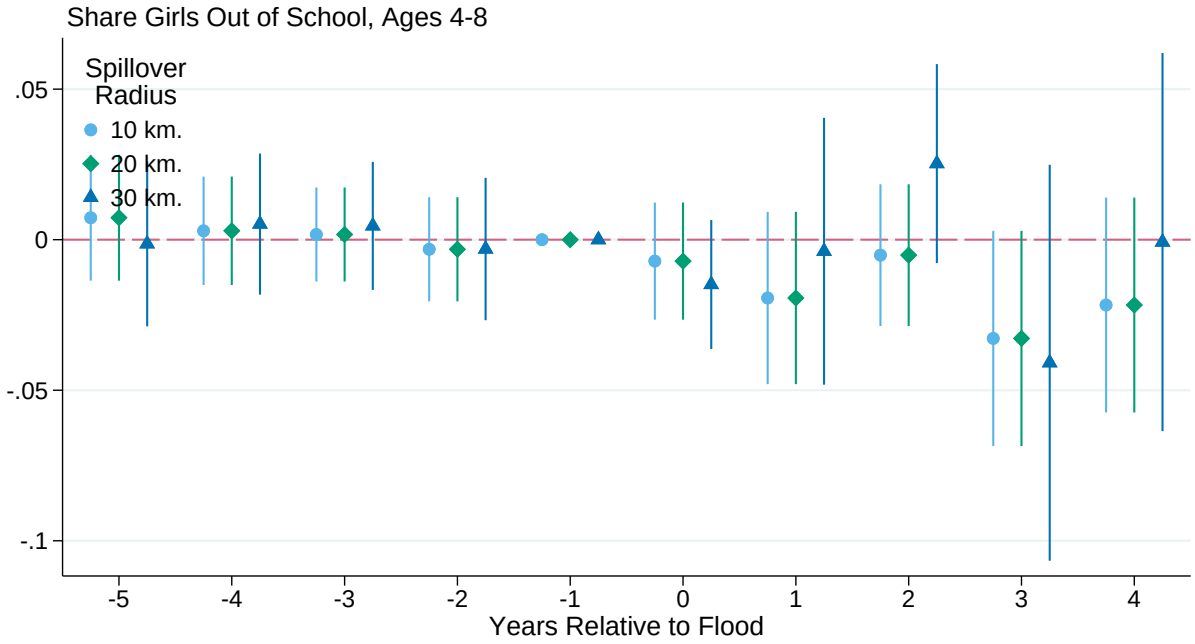


Figure E.19: The Marginal Impact of Floods on Children Without Schooling—Girls



Note: This figure plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. All regressions include age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.

Figure E.20: The Impact of Floods on Self-Reported Temporary Insolvency

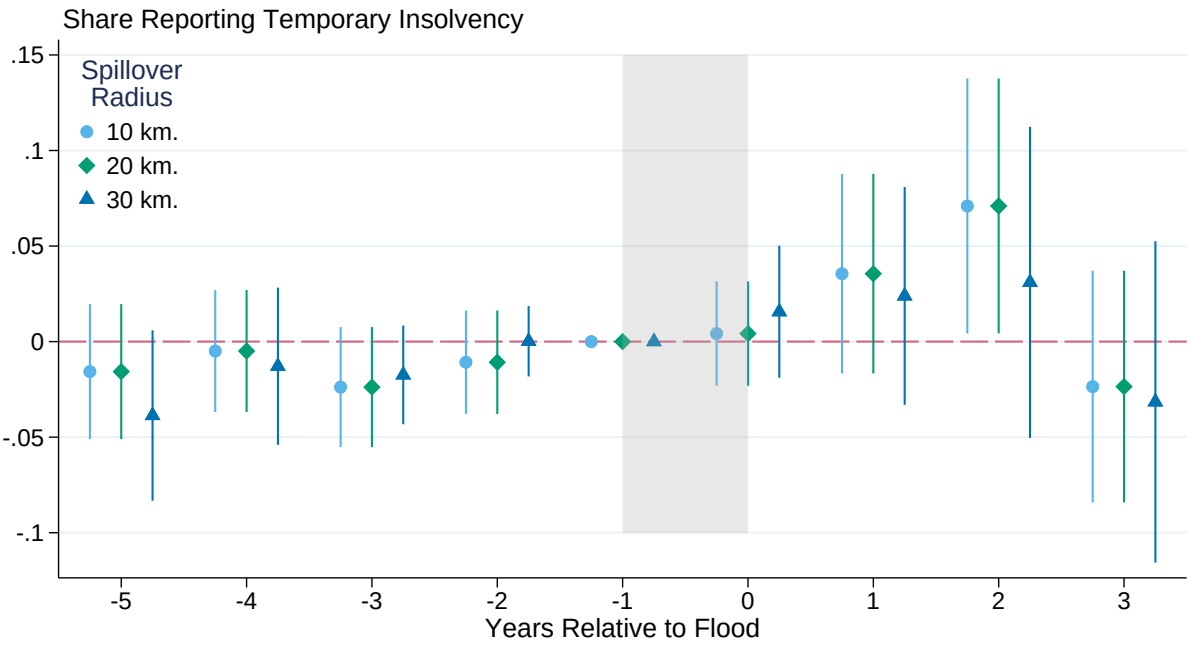
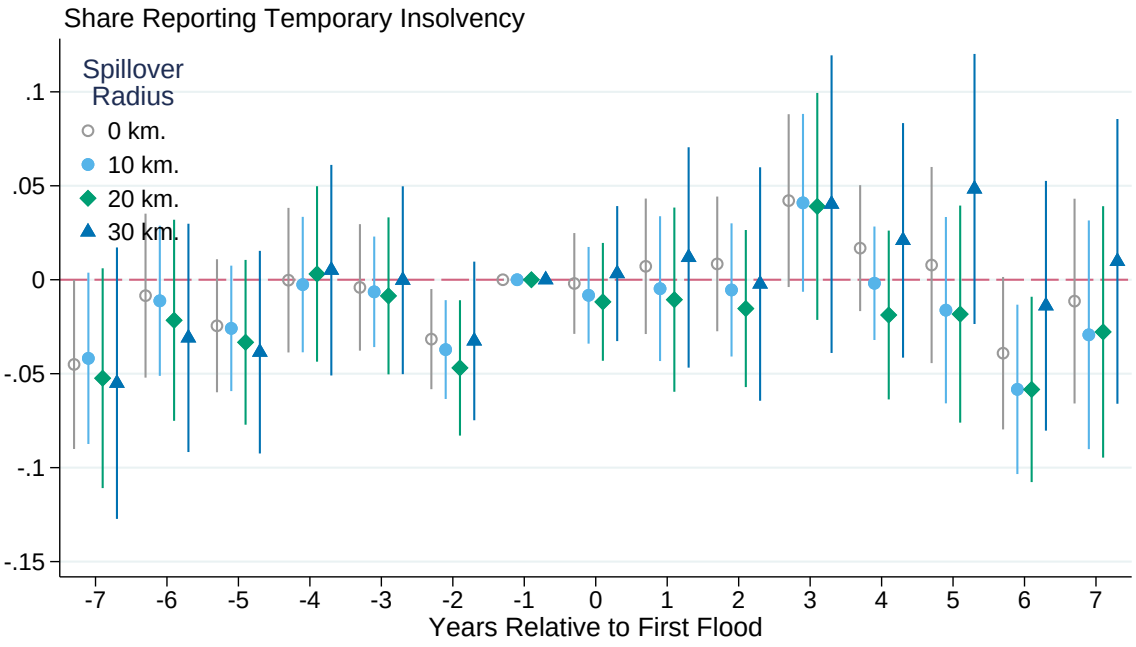


Figure E.21: The Impact of a First Flood on Self-Reported Temporary Insolvency



Note: Figure E.22 plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. Figure E.23 plots the coefficients from an event study of flooding using only the first-flood comparisons to isolate the impulse response function. Temporary insolvency comes from households' self-reported descriptions of their economic conditions in the Sample Vital Registration System microdata. All regressions include union, treatment cohort, relative time, and strata by relative time fixed effects, with standard errors clustered at the union level.

Table E.5: The Impact of Floods on Self-Reported Insolvency

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post × Flood	0.0120	0.0117	0.00975	0.0107	0.0145*	0.0193**	0.0118
	(0.00793)	(0.00726)	(0.00777)	(0.00730)	(0.00873)	(0.00817)	(0.0105)
Sample	All	Marginal	All	Marginal	All	Marginal	All
Weights	Propensity	None	Propensity	None	Propensity	None	Propensity
Spillover Radius	0 km.	0 km.	10 km.	10 km.	20 km.	20 km.	30 km.
Control Mean	.171	.171	.172	.172	.177	.177	.184
Observations	1916721	1916721	1817327	1817327	1559496	1559496	1299729
Clusters	1,110	1,110	1,048	1,048	894	894	752

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post × Flood	0.0217**	0.0154*	0.0180**	0.0130	0.0209**	0.0218**	0.0189
	(0.00884)	(0.00865)	(0.00859)	(0.00852)	(0.00962)	(0.00917)	(0.0116)
Sample	First	First	First	First	First	First	First
Weights	Propensity	None	Propensity	None	Propensity	None	Propensity
Spillover Radius	0 km.	0 km.	10 km.	10 km.	20 km.	20 km.	30 km.
Control Mean	.167	.167	.168	.168	.173	.173	.183
Observations	1590265	1590265	1516851	1516851	1287925	1287925	1066996
Clusters	937	937	887	887	750	750	626

Note: Table E.5 presents stacked difference-in-differences estimates of the impact of flooding on households' self-reported economic condition in the SVRS microdata (see Appendix Section C.3 for details). All specifications include union fixed effects, calendar year fixed effects, relative time fixed effects, and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights calculated from the strata or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Table E.6: The Heterogeneous Impact of Floods by Experience on Self-Reported Insolvency

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post × Flood	0.0195**	0.0141*	0.0161*	0.0120	0.0204**	0.0209**	0.0191*
	(0.00866)	(0.00810)	(0.00851)	(0.00810)	(0.00944)	(0.00879)	(0.0114)
Post × Flood × Experience	-0.0244**	-0.00687	-0.0210*	-0.00392	-0.0193	-0.00469	-0.0229
	(0.0106)	(0.0107)	(0.0116)	(0.0124)	(0.0148)	(0.0149)	(0.0180)
Sample	All	All	All	All	All	All	All
Weights	Propensity	None	Propensity	None	Propensity	None	Propensity
Spillover Radius	0 km.	0 km.	10 km.	10 km.	20 km.	20 km.	30 km.
Control Mean							
Observations							
Clusters	.171	.171	.172	.172	.177	.177	.184
	1916721	1916721	1817327	1817327	1559496	1559496	1299729
N_clust	1,110	1,110	1,048	1,048	894	894	752

Note: Table E.6 presents stacked difference-in-differences estimates of the impact of flooding on households' self-reported economic condition in the SVRS microdata (see Appendix Section C.3 for details). I define experience in the interaction term as the number of flooded years in that strata experienced in the five years leading up to treatment. All specifications include union fixed effects, calendar year fixed effects, relative time fixed effects, and stratum by relative time fixed effects. I cluster standard errors at the union level. The spillover radius denotes both the minimum distance I require between a control union and any treated union as well as the radius of the Conley (1999) standard errors. Stars denote $p-$values below .1 (*), .05 (**), and .01 (***). The weights used vary across specifications between propensity score weights calculated from the strata or none. The top row includes the sample of marginal effects; the bottom row limits to the first-flood comparison samples.

Figure E.22: The Marginal Impact of Floods on Literacy

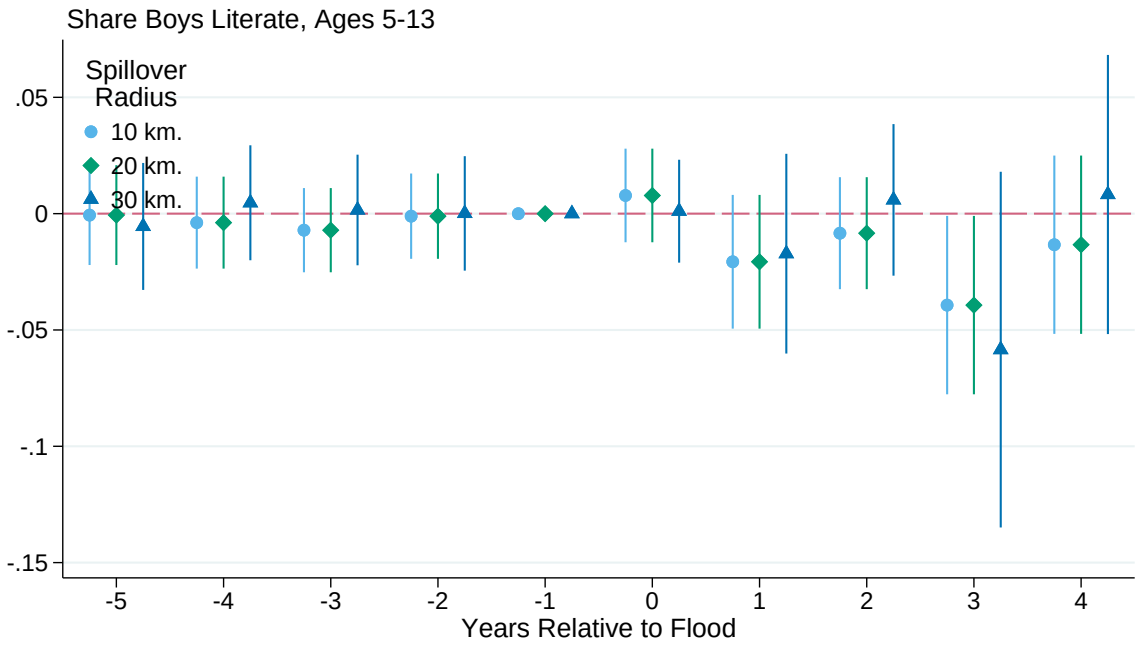
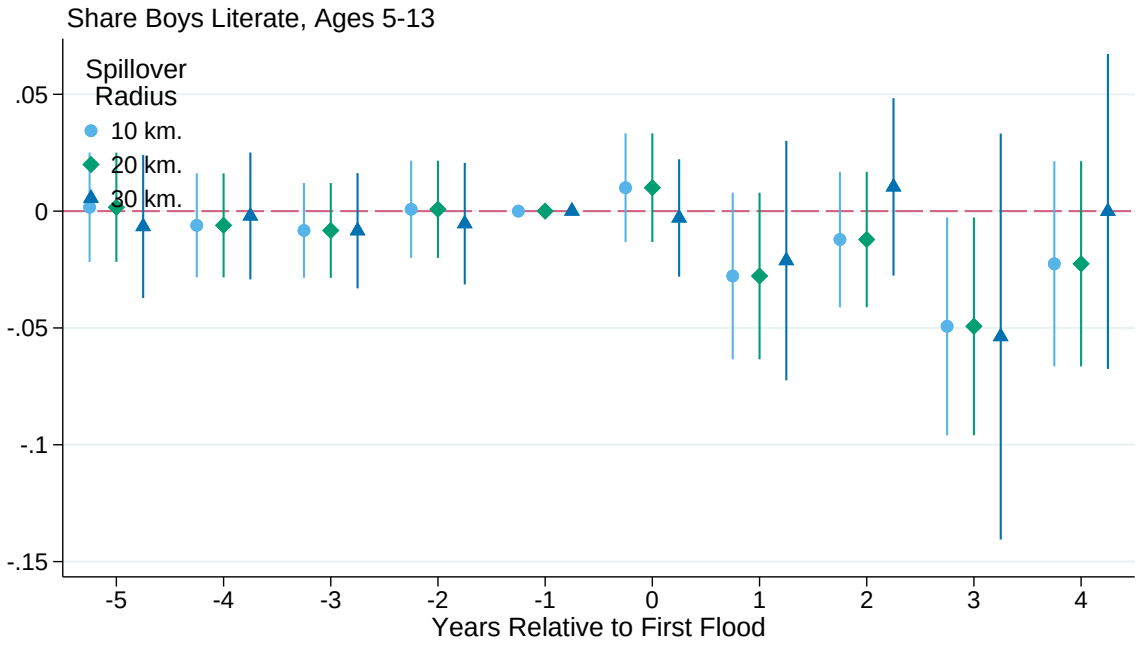


Figure E.23: The Impact of a First Flood on Self-Reported Temporary Insolvency



Note: Figure E.22 plots the coefficients from an event study of flooding using the sample to estimate the marginal effect of an additional flood. Figure E.23 plots the coefficients from an event study of flooding using only the first-flood comparisons to isolate the impulse response function. Literacy comes from the Sample Vital Registration System microdata, restricting to children with ages 5 to 13. All regressions include gender, age, union, and strata by relative time fixed effects, with standard errors clustered at the union level.