The Long-Run Environmental Consequences of Land Development*

Congyan Han, Lu Han, Vincent Yao†

January 13, 2024

Preliminary and Incomplete. Please do not share!

Abstract

We study the long-run environmental impacts of land development activities using flood claims and land use data at the zip code level. Employing long differences and instrumental variable approaches, we find that an increase in developed land is associated with a significant increase in flood claims after 8 years over a 15-year span. This effect is primarily observed in land developed from cropland and tree cover and varies by the initial development conditions of an area. By linking land with the demographic characteristics of residents, our study demonstrates that the flood cost attributable to land development is not borne equally spatially or between demographic groups, with a few hotspots and minority neighborhoods experiencing higher costs.

Keywords: Land Use, Flood Damage, Environmental Effects, Right to Work Law

JEL Classification: K25, Q24, Q28, Q51, Q56, Q58

*

†Han: Wisconsin School of Business, University of Wisconsin-Madison, chan74@wisc.edu. Han: Wisconsin School of Business, University of Wisconsin-Madison, lu.han@wisc.edu. Yao: J. Mack Robinson College of Business, Georgia State University, wyao2@gsu.edu.

1. Introduction

Solow's growth model, as highlighted in Solow (1956), underscores the importance of technological progress and factors of production, including land, in economic development. Land serves as a finite and essential resource that underpins various facets of economic growth. It provides the foundational support for agriculture, not only ensuring food security, but also generating income for millions of people globally (Deininger and Feder, 2001). Ample land supply also constitutes a requisite condition for attracting domestic and foreign investments, igniting urbanization, and cultivating employment opportunities, all of which are indispensable for economic advancement (e.g., Banerjee and Duflo, 2007). Furthermore, land serves as a valuable asset for real estate development and industrial activities, significantly contributing to a nation's GDP (Galor and Weil, 2000; Gollin et al., 2002).

In the United States, land use is subject to rigorous regulation by local governments, aiming to achieve a delicate balance between economic development, environmental sustainability, and social equity. However, concerns persist regarding the effectiveness and efficiency of these regulations (Gyourko et al., 2008, 2021). One crucial but often overlooked aspect is the long-term consequences of land use decisions made in the name of economic development. Since politicians are constrained by term limits and the costs and benefits of land use policies unfold over varying timeframes, most analyses tend to focus primarily on short-term economic benefits while neglecting the flood costs that take longer to materialize. For instance, in the extended debate over Right-to-Work (RTW) laws, a notable study by Holmes (1998) compares the differential growth rates of manufacturing activity when crossing state borders from an antibusiness state to a probusiness state. This unbalanced approach can lead to biased policy evaluations and have significant repercussions for intergenerational equity, as the long-term costs associated with land use decisions are borne by future generations.

This study assesses the long-term environmental consequences and temporal aspects of land development activities. Our assessment quantifies flood costs by analyzing flood damages utilizing data from the Federal Emergency Management Agency's (FEMA) National Flood

Insurance Program (NFIP) claims. Meanwhile, land use data are derived from high-resolution geospatial information sourced from the Land Change Monitoring, Assessment, and Projection (LCMAP). We conduct a comprehensive analysis encompassing the entire contiguous United States, connecting zip code-level data on these variables over the period 2001 to 2019.

The primary empirical challenge we face is the correlation between changes in land development and numerous confounding factors that also impact flood risk. To overcome these challenges, we employ several strategies to identify variations in local land development. Our first strategy involves using stacked long differences (LD) to capture the impact of land development on flood damages over different horizons, ranging from one year to 15 years. This approach controls for state-by-cohort year fixed effects and incorporates various zip code-level covariates. In addition, we distinguish between various sources of developed land, which depend on local natural endowments, and initial development conditions, which affect the economic incentives of land development. The second strategy entails a panel data analysis that leverages state-level differences in the adoption of RTW laws, a prominent pro-business policy, and the initial development states for a given zip code as instrumental variables based on matched data. We apply a Two-Stage Least Squares (2SLS) approach, demonstrating that these two instruments are highly predictive of land development and have no direct effect on flood damages, thus satisfying the exclusion restrictions.

Our baseline results reveal a progressive increase in the relationship between changes in land development and NFIP flood claims over longer horizons. The relationship lacks statistical significance over a 1- to 8-year horizon, but becomes positive and significant after 8 years, with both the magnitude of the estimates and level of significance demonstrating a continual rise over time. Over the 15-year horizon, a one-percent increase in developed land is associated with a 2.08-percent increase in NFIP flood claims, a significant result at the 1% level. Consistently, the estimates from panel data regressions capture a smaller short-term relationship between the contemporaneous change in land development and NFIP flood claims.

In the 2SLS regression, the first-stage results demonstrate that, compared to the control group containing zip codes in non-RTW states, low-developed zip codes in RTW states experience a significantly higher increase in land development by 0.065 percent, significant at the 1% level; high-developed zip codes in RTW states show a significantly lower increase in land development by -0.225 percent, also significant at the 1% level. Consistently, the 2SLS regression yields an estimate of 1.640 percent, slightly larger in magnitude than the OLS estimate. This difference is likely attributed to the instruments capturing the change in developed land over the long run after the enactment of RTW laws, rather than solely representing the contemporaneous change in land development in the OLS regressions. This result suggests that an increase in land development has a causal positive effect on the change in flood risk.

One confounding factor that might explain the results is the change in non-developed land, rather than developed land, due to the zero-sum nature of changes in developed and non-developed land. A recent study by Taylor and Druckenmiller (2022), which investigates the value of wetlands as a public good in the form of ecosystem services, reveals a significant positive correlation between the loss of wetlands and an increase in flood claims. Employing an upstream-downstream difference-in-differences (DID) analysis, they identify the mechanism through which wetland loss amplifies flood damage, highlighting the critical role of wetlands in providing valuable protective ecosystem services that mitigate flood damage, as the removal of such wetlands for alternative uses could lead to an escalation in flood-related damage.

When we include the change in non-developed land and also by individual land covers in the regression, the coefficient on changes in developed land ranges between 1.993-2.094 and remains statistically significant at the 1% level. The coefficients on changes in undeveloped land, either as a whole or by individual types of source land, are small and statistically insignificant. These findings suggest that the increase in flood damages is influenced by the expansion of developed land—representing the increase in human development activities—rather than the reduction of natural land, such as wetlands and tree cover (treeland), for non-development reasons.

Notably, when we examine the sources of developed land from 2001 to 2019 in the US, treeland and cropland stand as the two predominant sources, contributing to nearly 90% of all land developed in recent decades. Furthermore, the coefficients of change in developed land from cropland or treeland are 4.309 and 3.578 percent, respectively, and both are significant at the 1% level. The coefficient of change in developed land from wetland is 4.827 percent, though it is not statistically significant. The coefficients for developed land from other sources are small and not significant. This suggests that not all land developments have equal effects on flood risk, and the initial land endowment and, consequently, the source of land development matter.

The effect of land development on flood risk also depends on the initial development level and density. Our analysis demonstrates that the initial conditions in land development are highly predictive of subsequent land development. Zip codes with higher initial development conditions have experienced significantly higher growth in developed land in the following years. In total, approximately 80% of the land developed during the 15-year period is in zip codes with high initial conditions. There is also considerable variation in land development sources between groups. For areas with high initial conditions, cropland and treeland are the two most important sources of development, while for areas with medium initial conditions, treeland becomes a more crucial source than cropland. In contrast, the significant effect observed in the baseline analysis is mainly due to zip codes with high initial conditions, where the majority of land development has occurred, and to a lesser degree from those with medium initial conditions. During the 15-year horizon, a one-percent increase in developed land is associated with a 1.307-percent increase (significant at the 5% level) in NFIP flood claims for areas with medium initial conditions and a 3.330-percent increase (significant at the 1% level) for areas with high initial conditions.

Next, we assess the distributional effects by linking developed land with the demographic characteristics of residents and also quantify the flood risk attributable to land development. Our analysis indicates that the most important land developed in the past two decades—those from cropland and treeland—happen to be in areas where there is a low proportion of minority and a high proportion of low-income population and where housing and population are sparse. As more land is developed, these areas see a significant decrease in the share of the minority and low-income population. There are also demographic differences between areas with more cropland or more treeland available. Overall, land development, particularly from treeland, is associated with a significant decrease in the proportion of minorities and the proportion of low-income population. Furthermore, areas with high initial conditions have a higher proportion of minorities but a lower proportion of low-income population, and have experienced a significant increase in minorities and low-income population as more land is developed.

Using the coefficients for the change in developed land interacted with land development sources and initial conditions, we calculate the lifetime flood cost of the developed land to be \$2.59 billion in total and \$2,164 per hectare (ha, equivalent to 2.47 acres). Based on estimated and actual flood damage, 14.7% of the flood damage that has increased from 2001 to 2016 is attributable to land development. The unit costs per ha, representing the long-run social cost of land development not paid by the private market, are equivalent to 22% of the market value of land. However, there is significant variation in the estimated flood cost spatially and between demographic groups, primarily due to differences in natural endowment and initial development conditions. The top 1% of counties have to bear 95% of the cost, with the highest burden falling on counties in the Houston metro area. Most of the hotspots, using different transformations, are located in Texas, Louisiana, Virginia, Maryland, Maine and Florida, all coastal states. Zip codes with a higher proportion of black population and high-income majority neighborhoods also experience a disproportionately higher flood risk.

Very few decisions influence the lives of local residents more than a land development decision, given the essential role of land development in economic growth. However, land development is costly in terms of land acquisition cost, construction cost, and long-run environmental cost. While the literature on land development has focused almost exclusively on the first two cost components, this paper provides a first measure of the long-run environmental cost associated with land development. We find enormous spatial variation in land development as well as its flood costs across the US. Our reduced-form analysis establishes three novel stylized facts. First, areas with a higher proportion of minorities tend to have higher initial development conditions and less land available for future development, but relatively more cropland than other covers among the available land. Our analysis indicates that areas with high initial conditions and more land developed from cropland are associated with significantly higher flood damage. This exposes original minority residents to high flood risk in the long run as land is developed. Second, these same areas have seen a significant increase in the proportion of minority as new land is developed, increasing the share of minority that have been exposed to the long-run flood risk of land development. Third, while the low-income population tends to live in areas with lower initial conditions initially, land development in areas with high initial conditions have attracted more low-income population, exposing them to the long-run flood risk of land development.

In the remainder of the paper, we propose a framework to rationalize the land development decisions of households and hence the variation in the long-run flood risk across space and social demographics. At the center of the model is the intertemporal trade-off involved in land development decisions and the heterogeneity in the initial land endowments, as documented above. We start by proposing a simple three-period economy where households have heterogeneous income and initial fraction of developed land. Higher-income and minority households are more likely to be in areas with more developed land. The marginal cost of land development increases with the initial developed fraction while the marginal utility from land development decreases with the initial income. Residents collectively choose the amount of land development, which in turn determines their income next period and the flood risk in the third period, both as a function of the land development. The optimal land development is determined by local residents' marginal willingness to give up income in the second period in exchange for a lower risk of flooding in the third period. In equilibrium, lower-income and minority households choose to develop more land, as their marginal utility from the short-run growth outweighs marginal cost of land development. To the extent that migration costs are sufficiently high, the model predicts that marginalized households are more likely to reside in the high flood risk areas.

(In progress) We then take the model prediction to the data. By merging the individual-level demographics and mobility data from the InfoUSA with the land development and flood risk data, our dataset constitutes one of the first large-scale panel datasets with detailed information on household demographics, mobility, climate risk and land development decisions for the entire US. Using these rich data, we estimate a land development model where households choose how much land to develop given their location, where locations differ in the short-run income growth and in the long-run flood risk. Preliminary results suggest that there are economically and statistically significant differences in lower- and higher-income households' marginal-willingness-to-pay (MWTP) to avoid the flood risk associated with land development, which in turn affects their land development decisions. Compared to higher income and white households, the observed flood risk in residential locations of low-income and minority households reveals a smaller willingness to give up current consumption in exchange for a lower flood risk 15 years from now. These MWTP differences are consistent with the intertemporal trade-off highlighted in the model. We further quantify the influence of heterogeneity of MWTP on land development outcomes with a counterfactual exercise. In particular, we ask how the distribution of the land development would differ if we gave lower-income households the same MWTP to avoid flood risk as high-income households when they make their land development decisions in 2001. In the next version of this paper, we will report the results.

The paper is organized as follows. After the literature review, Section 2 discusses some institutional details of land use in the United States, along with a summary of scientific research in the relationship between land development and flood risk. This is followed by the explanation of the data and sample design in Section 3. We present the empirical identification and the empirical baseline analyses in Section 4 and the heterogeneity analysis in Section 5. Section 6 presents various distributional analysis of flood risk. Section 7 concludes.

Literature Review We study the long-term effect, particularly the environmental implications, of economic development specifically through land development. Existing literature has predominantly emphasized the employment benefits associated with various economic development programs and policies aimed at attracting manufacturing plants to specific regions. For example, Holmes (1998) shows that on average, the manufacturing share of total employment in a county increases by approximately one third on the probusiness side, as evidenced by border analysis, due to an overall effect of the state's probusines policies.[1] Another example is Kline and Moretti (2014a) who examine the long-term consequences of the Tennessee Valley Authority's initiatives, a prime example of the "big push" development strategy, and uncover differential effects on agricultural and manufacturing employment stemming from federal transfers.[2] Our paper adds a balanced view to the literature by focusing on the long-term consequences of flood risks associated with increased land development activities. We find that although land development does not exert significant flood damage in the short run, the effect becomes pronounced in the long run. Overlooking these long-term costs can lead to an incomplete evaluation of economic development policies, as they entail intergenerational transfers with far-reaching consequences.

Our paper is also situated within a substantial body of literature examining the relationship between land use regulation and housing prices, as well as broader economic outcomes. Saiz (2010) employs comparable satellite-derived geospatial data to estimate local housing supply elasticities, underscoring geography as a crucial factor in determining housing supply inelasticity. Additionally, various studies investigate the impact of land use regulations on land values by analyzing historical shifts in regulations (Zhou et al., 2008; Libecap and Lueck, 2011). Turner et al. (2014) provides estimates suggesting significant adverse effects of regulation on land values and overall welfare in these regions. A large complementary literature examines the spillover effects of nearby open spaces and amenities on land prices (e.g., Greenstone and Gallagher, 2008; Rossi-Hansberg et al., 2010; Gamper-Rabindran and Timmins, 2013). These studies collectively provide insights into the impact of land use regulations on

[1] Relatedly, Bloom et al. (2019) find that there is considerable variation in management practices across plants within the same firm and business environment, especially RTW rules, can increase structured management practices around pay, promotion, and dismissals.

[2] Numerous other studies have evaluated the impacts of various local, state, and federal economic development policies or place-based programs. These include studies by Kline and Moretti (2014b); Busso et al. (2013); Gobillon and Magnac (2016); Suárez Serrato and Zidar (2016); Slattery and Zidar (2020); Chyn and Katz (2021), among others. These works collectively highlight the significant potential of public policies to spur productivity and enhance welfare in severely underdeveloped regions through agglomeration forces, particularly productive spillovers between workers and firms, as emphasized by Ellison and Glaeser (1997), Ellison et al. (2010) and Greenstone et al. (2010).

land prices and, in turn, their broader economic implications. Our findings reveal that land development has a multifaceted influence, manifesting in both benefits and costs, with these effects materializing at different time frames. This temporal variation presents a challenge when attempting to assess the net impact at any given point.

Our analysis is also related to the literature that studies interactions between climate change, environment and aggregated economy (e.g., Nordhaus, 1977; Golosov et al., 2014; Nordhaus, 2019). Our study specifically focuses on the impact of land development on future flood damage, one of the most acute climate-related events. Our findings reveal a significant and positive association between increased land development and higher flood claims in the long term, with this effect somewhat attenuating in the short and intermediate terms. This effect stands apart from the findings in existing literature, which primarily emphasize the role of protective ecosystem services in mitigating flood damages.

2. Background

2.1. Land Use in Economic Development

Economic development policies often incorporate a range of land-use strategies and initiatives designed to stimulate private economic growth and foster job creation. One prominent example is the establishment of enterprise zones, which receive special privileges such as tax breaks, regulatory exemptions, or public assistance to encourage business operations within the designated area. These zones are commonly utilized to revitalize neighborhoods that have experienced a decline in essential businesses or quality housing (e.g., Jersey City, NJ in the 1980s), or to aid areas struggling to recover from natural disasters (e.g., New Orleans after Hurricane Katrina). Additional land-use policies include Tax Increment Financing (TIF) Districts, created to harness the augmented property tax revenue within a specified region for infrastructure improvements and economic development projects, as well as Foreign Trade Zones designed to facilitate international trade, promote manufacturing, and generate employment by offering tariff and tax incentives to businesses engaged in import/export activities.

RTW laws and land use policies represent two distinct yet interrelated facets of local economic development policies. RTW laws primarily impact labor relations by forbidding employers and labor unions from mandating union membership or the payment of union dues or fees as a condition of employment. Consequently, RTW states often appear more attractive to businesses when making location decisions. Furthermore, RTW laws frequently coincide with other pro-business packages, including the designation of specific zones or areas for economic development, such as business parks or enterprise zones. These designated areas are strategically established to draw businesses, stimulate investment, and foster job creation.

2.2. Land Use and Flood Risk

The relationship between land development, urbanization and flood risk is complex and interconnected. As urban areas expand and land is developed for various purposes, such as residential, commercial, or industrial use, the natural landscape often undergoes significant alterations. These alterations, including the construction of buildings, roads, and drainage systems, can disrupt the natural flow of water and increase the risk of flooding (e.g., Konrad et al., 2003; Khan, 2005). Impermeable surfaces like pavement and concrete prevent rainwater from being absorbed into the ground, leading to increased runoff into rivers and streams during heavy rainfall (e.g., Yan and Edwards, 2013). Poorly planned land development can exacerbate flood risks by channeling water into vulnerable areas and reducing natural flood buffers such as wetlands and floodplains (Taylor and Druckenmiller, 2022). In addition, urbanization has the potential to directly influence the magnitude of extreme precipitation. For example, Zhang et al. (2018) find that the probability of such extreme flood events across the studied basins increased on average by about 21 times in the period 25–30 August 2017 because of urbanization.

3. Data

3.1. Sources

LCMAP Data We utilize the LCMAP data from the USGS's Earth Resources Observation and Science Center. This dataset provides land cover information at the pixel level and annually for the years 2001-2019. Covers eight classes, including developed, cropland, grass/shrub (grassland), tree cover (treeland), water, wetland, ice/snow, and barren. In Figure 1, it is evident that the United States possesses a substantial amount of undeveloped land that could potentially be available in the future. In 2001 (Panel A), only 4.9% of the land was developed. The majority of the landscape consisted of grassland (29.4%), cropland (29.2%) and treeland (28.1%), representing a total of 87%. Over a 15-year period, the proportion of developed land slightly increased to 5.1% in 2016, marking a mere 0.2 percentage point increase. This expansion in developed land, as well as in cropland and grassland, occurred at the expense of treeland and, to a lesser extent, wetland (Panel B). While the increase in developed land is smaller in terms of area compared to cropland, the relative growth of developed land is 4.09%, exceeding the growth of 0.89% observed for cropland.

NFIP Data The National Flood Insurance Program (NFIP) was established in 1968 by the federal government with the aim of providing affordable flood insurance to homeowners. As of 2017, the program boasted more than 5 million policies in force nationwide, covering nearly $1.27 trillion. However, slightly over 80% of the policies within the program are charged what FEMA terms as full-risk rates, while the remaining 20% receive substantial subsidies at less than half of the full rates. The NFIP carries various cross-subsidies, both explicit and implicit, sparking ongoing debates in legislative discussions regarding affordability, rise-based pricing, and fiscal sustainability - a clear tradeoff (Kousky, 2018).[3] The NFIP faces a substantial debt burden, exceeding $20 billion as of early 2018, with no foreseeable means of

[3] In 2021, FEMA introduced a new pricing approach known as Risk Rating 2.0 (RR2.0), which enhances accuracy by assessing risk at the property level. NFIP premiums are now determined through simulation-based catastrophe models combined with claim history, resulting in insurance pricing that more precisely mirrors a property's anticipated losses.

repayment. This financial scenario has significant implications for land development. First, due to discounts and cross-subsidies, the financial interests of land developers and landowners are not aligned with those of the public, who ultimately bear the costs of flood risk. Second, if land development contributes to increased flood damage, any economic development policies based on an unsound economic analysis could essentially represent a transfer from taxpayers to local governments and developers.

Data from the NFIP by the FEMA are used to construct our flood damages (FD) measure. Specifically, we rely on the NFIP Redacted Claims - v1 dataset, which includes over 2,000,000 claims transactions, providing location and claim amount information. To calculate the total claim amount, we sum the amounts paid for building and contents claims, aggregating this data to the zip code-by-year level. To address the infrequency of flood events, we create a 3-year moving average of the NFIP total paid amount for the dependent variable. From 2001 to 2016, the total NFIP claims increased from \$1.03 billion to \$5.03 billion or by 415%.

Census Data To mitigate concerns about endogeneity resulting from NFIP insurance take-up rates, we control for various demographic and socioeconomic factors. These data are sourced from the 2000 Decennial Census and the 5-year American Community Survey (ACS) at the Zip Code Tabulation Area (ZCTA) level. We construct measures for the population, the number of housing units, the median household income, and the median housing value. For distributional effect analysis, we use household demographics to create measures for minority vs. majority populations and different income classes.

3.2. Sample

Our main sample is based on all zip codes in the LCMAP data. We also used a supplemental RTW subsample that consists of zip codes located in states that had enacted RTW laws and their neighboring states.

National Sample The national sample employs all land cover data spanning from 2001 to 2019, creating a 19-year zip code panel. Subsequently, we computed the differences, denoted as , for all variables considered in the analysis. This process generates multiple stacked LD samples. For instance, the 15-year LD sample comprises four stacks: 2001-2016, 2002-2017, 2003-2018, and 2004-2019, constituting the primary dataset for our analysis. As the LD horizon shortens, the number of stacks and observations within the LD sample increases. Consequently, we can conduct analyses across the entire horizon to explore the short and long-term effects of land development on flood risk while controlling for other factors.

The LCMAP dataset initially contains 31,539 zip code-level observations. After merging with control variables and to maintain strong balance across stacks, the final sample is reduced to 26,403 zip codes. Table 1 Panel A provides summary statistics for all variables in our main analysis based on the 15-year LD sample. Over the course of 15 years, developed land has exhibited an average increase of 2.4% (logarithmic growth weighted by population), while FD have shown an increase of 38.4%.

These mean values encompass considerable spatial variability in both series. We begin by visualizing the two primary variables, the change in FD ($\Delta Log(FD)$) and the change in developed land ($\Delta Log(Dev)$), from 2001 to 2016, in the two panels of Figure 2. Panel A shows that changes in developed land exhibit a smaller regional dispersion, with 62% of zip codes reporting increases in developed land from 2001 to 2016. Five states have seen the most rapid growth in Dev, driven by their ample land availability, including Wyoming, Montana, Delaware, Utah, and Texas. Texas stands out with the largest share, contributing 21% to the total of 1.2 million hectares of land developed in the past 15 years. Following closely is Florida, which ranks second with 5.6%. On the opposite end of the spectrum, Alabama, Mississippi, and West Virginia have experienced the slowest rates of Dev, with either negative or near-zero growth.

Panel B illustrates that changes in FD are geographically concentrated in the South Atlantic and two Central South regions. From 2001 to 2016, six states have witnessed increases in flood damage that exceeded ten times, including South Carolina (57 times), Idaho (21 times), Florida (15 times), Wyoming (14 times), Georgia (12 times), and Montana (11 times). Analyzing changes in FD dollars during this period, Texas, Louisiana, and Florida emerge as the top-ranking states, contributing to 62%, 20%, and 10% of the total $4.3 billion in flood damage, respectively. In contrast, fourteen states have experienced declines in flood damage during the same 15-year period, with notable examples being North Dakota (-99%), Delaware (-89%), Maryland (-87%), and Minnesota (-80%). Confirming the substantial spatial variation depicted in our primary variables in Figure 2, $\Delta Log(FD)$ exhibits a standard deviation of 5.281 (13.8 times the mean), whereas $\Delta Log(Dev)$ shows a standard deviation of only 0.087 (3.6 times the mean). Nevertheless, Panel C of Figure 2 reveals that the two series demonstrate a significant positive correlation at 55%.

RTW Sample Panel B of Table 1 presents summary statistics for the variables in the supplementary RTW sample. During our sample period (2001-2019), there are five states that passed and enacted the RTW Law in the state, including Indiana (2012), Michigan (2013), Wisconsin (2015), West Virginia (2016), and Kentucky (2017). These states constitute the treatment group, while their neighboring non-RTW states serve as the respective control groups.[4] This approach results in some states being repeated as control states for different treatment states, albeit with distinct treatment dates. The final sample encompasses 8,938 unique zip codes, divided into 14,498 zip code observations distributed across five state-event groups.

The treatment group comprises all zip codes located in states that enacted RTW laws during the sample period, making treatment a time-varying event. Our analysis reveals that in areas where the zip code's initial development level is above the 75^{th} percentile, there tends to be a decreasing relationship between the change in developed land and the initial development level. Conversely, in areas falling below this "tipping point" threshold in terms of initial development, the relationship tends to increase. Consequently, we further categorize these zip

[4]The control groups are composed of Illinois and Ohio for Indiana, Ohio for Michigan, Minnesota and Illinois for Wisconsin, and Ohio, Pennsylvania, and Maryland for West Virginia, as well as Ohio, Illinois, and Missouri for Kentucky.

codes into high and low developed areas based on their initial development status. A zip code is classified as "low developed" if its proportion of developed land in 2000 falls below the 75^{th} percentile of the state; otherwise, it is considered "high developed." Thus, we categorize the primary variables within the RTW sample into six distinct subgroups, depending on whether the zip code belongs to one of the two treatment groups or control state, and whether the observation is from a pre- or post-RTW year.

Panel B indicates that low-developed zip codes in treatment states have seen a change in developed land from 836 to 895, or 7. 2%, while high-developed areas in the treatment states have experienced a change from 1,674 to 1,736, or 3. 6%, similar to the control group that has seen a change from 991 to 1,023, or 3. 2%. Meanwhile, the relative change in FD from the pre-RTW to the post-RTW years for these three groups is -35%, -51%, and -41%, respectively. On the surface, there is a high correlation between changes in developed land across the RTW groups, while there is very little correlation between changes in FD across these groups.

4. Main Analysis

4.1. Empirical Strategy

Our empirical work aims to estimate the causal effect of land development on climate risk using flood insurance claims. However, we face the challenge that changes in land development are correlated with many confounding factors that also affect flood risk. For example, Florida, a coastal state traditionally associated with high flood insurance claims, has experienced faster land development than many other states in the past decades. However, we cannot interpret the positive correlation between land development and flood risk as land development directly causing a higher flood risk, as Florida's geographic characteristics, such as its low elevation, porous limestone underlying its land, extensive coastline, and susceptibility to hurricanes, also contribute significantly to its vulnerability to global sea level rise and increased flood risk. Furthermore, various local covariates, which may or may not be correlated with land development, can explain changes in flood risk.

To address these empirical challenges, we employ several strategies to identify the variation in local land development that affects flood risk. The first strategy involves stacked LD to capture the effect of land development on flood risk over different horizons since the land is developed ($t_0 + n, n = 1, 2, ..., 15$). This approach controls for state (s) by cohort year (t_0) fixed effects and includes many zip code-level covariates. We also control for changes in non-developed land, which may contribute to flood risk as well. In addition, we distinguish between various sources of developed land, which depend on local natural endowments, and initial development conditions, which affect the economic incentives of land development. The second strategy entails a panel data analysis that leverages state differences in the adoption of RTW laws—one of the signature pro-business laws—based on matched data and employs a 2SLS approach.

We will elaborate on the rationale for using state differences in the adoption of RTW laws as a valid instrumental variable (IV). For the estimation of IV to be valid, the instrument Z must meet two criteria: (1) it must affect the endogenous treatment variable X, and (2) it must satisfy the exclusion restriction, which states that Z should not have a direct effect on the outcome variable Y, or if it does, the effect should only be through X (e.g., Newhouse and McClellan, 1998; Angrist and Krueger, 2001). Violating the exclusion restriction can lead to biased IV estimates.

We choose two instruments: the indicator of RTW states and the indicator of initial development state for a given zip code (high vs. low developed). The former is a time-varying variable, defined as 1 if a zip code is located in a state that has enacted RTW laws in a given year, and 0 otherwise. The latter is a cross-sectional variable, defined as low developed if the share of developed land in a given zip code in the year 2000, one year before our sample begins, is below a certain threshold, and high developed if it is above the threshold. The rationale for using RTW states as an instrument for land development is that land is a critical resource and a key element in any pro-business policy aimed at attracting businesses. However, it should not directly affect the local flood risk because it is a statewide policy. The second variable captures historical land availability—high–developed places have less land available (and possibly less economic incentives) for further development—which should not be correlated with changes in future flood risk. Nonetheless, it can influence the cost and price elasticity of land development, akin to other land supply measures used in the literature (e.g., Saiz, 2010; Gyourko et al., 2021; Lutz and Sand, 2022). Since our instrument is an interaction between the initial development conditions and an RTW indicator, the identification requirement here is even weaker. We need that, else equal, the marginal effect of the initial developed fraction on the subsequent land development varies across RTW and non-RTW areas; but its marginal effect on flood risk, if any, does not change with the pro-business policies.

Figure 3 presents the first-stage results, which involve regressing the outcome and treatment variables on the interaction of the two indicator variables over a 19-year window (10 years before through 8 years after the enactment of the RTW laws) in the following specification:

$$\begin{aligned} Y_{i,t} \\ &= \sum_{t=-10}^{8} \left(\beta_{1,t} \cdot I(\text{Post RTW}_{i,t} \times \text{Low Dev}_{i,0}) + \beta_{2,t} \cdot I(\text{Post RTW}_{i,t} \times \text{High Dev}_{i,0})\right) \\ &\quad + \theta \cdot Log(R_{i,t}) + \mu_{g\times t} + \mu_s + \varepsilon_{i,t} \end{aligned} \tag{1}$$

where $Y_{i,t}$ represents the logarithm of either FD or land development (Dev), I(Post RTW$_{i,t}$ $\times$ Low Dev$_{i,0}$) is an indicator for zip codes located in RTW states with low initial land development before our sample period (low developed areas), and I(Post RTW$_{i,t}$ $\times$ High Dev$_{i,0}$) is an indicator for zip codes located in RTW states with high initial land development (high developed areas). The reference group comprises zip codes located in non-RTW states adjacent to the RTW states. Each RTW state and its adjacent non-RTW states are matched to form a state pair named after the RTW state (g, e.g., WV group includes WV and its OH, PA, and MD neighbors), and some states are thus included in different state pairs. The model includes state pair by relative year fixed effects ($\mu_{g\times t}$), state fixed effects (μ_s), and zip code covariates $R_{i,t}$, such as housing units, housing value, income, population, and NFIP risk ratings.

Panel A of Figure 3 shows no significant differences in FD between the two treatment groups and the control group, both before and after the event date. The confidence intervals for the estimates in high and low developed zip codes also exhibit overlap, suggesting that there are also no significant differences between the two treatment groups in the pre- and post-RTW years. These results indicate that the two instruments meet the exclusion restriction criteria.

Panel B shows that low-developed zip codes located in RTW states experience significantly higher growth in land development after the enactment of the law compared to control states in the same region. The growth in land development becomes significantly higher starting from the third year after the state adopts the RTW laws, relative to its own level before the law and to its neighboring states. By the eighth year, the differential growth rate between developing zip codes in RTW states had reached 1. 5%, corresponding to approximately 13 hectares of land development before the law. In contrast, there are no pretrends in the years before the law was passed, validating the parallel trend assumption. High-developed zip codes in RTW states also exhibit significantly higher growth in land development following the law, although to a much smaller extent (0.3%). These results suggest that while RTW laws, along with other pro-business policies, may provide a necessary condition for land development, the initial development state and supply constraints play crucial roles in the effect on land development.

4.2. OLS Results

We begin by employing a stacked LD approach to estimate the short- and long-term relationship between changes in land development and FD using the following specification:

$$\Delta Log(FD_{i,t_0+n}) = \beta \cdot \Delta Log(Dev_{i,t_0+n}) + \theta \cdot \Delta Log(X_{i,t_0+n}) + \mu_{s,t_0} + \varepsilon_{i,t_0+n}, \tag{2}$$

Here, $\Delta Log(FD_{i,t_0+n})$ and $\Delta Log(Dev_{i,t_0+n})$ represent the change in the natural logarithm of flood damage and developed land, respectively, in a zip code i located in state s and cohort year t_0 over an n-year ($n = 1, 2, ..., 15$) horizon. The vector of covariates X_{i,t_0+n} includes the logarithmic changes in zip code-level number of housing units, average housing value, median household income, the population, and the local participation in the NFIP's CRS. We also control for state by cohort year (t_0) fixed effects to isolate any state $\times$ year-level factors that could affect $\Delta Log(Dev)$. To account for potential correlation in the error term over space, standard errors are clustered at the county level.

The results are displayed in Table 2. Panel A reports the analysis using the 1-year, 5-year, 10-year and 15-year differences, and Figure 4 plots the coefficient on $\Delta Log(Dev_{i,t_0+n})$ across various horizons for $n = 1, 2, ..., 15$, according to Equation (4). These findings reveal a progressive increase in the relationship between $\Delta Log(Dev)$ and $\Delta Log(FD)$ over longer horizons. Initially, the relationship is negative, but lacks statistical significance on a 1- to 3-year horizon. Subsequently, it turns positive, but remains statistically insignificant over a 4- to 8-year horizon. The relationship achieves statistical significance after 8 years, with both the magnitude of the estimates and the level of significance demonstrating a continual rise over time. Over the 15-year horizon, a one-percent increase in developed land is associated with a 2.08-percent increase in flood damage, a significant result at the 1% level.

4.3. 2SLS Results

Alternatively, we utilize a panel fixed effects model that relies on within-zip code year-over-year variation in land development to estimate its short-term effect on flood risk:

$$Log(FD_{i,t}) = \beta \cdot Log(Dev_{i,t}) + \theta \cdot Log(X_{i,t}) + \mu_i + \mu_{s,t} + \varepsilon_{i,t}, \tag{3}$$

In this equation, $Log(FD_{i,t})$ and $Log(Dev_{i,t})$ represent the logarithm of flood damage and land development in zip code i located in state s and year t, respectively. The vector of covariates $Log(X_{i,t})$ includes the logarithm of zip code-level numbers of housing units, the average housing value, the median household income, the population, and the local participation in the NFIP's CRS. To control for unobserved zip code-level and state $\times$ year-level confounding factors, we include state $\times$ year ($\mu_{s,t}$) fixed effects. Once again, we cluster standard errors at the county level to account for possible correlation in the error term over space and time.

Panel B of Table 2 summarizes the analysis based on panel data regressions as specified in Equation (3). Columns (1) and (2) present the OLS regressions for the full sample and the RTW sub-sample, respectively. Column (2) further includes state pair by relative year fixed effects to account for the fact that some control states may appear in another state pair for different events. The OLS regression results in estimates of 0.545 and 0.506, indicating that a one-percent annual increase in land development is associated with a 0.545 and 0.506-percent increase in flood damage within the same zip code, both estimates are statistically significant at the 1% level. Taking into account the structure of the panel, these estimates capture the short-term relationship between the contemporaneous change in land development and flood damage and, as a result, the estimate is comparatively smaller in magnitude compared to the LD analysis in panel A.

Columns (3) and (4) present the first and second stage results in the 2SLS regression, where $Log(Dev_{i,t})$ is instrumented with $I(\text{Post RTW}_{i,t} \times \text{Low Dev}_{i,0})$ and $I(\text{Post RTW}_{i,t} \times \text{High Dev}_{i,0})$ in Equation (1). The results in the first stage indicate that, compared to the control group comprising zip codes in non-RTW states, low-developed zip codes in RTW states experience a significantly higher increase in land development of 0.065 percent, significant at the level 1%. In contrast, high-developed zip codes in the treatment states exhibit a significantly lower increase in land development by -0.225 percent, also significant at the 1% level. Both the underidentification test and the weak instruments test strongly reject the null hypothesis, signifying that the instruments possess sufficient explanatory power to predict the X variable. This suggests that our two instrument variables effectively predict changes in land development.

Consistently, the 2SLS regression yields an estimate of 1.640 percent, significant at the 10% level. This estimate is larger in magnitude than the OLS estimate in column (2), likely attributed to $\widehat{Log(Dev_{i,t})}$ now capturing the change in developed land over 8 years after the enactment of RTW laws for zip codes in the treatment group, rather than solely representing the contemporaneous change in land development in column (2). However, this result suggests that an increase in land development has a causally positive effect on the change in flood risk.

5. Heterogeneity Analysis

The baseline results in Table 2 consistently show a positive and significant relationship between changes in land development and flood risk, while controlling for state fixed effects and local covariates, as well as employing exogenous instrument variables that provide quasi-randomization of the treatment. However, concerns remain about unobserved confounding factors that might explain the results.

5.1. Different Land Uses

5.1.1. Change in Non-Developed Land

One confounding factor that warrants consideration is the change in non-developed land, rather than developed land, which could drive changes in flood risk. This is due to the zero-sum nature of changes in developed and non-developed land. Consequently, an increase in flood risk may not be primarily due to population growth and industrial activities, but rather to the reduction of land that could have been conserved as natural habitat and climate mitigants, as demonstrated by the case of wetlands studied by Taylor and Druckenmiller (2022). It is essential to recognize that other land uses might similarly contribute to the natural ecosystem, impacting flood risk exposure.

To test this hypothesis, we include the change in non-developed land in Equation (1) using the 15-year LD specification. The results are presented in Table 3.[5] Column (1) is our baseline specification in Table 2. Column (2) shows that the coefficient on $\Delta Log(Dev_{i,t_0+15})$ remains positive at 2.014 and statistically significant at the 1% level, slightly lower than the baseline estimate of 2.080. In contrast, the coefficient on $\Delta Log(NonDev_{i,t_0+15})$ is negative but not statistically significant, suggesting that the positive effect of land development observed in the baseline regressions is not explained by changes in non-developed land as a whole.

In columns (3) and (4), we break down $\Delta Log(NonDev_{i,t_0+15})$ into changes in wetland

[5] Since both our sample and empirical specification differ from those in Taylor and Druckenmiller (2022), we adopt their small sample so that we can replicate similar results using our specification. We also report the estimated results using their specifications in Panel A of Table ??.

and other non-developed land uses, such as cropland, treeland, grassland, water, and barren. The findings indicate that none of the coefficients for the changes in non-developed land are statistically significant. Collectively, these findings suggest that the increase in flood damage is associated with the expansion of developed land rather than changes in other land uses.

5.1.2. Sources of Developed Land

Panel B of Table 3 provides insight into the sources of developed land from 2001 to 2019 in the US. In particular, treeland emerges as the predominant contributor, representing a significant portion ranging from 47% to 61% of the total developed land on various horizons. Follow closely, cropland consistently represents the second most important source, contributing around 40% to the total developed land. Barren land trails as a distant third, making up 6.5% to 10% of all developed land. The wetland accounts for 2.4% to 2.6% of the developed land. In contrast, land developed from grassland or water has either declined or shown negligible growth in recent decades.

In column (5) of Panel A, we replace the change in developed land with the changes in developed land from different sources, such as those from wetland, cropland, grassland, treeland, water, and barren land. First, when estimating the effect of the change in developed land from wetland, we observe a substantial effect of 3.361 percent, although it is not statistically significant. Second, several other development sources are also positively associated with changes in flood damage. In particular, the coefficients for the change in developed land from cropland and treeland are 4.133 and 3.076 percent, respectively, and both are statistically significant at the 1% or 5% level. On the contrary, the coefficients for changes in other development sources are not statistically significant. These results underscore that the significant and positive impact of land development on flood damage in the baseline analysis is mainly driven by changes in land developed from croplands and treelands. Third, the estimates of changes in developed land from cropland and treeland become even more pronounced when accounting for changes in other land uses in columns (6) and (7). In column (7) with full controls, the coefficients of change in developed land from cropland or treeland are 4.309 and

3.578 percent, respectively, and both are significant at the 1% level.

Figure 5 plots the estimated effect of the change in developed land from cropland and treeland, the two most significant development sources, over the entire 1 to 15-year horizon. Compared to the general effect of the change in developed land from all sources shown in Figure 4, the estimated effects of these two development sources are not only much greater in magnitude, but they also impact flood damage much earlier. Development sourced from treeland exhibits a positive and significant effect on flood damage as early as the 3^{rd} year, increasing in magnitude and significance over time, in contrast to the 9^{th} year for all sources. The same pattern applies to the development sourced from cropland as well, which is with a larger magnitude compared with all sources and shows a significant effect as early as the 5^{th} year.

To assess the relative importance of changes in non-developed land and various development sources in explaining flood damage, we calculate the proportion of the total XBETA explained by each variable. In column (8) in Table 3 Panel A, it is revealed that, among all land uses, the change in developed land from land covers with trees has the most substantial effect, registering an XBETA of 5.9. The change in developed land from cropland is closely followed, with an XBETA of 5.6. On the contrary, the combined impact of the other four sources amounts to an XBETA of 0.7. In comparison to developed land, the contribution of changes in non-developed land is relatively small.

5.2. Initial Development Conditions

The previous finding that the effect of the RTW laws on land development depends on the initial development state in Section 4.3 suggests that the initial conditions of a local area may also affect flood risk indirectly by affecting the economic incentives of land development. In this section, we examine the variations in this relationship between zip codes with different initial conditions. Here, the term "initial conditions" encompasses both the initial development state and the initial development density. The former captures the varying availability of non-developed land (and associated economic incentives) for future development, while the latter reflects the intensity in the use of existing developed land, although both may be correlated with the natural endowments of the area.

5.2.1. Land Development Patterns Across Difference Initial Conditions

As such, we categorize all zip codes into three groups based on their initial conditions: zip codes in the low initial conditions group characterized by a low share of developed land and low development density; those in the high initial conditions group characterized by a high share of developed land and high development density; and the remaining zip codes are classified as medium initial conditions. Panel A of Table 4 presents the land characteristics based on these initial conditions. According to defined criteria, zip codes in the low, medium, and high initial conditions group had 3.45%, 17.3%, and 42.9% of their land developed in 2001, respectively. In addition, the three groups show an ascending trend in land density, measured by the number of people living in each square mile. Concerning flood risk, while the average flood damage at the beginning of the sample period is higher for zip codes with higher initial conditions, there is no apparent difference in the percentage change in flood damage between the high and low initial conditions groups. Furthermore, during the 15-year period, the flood damage for the low and high initial conditions groups decreases by a smaller percentage than that for the medium initial conditions group.

The initial conditions are shown to be highly predictive of future land development. During the 15-year span, developed land has experienced an average increase of 81.8 hectares or 4.11% for zip codes in the high initial conditions group, 16.81 hectares or 1.97% for the medium group, and merely 2.56 hectares or 0.82% for the low group per zip code. In total, approximately 80% of the land developed during the 15-year period is in zip codes with high initial conditions. There is also considerable variation in land development sources between groups. For areas with high initial conditions, cropland and treeland are the two most important sources of development, each accounting for about 41% of the developed land. For areas with medium initial conditions, treeland becomes a more crucial source than cropland, accounting for twice the share of developed land. For areas with low initial conditions, treeland is even more important, accounting for more than 200% of developed land. In areas with low and medium conditions, the development on treeland and cropland is offset by the conversion of developed land to grassland, suggesting that the cost of land development in these areas is not high.

5.2.2. Estimates Across Difference Initial Conditions

Taking advantage of these distinct land development patterns, we explore differences in the relationship between $\Delta Log(Dev)$ and $\Delta Log(FD)$ by estimating separate coefficients on the change in developed land between the initial condition categories. The results are presented in Panel B of Table 4. It reveals that there is no significant relationship between $\Delta Log(Dev)$ and $\Delta Log(FD)$ in the short and long term for zip codes with low initial conditions. In contrast, the significant effect observed in the baseline analysis is mainly due to zip codes with high initial conditions, where the majority of land development has occurred, and to a lesser degree from those with medium initial conditions. Over the 15-year horizon, a one-percent increase in developed land is associated with a 1.307-percent increase (significant at the 5% level) in flood damage for the medium conditions group and a 3.330-percent increase (significant at the 1% level) for the high conditions group, respectively. Over the 10-year horizon, a one-percent increase in developed land is associated with a 1.289-percent increase (significant at the 5% level) in flood damage for the medium conditions group and a 2.054-percent increase (significant at the 5% level) for the high conditions group, respectively. In other periods with shorter horizons, the estimates for either group are not statistically significant.

Figure 6 plots the estimated effect across the horizon for the three initial condition groups. Compared to the general effect in Figure 4, the effect is much greater and more significant for higher initial conditions. It is not statistically significant and close to zero in most years for zip codes in the initial low group. In contrast, the effect is positive and significant after the 9^{th} year for the medium initial conditions group and even greater in magnitude and earlier in timing for the high initial conditions group.

Taking stock, we observe considerable variation in the estimated effect of developed land across different development sources, as well as the initial development conditions of the local area. Lands developed from croplands and treelands are associated with the most significant impact on flood damage, while other development sources are not as damaging to flood risk. Similarly, the initial development conditions, which prove to predict development activities in the following years, also influence the estimated effect, with the majority of the impact stemming from development activities in areas with high initial conditions.

6. Distributional Analysis

In this section, we link the change in developed land with the demographic characteristics of residents who ultimately bear the cost of flood damage.

6.1. Demographic Characteristics of Developed Land

6.1.1. Initial Demographic Characteristics

We identify the demographic characteristics of the land developed in the past 15 years using Equation (3), but replace $\Delta(FD)$ with demographic variables at the zip code level as the dependent variable. Table 5 Panel A presents the analysis of various initial demographic characteristics in the year before our land cover sample begins. We adopt five specifications to analyze the relationship between the demographic characteristics of the area and (1) the change in developed land, (2) initial shares of different land covers that measure the natural endowment, (3) the change in developed land from different sources, (4) the initial conditions categories and (5) the interaction of the change in developed land and the initial conditions categories.

The regressions in specification (1) reveal that, first, areas with a high proportion of minorities, a high proportion of low-income population, lower housing values, lower income, and a lower population in 2001 have experienced slower land development in the past 15 years. Furthermore, the results of specification (5) show that the slower land development experienced by areas with more minorities and a higher low-income population is significant in all three initial condition categories. Second, areas with a higher proportion of minorities and a larger population tend to have less land available for future development, but relatively more cropland than treeland available, as shown in Specification (2). In contrast, areas with a higher proportion of low-income population tend to have more land available, particularly treeland and other land covers except cropland. Third, despite having different natural endowments, Specification (3) indicates that areas with a higher proportion of minorities, low-income population, and a larger overall population have experienced a relatively higher amount of land developed from cropland compared to treeland. Fourth, Specification (4) shows that areas with high initial conditions are associated with a higher proportion of minorities, more housing units, a larger population and higher housing value, but a lower proportion of low-income population and a higher income.

6.1.2. Changes in Demographic Characteristics

These demographic characteristics are highly likely to change as more land is developed. Focusing on the net change from the initial level to account for people moving in and out, Table 5 Panel B presents the analysis of changes in various initial demographic characteristics. The results in Specification (1) indicate that, first, areas with more land development are associated with a significant increase in the proportion and minorities, the proportion of low-income population, housing units, housing value, income and population over the 15-year span. Furthermore, the increase in minorities and the low-income population associated with land development is more pronounced in areas with high initial conditions than elsewhere (see Specification (5)). Second, Specification (2) shows that areas with more non-developed land available have seen a significant decrease in the proportion of minorities and low-income population, indicating that these populations have relocated from these areas to more developed ones. Specifically, areas with a high share of treeland have seen a greater decrease in minorities, a smaller increase in population and median income compared to those with a significant amount of cropland, while areas with a high share of cropland have seen a greater reduction in low-income population compared to those with a significant amount of treeland.

Third, areas with more land developed from treeland are associated with a significant increase in the proportion of minorities, housing units, population, housing value, and median income, exceeding the impact observed in areas with more land developed from cropland and other sources. Areas with more land developed from other sources, such as wetland, are associated with a significant increase in the proportion of low-income population, surpassing the impact observed in areas with more land developed from cropland and treeland, as seen in Specification (3). Fourth, areas with high initial conditions have experienced a significant increase in the proportion of minorities, the proportion of low-income population and the overall population, but a significant decrease in income, as seen in Specification (4).

6.1.3. Summary

In summary, the sources of developed land are largely determined by the natural endowments of an area, specifically, the availability of different land covers. Consequently, the development of these land covers inevitably affects the people who live in them. Our analysis indicates that the most important land developed in the past two decades—those from cropland and treeland—happen to be in areas where there is a low proportion of minority and a high proportion of low-income population and where housing and population are sparse. As more land is developed, these areas see a significant decrease in the share of the minority and low-income population. There are also demographic differences between areas with more cropland or more treeland available. Initially, areas with more treeland available have even fewer minorities and a higher low-income population, experiencing a greater decline in minorities and a lesser decline in the low-income population as more land is developed, compared to areas with more cropland. Overall, land development, particularly from treeland, is associated with a significant increase in income, population, and housing units, and a significant decrease in the proportion of minorities, the proportion of low-income population, and the value of housing.

The distribution of initial development conditions is related to, but different from, natural endowments in that areas with high initial conditions develop land equally from cropland and treeland, while areas with medium and low conditions predominantly develop from treeland. Relative to zip codes with low initial conditions, those with high initial conditions have a higher proportion of minority but a lower proportion of low-income population, and have experienced a significant increase in minorities and low-income population as more land is developed. Thus, land development is expected to have different distributional effects on different demographic groups due to differences in initial demographic characteristics as well as changes in them.

6.2. Methodology

We quantify the distributional effects across demographic groups by estimating the change in flood damage on developed land interacted with a grid of both land development sources and initial conditions, while controlling for state-by-cohort year fixed effects and zip code covariates. We focus on the 15-year LD as this represents the longest horizon in our data for which we can estimate the long-term effect. The estimation is specified below:

$$\begin{aligned}\Delta Log(FD_{i,t_0+15}) = \sum_j \sum_k \beta_{j,k} \cdot \Delta Log(Dev_{i,t_0+15}) \cdot 1(Source_j) \\ \cdot 1(Initial_k) + \mu_{s,t_0} + \theta \cdot \Delta Log(X_{i,t_0+n}) + \varepsilon_{i,t_0+15},\end{aligned} \tag{4}$$

where $1(Source_j)$ and $1(Initial_k)$ represent categorical variables defined based on sources of land development (i.e., from cropland, treeland, and others) and initial conditions (e.g., low, medium and high).

The coefficients for these 9 grids are presented in Table 6. Consistent with the results in Section 5, the change in developed land from croplands and treelands is associated with a significant increase in flood damage, while there is no significant link between developed land from other sources and flood damage. The grid analysis further shows that the significant association between the change in developed land from cropland and flood damage is valid in the medium and high initial condition categories, but that between land developed from treeland and flood damage is only positive and significant for the high initial conditions group.

During the 15-year horizon, a one-percent increase in developed land from cropland and treeland is associated with a 5.892 and 6.170-percent increase (both significant at the 1% level) in flood damage for the initial high group. The effect decreases to only a 2.595-percent increase (significant at the 5% level) in flood damage per one-percent increase in developed land from cropland for the medium initial conditions group and an insignificant increase for the low conditions group, respectively. The effect of the increase in developed land from treeland on flood damage is positive, but not significant for the low and medium initial conditions.

6.3. Estimated Cost

Using the predicted XBETA from the coefficients for these 9 grids as the basis, we calculate the long-run flood cost per hectare of developed land. This is done by transforming the XBETA, representing the percent change in flood damage, to total dollars in each zip code by multiplying the estimated percent change and FD_0 at the zip code level. At the national level, we estimate that the long-term flood cost of developed land through the flood damage channel to be \$627 million at the 15^{th} year, which accounts for 14.7% of \$4.27 billion in FD increased from 2001 to 2016. Based on the distribution of the estimated cost over different years depicted in Figure 4, we apply a factor of 4.1 to obtain the lifetime cost. This amounts to \$2.59 billion in total and \$2,164 per hectare cumulated over the 15-year span, representing the social cost of land development not paid by the private market. Based on the average land value of \$9,979 per hectare in 2016, the social cost represents 22% of the market value.

6.3.1. Spatial Distribution

However, there is significant spatial variation in the estimated flood cost, as shown in Figure 7, primarily due to differences in the natural endowment and initial development conditions. Panel A shows the total estimated lifetime cost by US county. If all counties were to equally share the total cost of \$2.59 billion, that would amount to \$839,000 per county. However, the top 1% of counties have to bear 95% of the cost, with the highest burden falling on Harris County, TX, in the Houston metro area, which has to bear $1.53 billion. Most of these hotspots are located in Texas, Louisiana, Virginia, Maryland and Florida. These hotspots also have the highest cost per hectare of developed land in Panel B. Controlling for total flood damage in Panel C, hotspots become more spatially clustered in Texas, Louisiana, Florida and Maine, all coastal states. When we take into account the land value in Panel D, new hotspots emerge in North Dakota and Washington besides Texas, Louisiana and Florida.

6.3.2. Distributional Effects Across Demographic Groups

Flood risk is not borne equally by all demographic groups. We utilize zip-code-level data from the 2016 ACS to examine the demographic characteristics, focusing on race and poverty, of present-day flood risk due to land development across the United States. Our analysis centers on the estimated lifetime flood cost per hectare and as a percentage of actual flood damage in the zip code. Panel A of Figure 8 indicates that zip codes with a higher proportion of minority population experience disproportionately higher flood risk. The estimated flood cost is $5,837 per hectare of developed land in zip codes in the top quartile by minority proportion, roughly 69 times higher than the risk in zip codes that fall into the least minority quartile. The share of total flood damage attributable to land development is 19.5% in the top quartile by minority proportion, compared to only 3.1% in the bottom quartile.

Concentrating on majority communities, characterized by zip codes with less than a 30% minority population, Panel B consistently reveals that high-income neighborhoods suffer disproportionately higher flood cost, whether assessed through the poverty rate or median family income. The estimated flood cost stands at $423 per hectare of developed land in zip codes within the quartile with the lowest poverty rate and $497 in those with the highest median income, each accounting for approximately 9.0% of the actual flood damage in these zip codes. These shares are approximately 3 to 4 times higher than the share in zip codes falling into the highest poverty rate and lowest median income quartiles.

7. Conclusion

We estimate the effect of land development on flood risk in the continental United States based on zip code-level data over the past two decades. Employing long differences and instrumental variable approaches, we find that an increase in developed land is associated with a significant increase in flood claims after 8 years through a 15-year span. During the 15-year horizon, a one-percent increase in developed land is associated with a 2.08-percent increase in flood damage based on the stacked LD approach and a smaller 1.64-percent increase using the instrument variable approach, both statistically significant.

The estimates remain robust when we include changes in non-developed land covers. The effect of developed land represents a new mechanism in which increased human development activities can have a direct effect on flood risk, aside from the loss of natural covers. Furthermore, the effect is predominantly due to two development sources: cropland and treeland, not other sources. The effect also varies by local initial development conditions, with zip codes with high initial development conditions experiencing significantly higher growth in developed land in the following years and also seeing the highest risk of floods per percent increase in developed land.

Based on these estimates, we quantify a substantial amount of flood cost attributable to land development land in the past decades. Due to the public nature of flood claims, the flood cost is never priced into the original land development and has ultimately been and is expected to continue to be borne by taxpayers. We find that the substantial amount of the long-run cost of land development is spatially clustered in a few hotspots but ultimately receive implicit transfers from the federal government. Zip codes with a higher proportion of black population and high-income majority neighborhoods also experience a disproportionately higher flood risk.

References

Angrist, Joshua D, and Alan B Krueger, 2001, Instrumental variables and the search for identification: From supply and demand to natural experiments, *Journal of Economic perspectives* 15, 69–85.

Banerjee, Abhijit V, and Esther Duflo, 2007, The economic lives of the poor, *Journal of economic perspectives* 21, 141–167.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen, 2019, What drives differences in management practices?, *American Economic Review* 109, 1648–1683.

Busso, Matias, Jesse Gregory, and Patrick Kline, 2013, Assessing the incidence and efficiency of a prominent place based policy, *American Economic Review* 103, 897–947.

Chyn, Eric, and Lawrence F Katz, 2021, Neighborhoods matter: Assessing the evidence for place effects, *Journal of Economic Perspectives* 35, 197–222.

Deininger, Klaus, and Gershon Feder, 2001, Land institutions and land markets, *Handbook of agricultural economics* 1, 287–331.

Ellison, Glenn, and Edward L Glaeser, 1997, Geographic concentration in us manufacturing industries: a dartboard approach, *Journal of political economy* 105, 889–927.

Ellison, Glenn, Edward L Glaeser, and William R Kerr, 2010, What causes industry agglomeration? evidence from coagglomeration patterns, *American Economic Review* 100, 1195–1213.

Elwell, Craig K, 2014, The distribution of household income and the middle class .

Galor, Oded, and David N Weil, 2000, Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond, *American economic review* 90, 806–828.

Gamper-Rabindran, Shanti, and Christopher Timmins, 2013, Does cleanup of hazardous waste sites raise housing values? evidence of spatially localized benefits, *Journal of Environmental Economics and Management* 65, 345–360.

Gobillon, Laurent, and Thierry Magnac, 2016, Regional policy evaluation: Interactive fixed effects and synthetic controls, *Review of Economics and Statistics* 98, 535–551.

Gollin, Douglas, Stephen Parente, and Richard Rogerson, 2002, The role of agriculture in development, *American economic review* 92, 160–164.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski, 2014, Optimal taxes on fossil fuel in general equilibrium, *Econometrica* 82, 41–88.

Greenstone, Michael, and Justin Gallagher, 2008, Does hazardous waste matter? evidence from the housing market and the superfund program, *The Quarterly Journal of Economics* 123, 951–1003.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, 2010, Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings, *Journal of political economy* 118, 536–598.

Gyourko, Joseph, Jonathan S Hartley, and Jacob Krimmel, 2021, The local residential land use regulatory environment across us housing markets: Evidence from a new wharton index, *Journal of Urban Economics* 124, 103337.

Gyourko, Joseph, Albert Saiz, and Anita Summers, 2008, A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index, *Urban Studies* 45, 693–729.

Holmes, Thomas J, 1998, The effect of state policies on the location of manufacturing: Evidence from state borders, *Journal of political Economy* 106, 667–705.

Khan, Shuhab D, 2005, Urban development and flooding in houston texas, inferences from remote sensing data using neural network technique, *Environmental Geology* 47, 1120–1127.

Kline, Patrick, and Enrico Moretti, 2014a, Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority, *The Quarterly journal of economics* 129, 275–331.

Kline, Patrick, and Enrico Moretti, 2014b, People, places, and public policy: Some simple welfare economics of local economic development programs, *Annu. Rev. Econ.* 6, 629–662.

Konrad, Christopher Peter, et al., 2003, Effects of urban development on floods .

Kousky, Carolyn, 2018, Financing flood losses: A discussion of the national flood insurance program, *Risk Management and Insurance Review* 21, 11–32.

Libecap, Gary D, and Dean Lueck, 2011, The demarcation of land and the role of coordinating property institutions, *Journal of Political Economy* 119, 426–467.

Lutz, Chandler, and Ben Sand, 2022, Highly disaggregated land unavailability, *Available at SSRN 3478900* .

Newhouse, Joseph P, and Mark McClellan, 1998, Econometrics in outcomes research: the use of instrumental variables, *Annual review of public health* 19, 17–34.

Nordhaus, William, 2019, Climate change: The ultimate challenge for economics, *American Economic Review* 109, 1991–2014.

Nordhaus, William D., 1977, Economic growth and climate: the carbon dioxide problem, *American Economic Review* 67, 341–346.

Pew Research Center, Washington DC, 2015, The american middle class is losing ground .

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens III, 2010, Housing externalities, *Journal of political Economy* 118, 485–535.

Saiz, Albert, 2010, The geographic determinants of housing supply, *The Quarterly Journal of Economics* 125, 1253–1296.

Slattery, Cailin, and Owen Zidar, 2020, Evaluating state and local business incentives, *Journal of Economic Perspectives* 34, 90–118.

Solow, Robert M, 1956, A contribution to the theory of economic growth, *The quarterly journal of economics* 70, 65–94.

Suárez Serrato, Juan Carlos, and Owen Zidar, 2016, Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms, *American Economic Review* 106, 2582–2624.

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

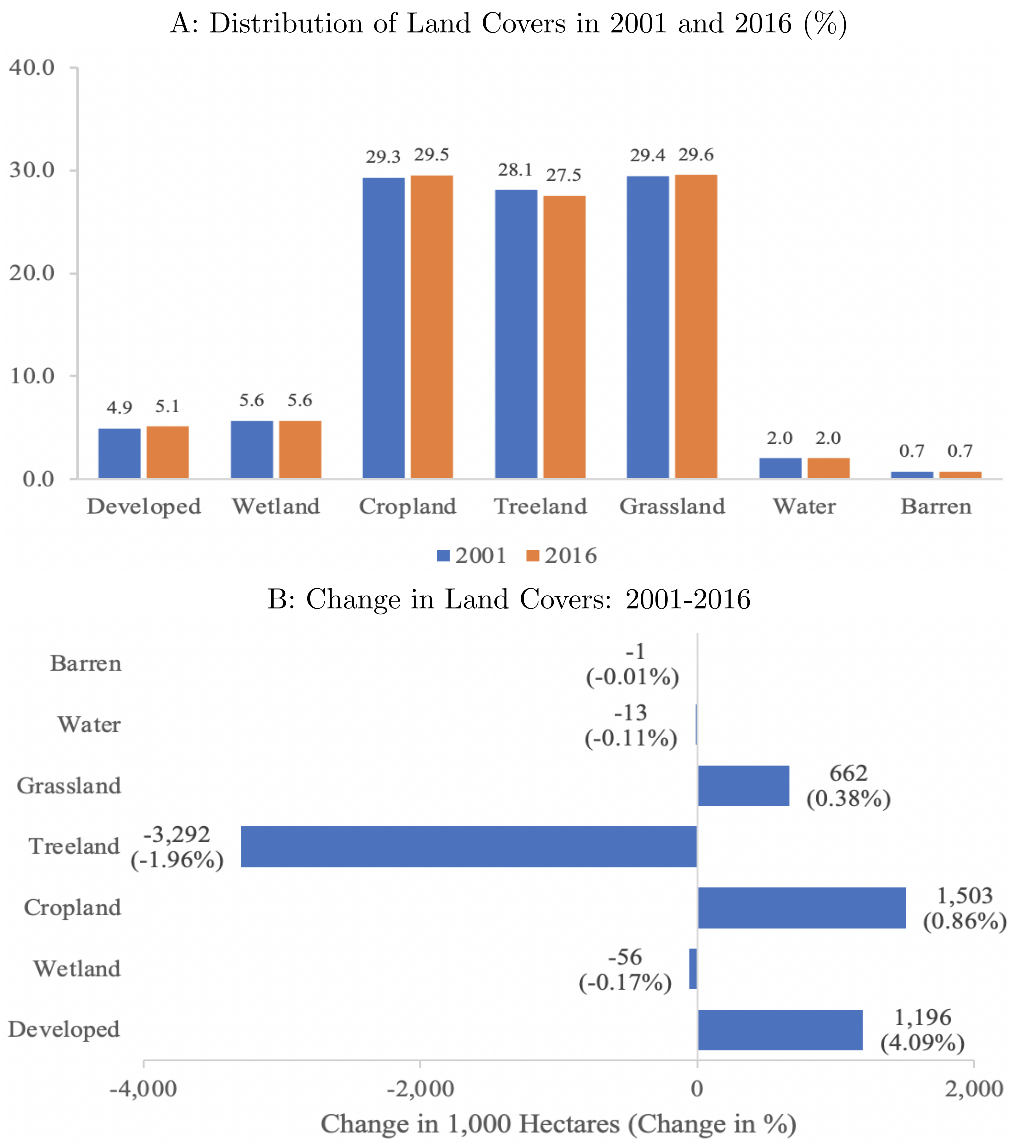
Turner, Matthew A, Andrew Haughwout, and Wilbert Van Der Klaauw, 2014, Land use regulation and welfare, *Econometrica* 82, 1341–1403.

Yan, Hongxiang, and Findlay G Edwards, 2013, Effects of land use change on hydrologic response at a watershed scale, arkansas, *Journal of Hydrologic Engineering* 18, 1779–1785.

Zhang, Wei, Gabriele Villarini, Gabriel A Vecchi, and James A Smith, 2018, Urbanization exacerbated the rainfall and flooding caused by hurricane harvey in houston, *Nature* 563, 384–388.

Zhou, Jian, Daniel P McMillen, and John F McDonald, 2008, Land values and the 1957 comprehensive amendment to the chicago zoning ordinance, *Urban Studies* 45, 1647–1661.

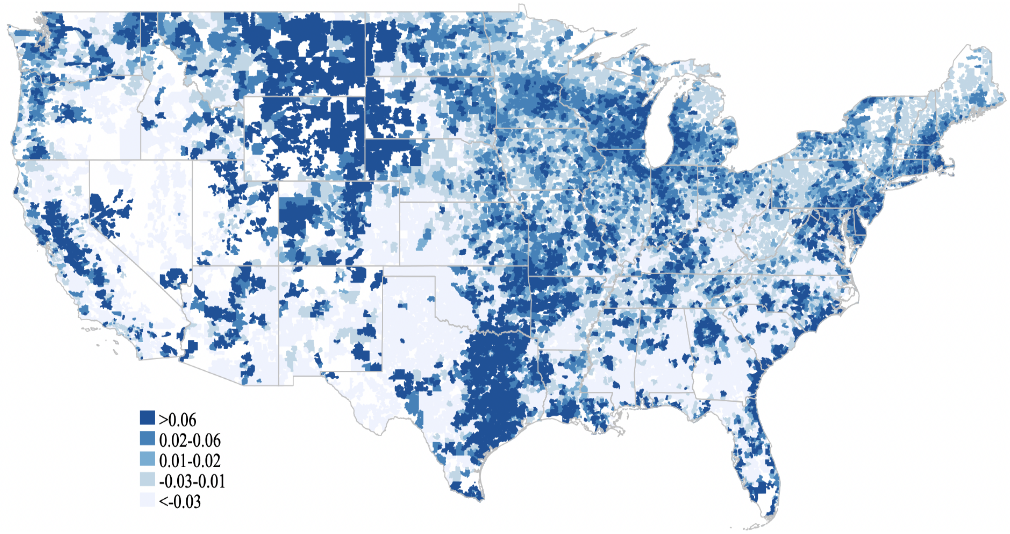
Figure 1. Land Uses in the United States



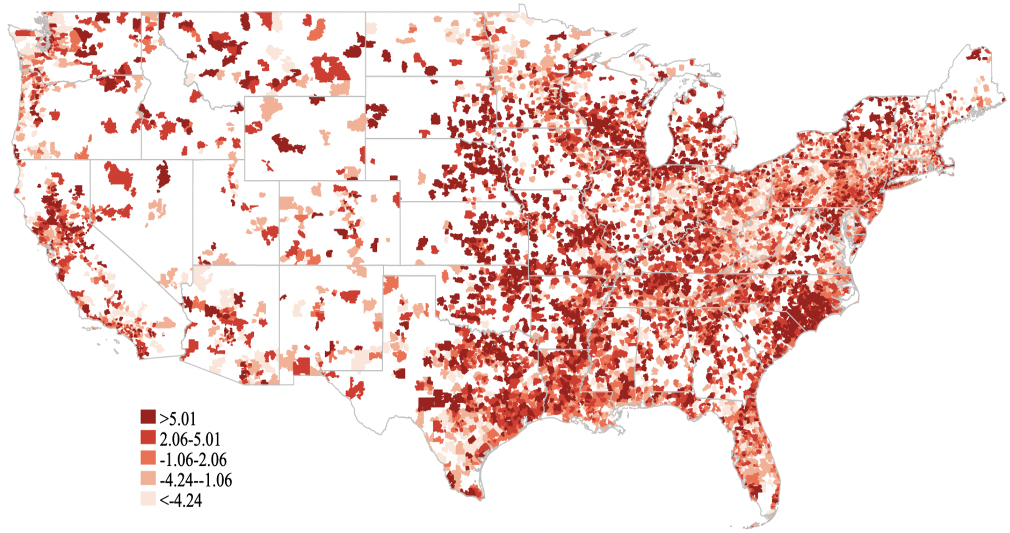
Notes: This figure shows the land uses in the United States. Panel A compares the shares of each land cover in 2001 and 2016. Panel B shows the change in land cover between 2001 and 2016 in the unit of 1,000 hectares, with the change in percentage denoted in parentheses.

Figure 2. 15-Year Change in Developed Land and NFIP Claims

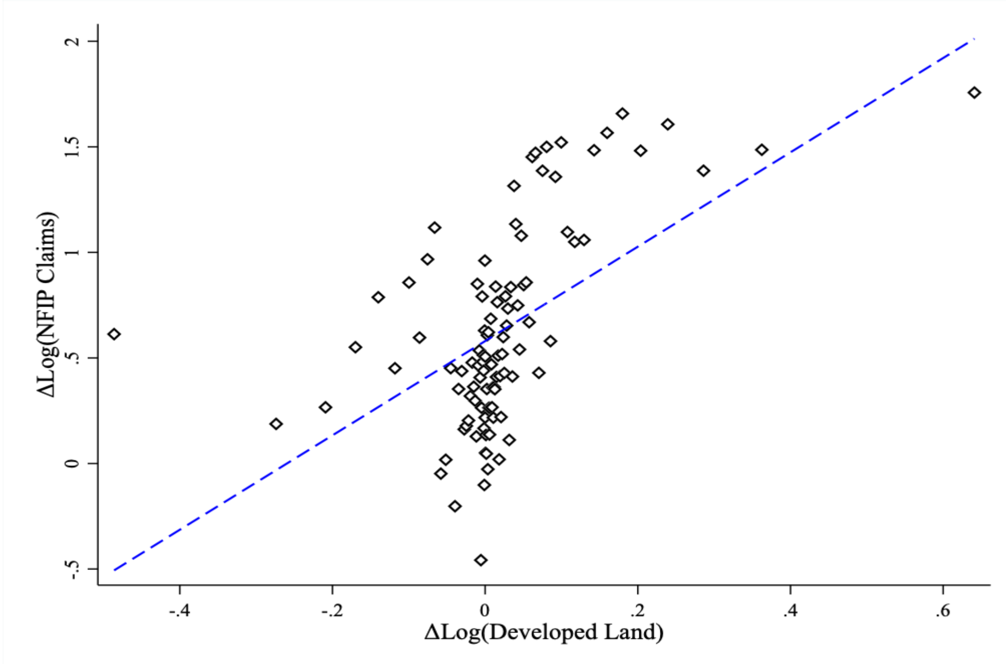
A: ΔLog(Developed Land)



B: ΔLog(Flood Damage)



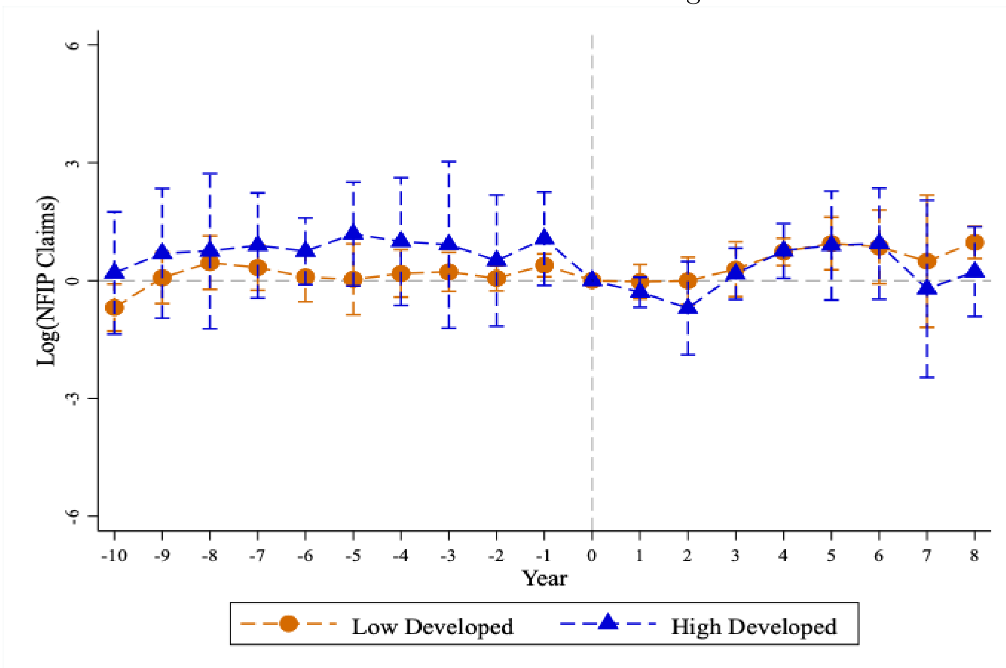
C: Correlation



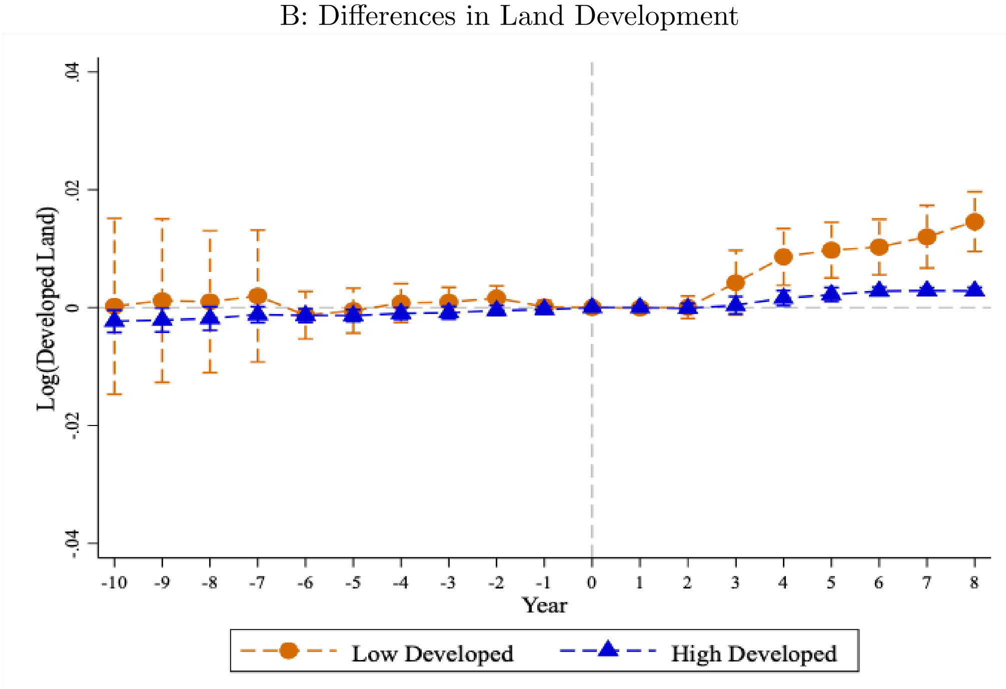
Notes: This figure shows 15-year changes in developed land and flood damage between 2001 and 2016. Panel A is the zip code-level map of the change in logarithm of developed land. Panel B is the zip code-level map of the change in logarithm of NFIP flood damage. Panel C shows the raw binscatter plot of the two variables, with the blue dashed line denoting the linear fitted line.

Figure 3. First-Stage Results of IV Regressions

A: Differences in Flood Damage

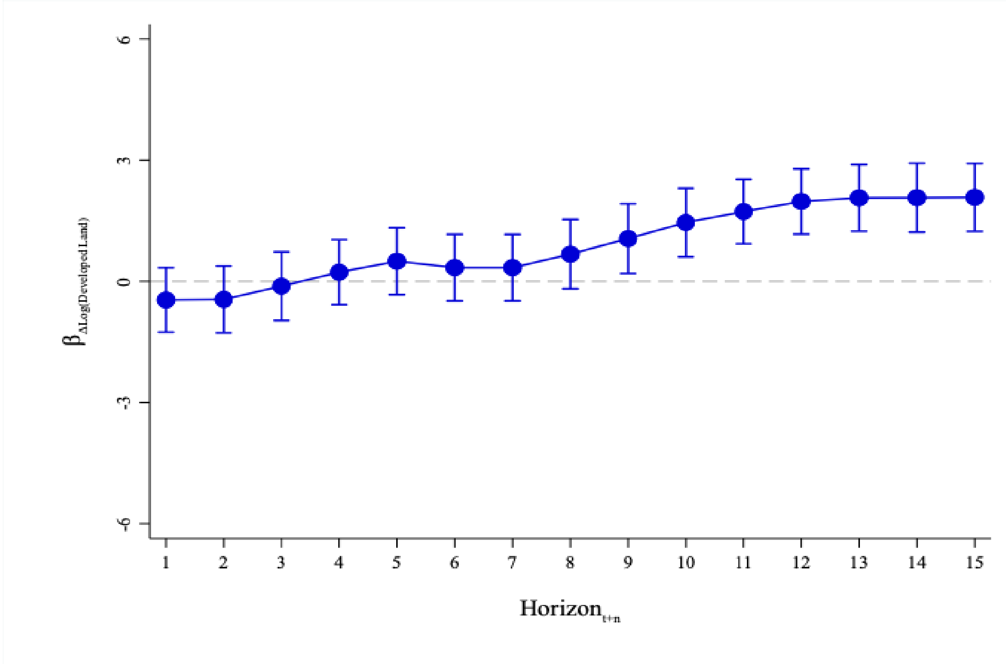


B: Differences in Land Development



Notes: This figure shows the first-stage results of IV regressions, which presents the regressions of the logarithm of flood damage (Panel A) and the logarithm of developed land (Panel B) on the indicators for Post-RTW × Low Developed and Post-RTW × High Developed by the year relative to the RTW enactment year. Lines show 90 percent confidence intervals. We control for state pair × relative year and zip code fixed years, and zip code covariates, including zip code-by-year level population, housing units, housing value, income, and NFIP CRS. Standard errors are clustered at the state pair level.

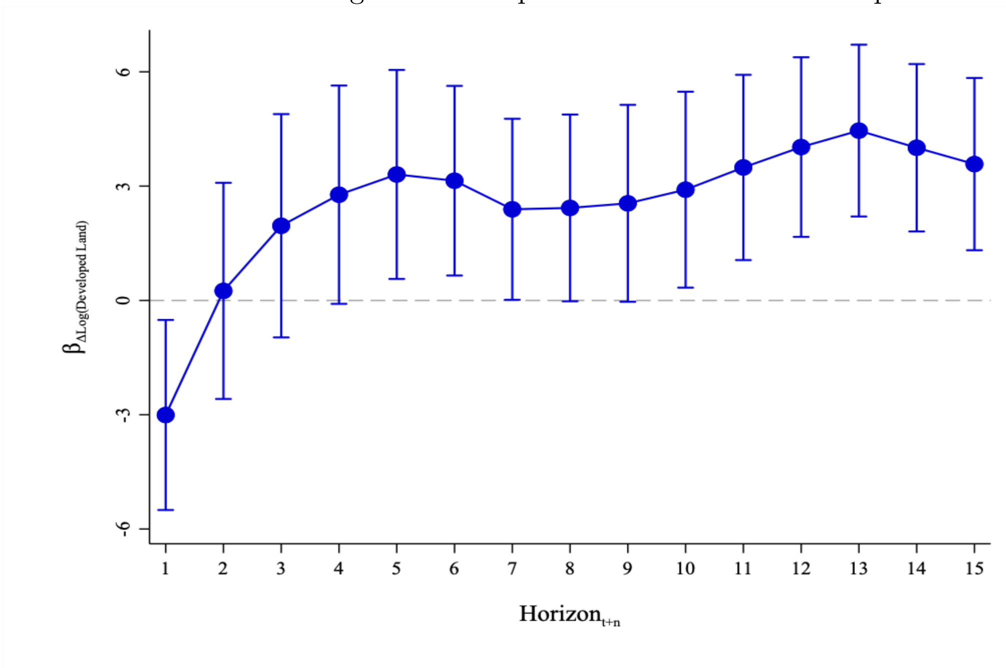
Figure 4. The Effect of Land Development on Flood Damage over Different Horizons



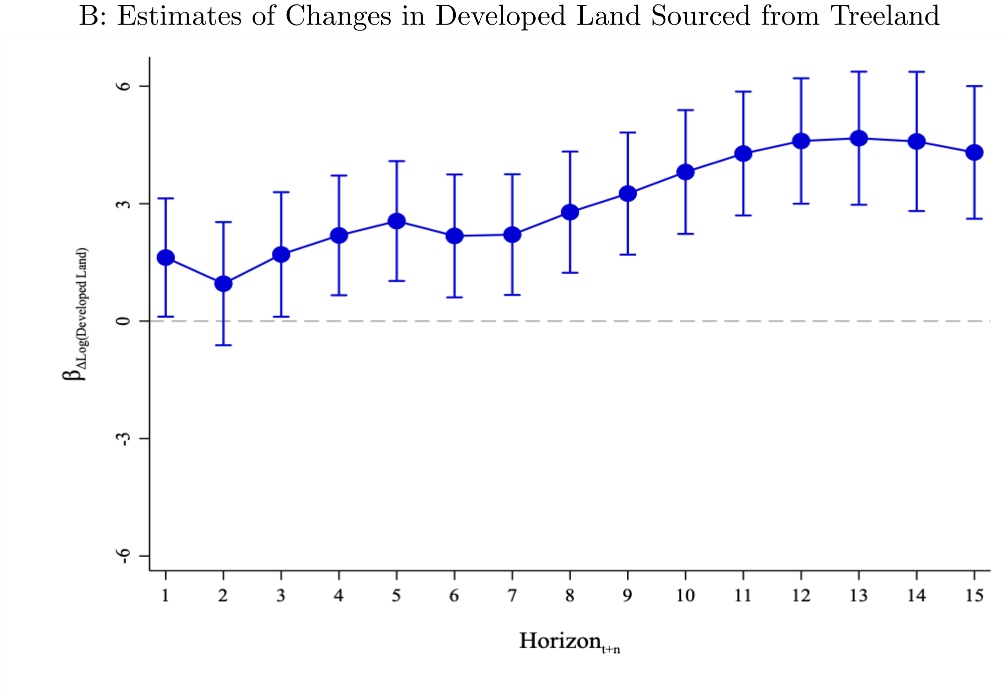
Notes: This figure shows the effect of land development on flood damage over different horizons. Each data point presents a regression of $\Delta Log(FD)$ on $\Delta Log(Dev)$ over a particular horizon, where the x-axis denotes the horizon for the long differences and the y-axis denotes the coefficient. Upper and lower bounds indicate the 90% confidence intervals. We control for zip code covariates, including zip code-by-year level population, housing units, housing value, income, and NFIP CRS. Regressions are based on the stacked LD samples containing differences from four cohort years: 2001 to 2016, 2002 to 2017, 2003 to 2018 and 2004 to 2019. All regressions are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level.

Figure 5. The Effect of Land Development on Flood Damage by Source

A: Estimates of Changes in Developed Land Sourced from Cropland



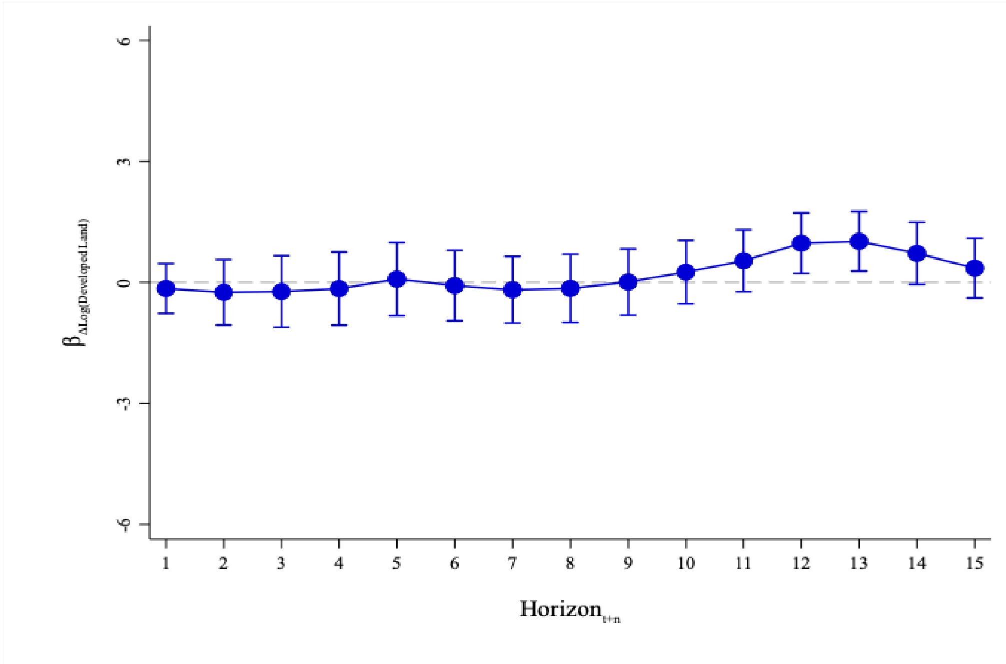
B: Estimates of Changes in Developed Land Sourced from Treeland



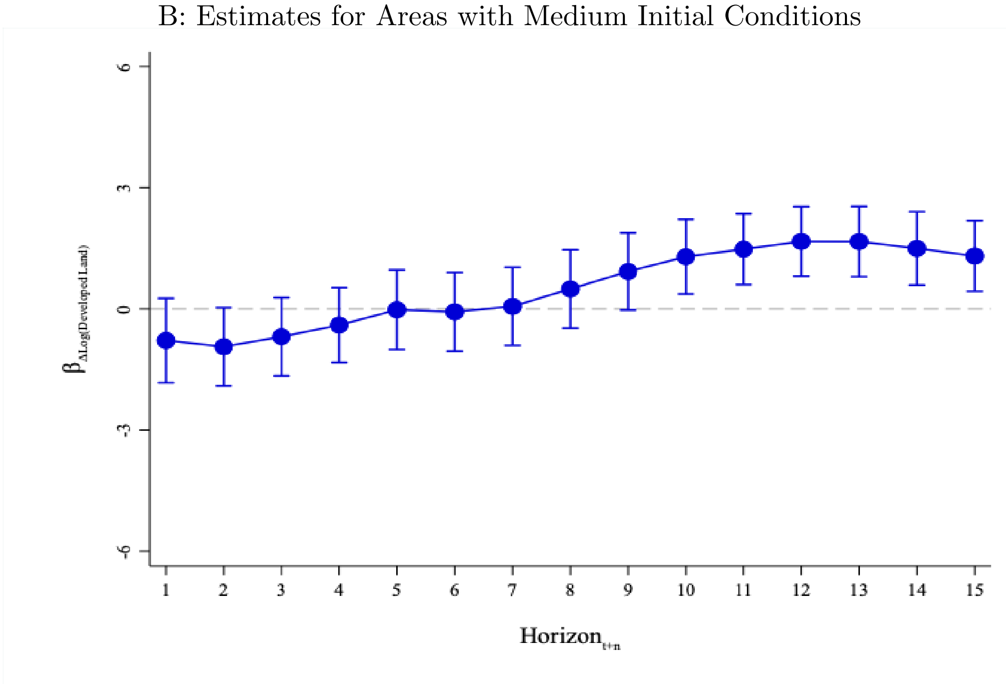
Notes: This figure shows the effect of land development by source on flood damage over different horizons. Each data point presents a regression of $\Delta Log(FD)$ on $\Delta Log(Dev)$ sourced from cropland (Panel A) and treeland (panel B) on a particular horizon, where the x-axis denotes the horizon for the long differences and the y-axis denotes the coefficient. The upper and lower bounds indicate the 90% confidence intervals. We control for zip code covariates, including zip code-by-year level population, housing units, housing value, income, and NFIP CRS. Regressions are based on the stacked LD samples containing differences from four cohort years: 2001 to 2016, 2002 to 2017, 2003 to 2018 and 2004 to 2019. All regressions are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level.

Figure 6. The Effect of Land Development on Flood Damage by Initial Conditions

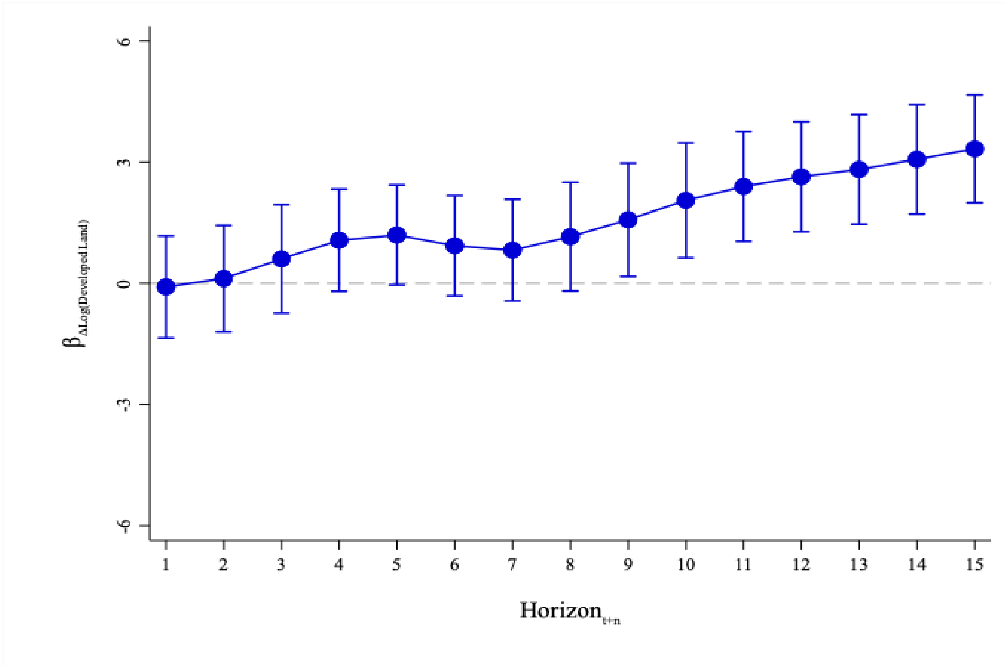
A: Estimates for Areas with Low Initial Conditions



B: Estimates for Areas with Medium Initial Conditions



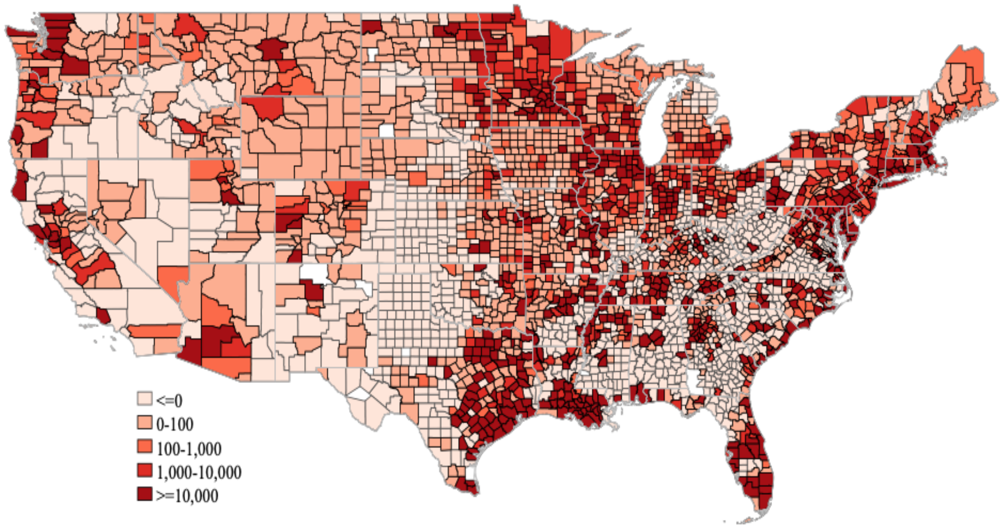
C: Estimates for Areas with High Initial Conditions



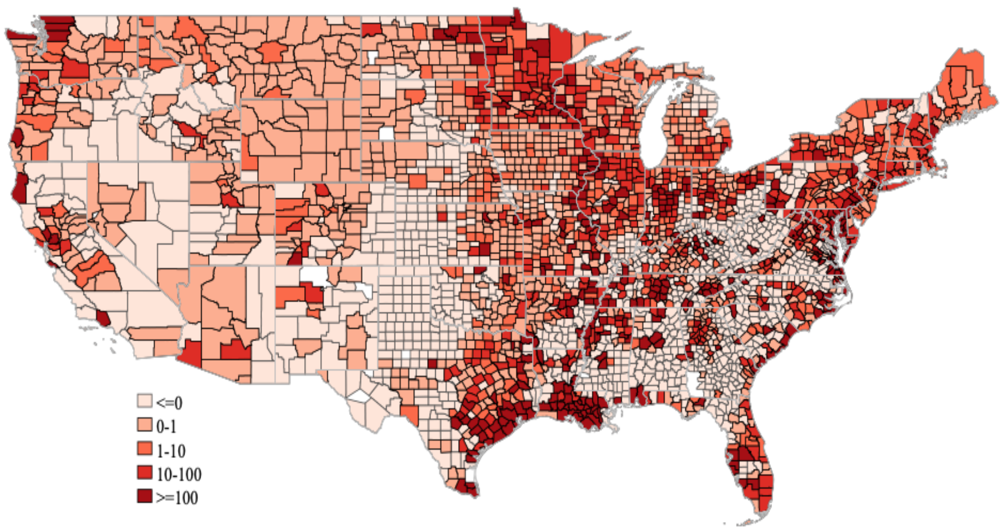
Notes: This figure shows the effect of land development on flood damage over different horizons by the initial development conditions. We classify all zip codes into low initial development conditions as those with initial development level below the median and developed land density above the median, high initial development conditions as those with initial development level above the median and developed land density above the median), and medium initial development conditions as those not in the high or low conditions group. Each data point presents a regression of $\Delta Log(FD)$ on $\Delta Log(Dev)$ over a particular horizon, where the x-axis denotes the horizon for the long differences, and the y-axis denotes the coefficient. The upper and lower bounds indicate the 90% confidence intervals. We control for zip code covariates, including zip code-by-year level population, housing units, housing value, income, and NFIP CRS. Regressions are based on the stacked LD samples containing differences from four cohort years: 2001 to 2016, 2002 to 2017, 2003 to 2018 and 2004 to 2019. All regressions are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level.

Figure 7. Spatial Distribution of The Estimated Long-Run Cost of Land Development

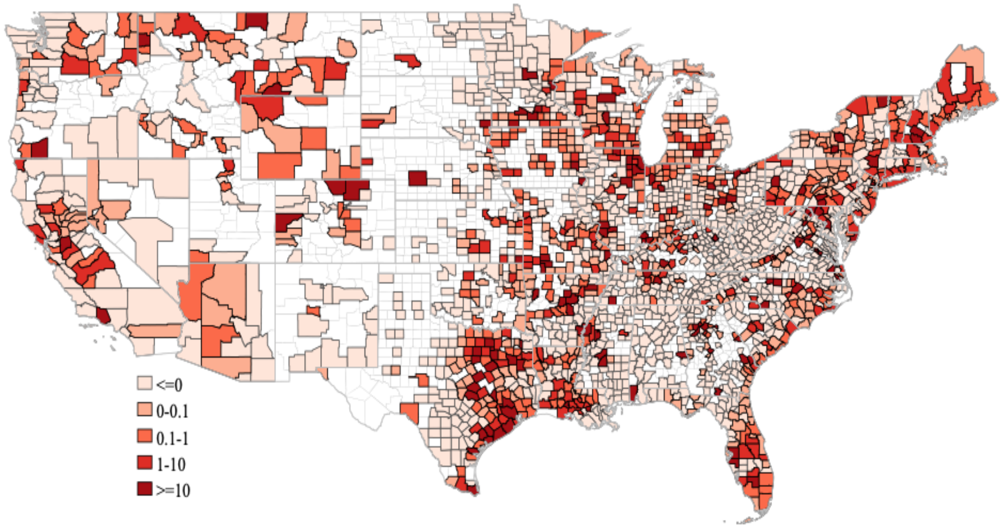
A: Total Estimated Cost ($)



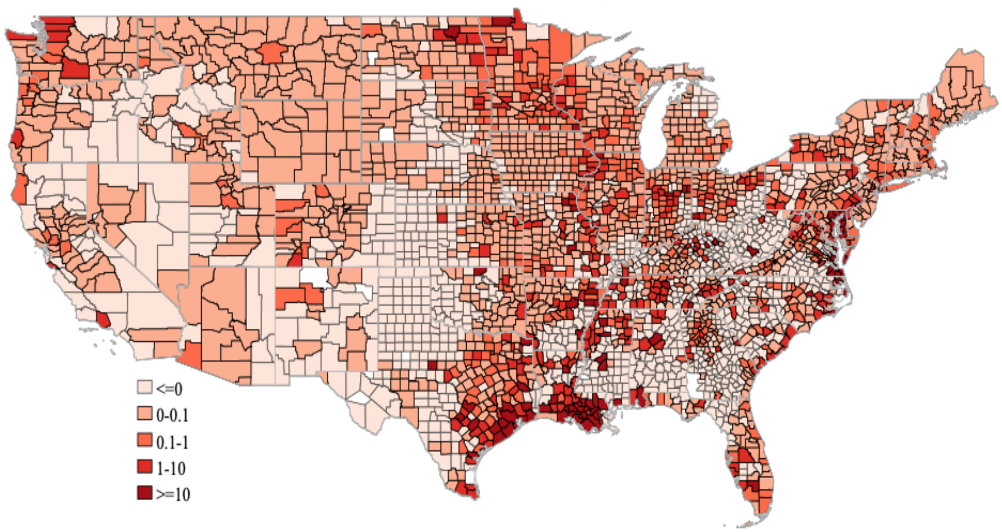
B: Estimated Cost Per Hectare ($)



C: Total Estimated Cost as % of Actual ΔFD



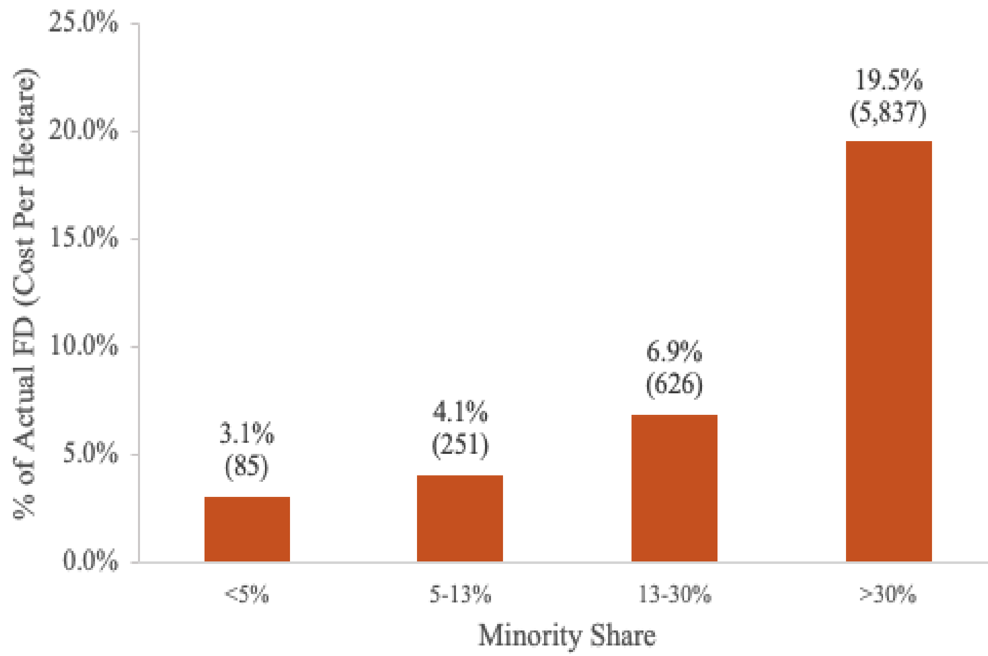
D: Estimated Cost Per Ha as % of Land Value



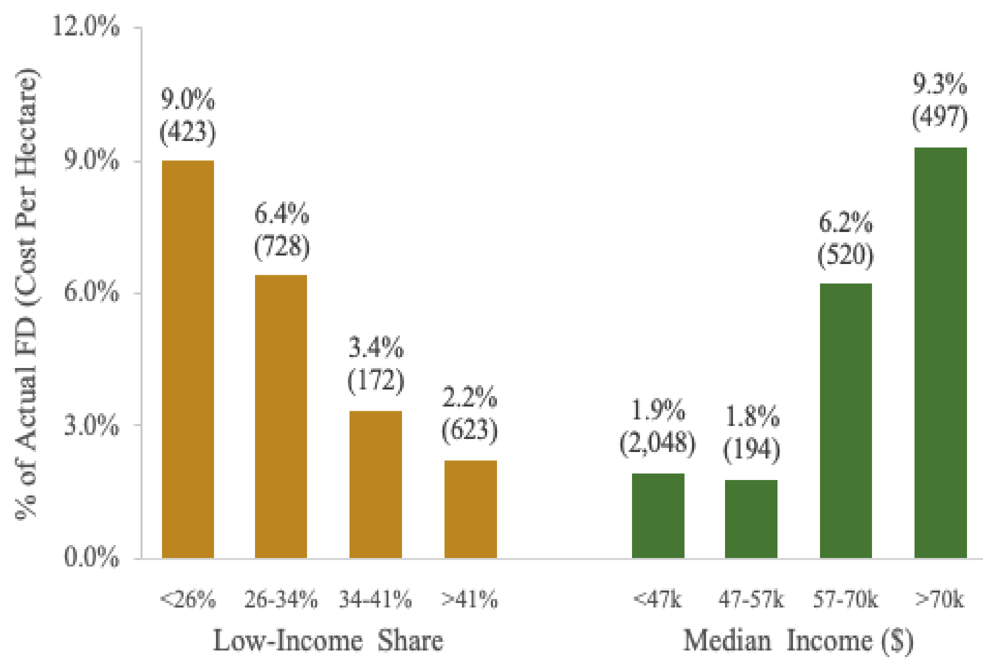
This figure shows the distribution of the estimated long-run cost of land development by county. Panel A plots the total estimated lifetime cost of land development. Panel B plots the estimated cost of land development per hectare. Panel C plots the estimated cost of land development as a percentage of the actual change in flood damage in the area. Panel D is the estimated cost of land development as a percentage of the land value. The long-run flood cost is predicted based on the regression in Table 6 and based on the LD sample containing 15-Year changes from 2001 to 2016.

Figure 8. Demographic Distribution of The Estimated Long-Run Cost of Land Development

A: Estimated Cost by Minority Population Share



B: Estimated Cost by Income Groups for Majority Neighborhoods



This figure shows the demographic distribution of the estimated long-run cost of land development at the zip code level. It reports the estimated cost of land development as a percentage of the actual change in NFIP claims, with the estimated cost of land development per hectare denoted in parentheses. Panel A shows the estimated cost by minority population share at the zip code level. Panel B shows the estimated cost by low-income population share and by median income, within the areas with relatively low share of minority share (quartiles 1-3 in Panel A). Minority is defined to include Black, American Indian, Asian, Pacific Islander, and Hispanics. The low-income population is defined as households whose incomes are less than two thirds of the national median (Pew Research Center, 2015; Elwell, 2014). The long-run flood cost is predicted based on the regression in Table 6 and based on the LD sample containing 15-year changes from 2001 to 2016.

Table 1: Summary Statistics

Panel A: Full Sample

Variable	N	Mean	SD	P1	P5	P50	P95	P99
Main Variables								
Dev_0	105,612	1,117	1,146	29	79	696	3,479	5,177
ΔLog(Dev)	105,612	0.024	0.087	-0.192	-0.064	0.005	0.164	0.352
FD_0 (ha)	105,612	128,991	3,991,956	0.000	0.000	0.000	68,036	629,503
ΔLog(FD)	105,612	0.384	5.281	-11.677	-9.456	0.000	9.891	11.910
Other Land Covers								
ΔLog(Wetland)	105,612	-0.003	0.082	-0.209	-0.052	0.000	0.025	0.213
ΔLog(Cropland)	105,612	0.128	0.442	-1.002	-0.343	0.010	1.024	1.708
ΔLog(Treeland)	105,612	-0.031	0.239	-0.822	-0.332	-0.008	0.193	0.716
ΔLog(Grassland)	105,612	0.131	0.473	-0.971	-0.469	0.025	1.030	1.723
ΔLog(Water)	105,612	0.009	0.230	-0.649	-0.218	0.000	0.259	0.779
ΔLog(Barren)	105,612	0.169	0.556	-1.410	-0.649	0.102	1.126	1.787
Controls								
ΔLog(Housing Units)	105,612	0.108	0.194	-0.385	-0.100	0.078	0.431	0.769
ΔLog(Housing Value)	105,612	0.518	0.251	-0.038	0.154	0.504	0.953	1.174
ΔLog(Income)	105,612	0.340	0.148	0.000	0.115	0.333	0.583	0.766
ΔLog(Population)	105,612	0.079	0.216	-0.495	-0.198	0.058	0.427	0.783
ΔLog(CRS)	105,612	0.011	0.838	-2.918	-1.792	0.000	1.792	2.773

Panel B: RTW Sample

Treatment vs. Control	Treatment				Control	
Initial Condition	Low Developed		High Developed		Both	
Time	Pre-RTW	Post-RTW	Pre-RTW	Post-RTW	Pre-RTW	Post-RTW
FD	10,593	6,914	16,644	8,182	15,871	9,398
	(98,155)	(75,076)	(157,960)	(44,463)	(143,293)	(64,552)
Dev	836	895	1,674	1,736	991	1,023
	(1,015)	(1,054)	(1,219)	(1,264)	(1,125)	(1,142)
Housing Units	2,782	2,938	8,967	9,121	4,231	4,301
	(3,616)	(3,783)	(5,898)	(6,006)	(5,668)	(5,872)
Housing Value	109,991	141,986	133,429	159,174	132,517	153,312
	(46,696)	(55,945)	(69,498)	(88,145)	(85,539)	(94,615)
Income	44,221	56,167	48,765	58,129	50,806	61,710
	(13,172)	(15,020)	(19,919)	(24,002)	(18,798)	(22,077)
Population	6,178	6,368	20,374	20,564	9,865	9,910
	(8,453)	(8,865)	(13,692)	(13,878)	(13,460)	(13,871)
CRS	0.170	0.116	1.197	0.737	0.248	0.260
	(1.434)	(1.114)	(4.815)	(3.269)	(1.879)	(1.925)
N	35,613	14,832	6,841	3,343	161,630	53,184

Notes: This table reports the summary statistics of the variables used in our analysis. Panel A includes the main variables of interest, other land covers, and controls used in our stacked 15-year LD regressions. All the ΔLog variables are the logarithmic change of the variable, weighted by the zip code level share of population, which is also the weight in all regressions throughout the paper. Panel B reports the mean and standard deviation (in parentheses) of the variables in the RTW subsample and used in panel regressions. The sample is further classified into 3 subsample based on whether the state has the RTW law enacted during 2001–2019 (treatment vs. control) and whether the zip code has a low or high initial developed land share. The high or low developed share is defined based on whether the zip code has an initial developed share above or below the 75^{th} percentile within the state.

Table 2: Baseline Results

Panel A: LD Regressions

Dep Var	ΔLog(FD)			
Horizon	1-Year	5-Year	10-Year	15-Year
	(1)	(2)	(3)	(4)
ΔLog(Dev)	-0.461	0.498	1.456***	2.080***
	(0.485)	(0.503)	(0.515)	(0.512)
ΔLog(Housing Units)	0.508*	0.702*	0.715**	0.515
	(0.288)	(0.399)	(0.362)	(0.461)
ΔLog(Housing Value)	-0.070	0.214	0.974***	0.436
	(0.155)	(0.253)	(0.282)	(0.310)
ΔLog(Income)	0.359**	0.265	0.398	0.419
	(0.152)	(0.225)	(0.285)	(0.343)
ΔLog(Population)	-0.155	-0.313	-0.279	-0.203
	(0.255)	(0.360)	(0.340)	(0.434)
ΔLog(CRS)	0.551***	1.163***	1.165***	1.155***
	(0.020)	(0.041)	(0.041)	(0.058)
State × Year FE	Yes	Yes	Yes	Yes
N	475,254	369,642	237,627	105,612
adj. R-sq	0.138	0.186	0.179	0.122

Panel B: Panel Regressions

Dep Var	Log(FD)			
Sample	Full	RTW Subsample		
			IV	
Specification	OLS	OLS	2nd-Stage	1st-Stage
	(1)	(2)	(3)	(4)
Log(Dev)	0.545***	0.506***	1.640*	
	(0.047)	(0.079)	(0.841)	
Log(Housing Units)	0.827***	0.611***	-0.114	0.634***
	(0.082)	(0.197)	(0.568)	(0.064)
Log(Housing Value)	0.627***	0.019	0.605	-0.520***
	(0.086)	(0.114)	(0.517)	(0.137)
Log(Income)	-0.156	0.223	-0.710	0.821***
	(0.107)	(0.203)	(0.788)	(0.120)
Log(Population)	-0.337***	-0.064	-0.261	0.180***
	(0.086)	(0.175)	(0.203)	(0.052)
Log(CRS)	1.949***	2.043***	2.040***	0.005
	(0.069)	(0.075)	(0.073)	(0.025)
I(Post RTW × Low Developed)				0.065***
				(0.015)
I(Post RTW × High Developed)				-0.225***
				(0.052)
State × Year FE	Yes			
State FE		Yes	Yes	Yes
State Pair × Relative Year FE		Yes	Yes	Yes
N	501,657	275,443	275,443	275,443
adj. R-sq	0.333	0.204	0.152	0.822
Underidentification test (Kleibergen-Paap rk LM stat.)				17.07
Weak identification test (Cragg-Donald Wald F stat.)				433.06
Weak identification test (Kleibergen-Paap Wald rk F stat.)				11.553
Overidentification test (Hansen J stat.)				0.733

Notes: This table reports the baseline results for the effects of developed land on flood damage. Panel A is based on the stacked ΔLog specification using 1-year, 5-year, 10-year, and 15-year differences. Panel B is based on the RTW sample, where the treatment states are those with RTW Law enacted during 2001-2019, and the control groups are the states adjacent to them but without RTW Law (with repeated observations). The dependent variable and the main explanatory variable are $\Delta Log(FD)$ and $\Delta Log(Dev)$ over different horizons, respectively, in Panel A and $Log(FD)$ and $Log(Dev)$ over different horizons, respectively, in Panel B. We control for state by cohort year fixed effects and zip code covariates include zip code by year level population, housing units, housing value, income, and NFIP CRS. Regressions are based on the stacked LD samples containing differences from four cohort years: 2001 to 2016, 2002 to 2017, 2003 to 2018 and 2004 to 2019. All regressions are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level. Asterisks denote significance levels (***=1%, **=5%, *=10%).

Table 3: The Impacts of Developed Land Sources

Panel A: Effect of Non-Developed Land and Land Development Sources

	ΔLog(FD)							$X\beta \times 100$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔLog(Dev)	2.080***	2.014***	1.993***	2.094***				
	(0.512)	(0.490)	(0.490)	(0.505)				
ΔLog(Dev: Wetland)					3.361	4.454	4.827	0.411
					(4.200)	(4.198)	(4.233)	
ΔLog(Dev: Cropland)					4.133***	4.024***	4.309***	5.602
					(1.038)	(1.022)	(1.030)	
ΔLog(Dev: Treeland)					3.076**	3.013**	3.578***	5.921
					(1.373)	(1.369)	(1.374)	
ΔLog(Dev: Grassland)					-0.186	-0.176	-0.262	0.009
					(0.757)	(0.761)	(0.759)	
ΔLog(Dev: Water)					-1.010	-1.946	-1.012	-0.001
					(4.324)	(4.292)	(4.204)	
ΔLog(Dev: Barren)					2.755	2.254	1.141	0.278
					(3.088)	(3.171)	(2.944)	
ΔLog(Wetland)			0.591	0.524		0.607	0.524	
			(0.442)	(0.454)		(0.444)	(0.453)	
ΔLog(Cropland)				0.003			0.053	
				(0.167)			(0.168)	
ΔLog(Treeland)				0.325			0.359	
				(0.236)			(0.234)	
ΔLog(Grassland)				-0.106			-0.133	
				(0.090)			(0.090)	
ΔLog(Water)				0.088			0.063	
				(0.146)			(0.144)	
ΔLog(Barren)				-0.078			-0.084	
				(0.071)			(0.071)	
ΔLog(Non-Dev)		-0.149						
		(0.171)						
ΔLog(Non-Dev/Wetland)			-0.177			-0.132		
			(0.164)			(0.168)		
State-by-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Zip Code Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
N	105,612	105,612	105,612	105,612	105,612	105,612	105,612	
adj. R-sq	0.122	0.122	0.122	0.122	0.123	0.123	0.123	

Panel B: Summary of Land Development Sources

Source of Developed Land	1-Year	5-Year	10-Year	15-Year
Change in Developed Land (ha)				
Dev: Wetland	0.06	0.30	0.51	0.95
Dev: Cropland	1.09	4.55	7.17	14.52
Dev: Treeland	1.24	6.05	11.12	18.48
Dev: Grassland	0.06	-0.36	-2.44	-0.37
Dev: Water	0.01	0.00	-0.04	0.01
Dev: Barren Land	0.17	0.92	1.83	2.72
Total	2.63	11.46	18.15	36.32
% of Change in Developed Land (%)				
Dev: Wetland	2.43	2.64	2.78	2.62
Dev: Cropland	41.54	39.73	39.50	39.98
Dev: Treeland	47.07	52.80	61.26	50.89
Dev: Grassland	2.15	-3.15	-13.42	-1.02
Dev: Water	0.32	-0.01	-0.22	0.04
Dev: Barren Land	6.49	7.99	10.10	7.50
Total	100.00	100.00	100.00	100.00
Percentage Change (%)				
Dev: Wetland	0.01	0.03	0.04	0.09
Dev: Cropland	0.10	0.40	0.63	1.30
Dev: Treeland	0.11	0.53	0.98	1.65
Dev: Grassland	0.00	-0.03	-0.21	-0.03
Dev: Water	0.00	0.00	0.00	0.00
Dev: Barren Land	0.01	0.08	0.16	0.24
Total	0.23	1.01	1.60	3.25

Notes: This table reports the impact of developed land from different sources on flood damage. Panel A reports the regression of $\Delta Log(FD)$ on $\Delta Log(Dev)$ over the 15-year horizon. We control for state by cohort year fixed effects and zip code covariates include the logarithmic changes in zip code by year level population, housing units, housing value, income, and NFIP CRS. Regressions are based on the stacked LD samples containing differences from four cohort years: 2001 to 2016, 2002 to 2017, 2003 to 2018 and 2004 to 2019. All regressions are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level. Asterisks denote significance levels (***=1%, **=5%, *=10%). Panel B reports the summary of land development sourced from other land covers over 1-year, 5-year, 10-year and 15-year horizon.

Table 4: The Impacts of Initial Conditions

Panel A: Land Covers by Initial Conditions

	Initial Condition		
	Low	Medium	High
Initial Land Attributes			
Initial Developed Share (%)	3.45	17.30	42.90
Initial Density Index	826	1998	3813
NFIP			
FD_0 ($)	19,504	68,687	273,419
FD Change ($)	-5,309	-26,872	-78,222
FD Change (%)	-27.22	-39.12	-28.61
Change in Developed Land (ha)			
Dev: Wetland	0.33	0.95	1.41
Dev: Cropland	0.29	6.63	33.35
Dev: Treeland	5.51	13.55	33.29
Dev: Grassland	-4.39	-6.57	9.09
Dev: Water	-0.02	0.01	0.04
Dev: Barren Land	0.85	2.24	4.62
Total	2.56	16.81	81.80
% of Change in Developed Land (%)			
Dev: Wetland	12.88	5.64	1.73
Dev: Cropland	11.25	39.43	40.77
Dev: Treeland	214.64	80.61	40.69
Dev: Grassland	-171.22	-39.09	11.12
Dev: Water	-0.68	0.07	0.05
Dev: Barren Land	33.12	13.36	5.65
Total	100.00	100.00	100.00
Percentage Change (%)			
Dev: Wetland	0.11	0.11	0.07
Dev: Cropland	0.09	0.78	1.68
Dev: Treeland	1.76	1.59	1.67
Dev: Grassland	-1.41	-0.77	0.46
Dev: Water	-0.01	0.00	0.00
Dev: Barren Land	0.27	0.26	0.23
Total	0.82	1.97	4.11
No. of Zip Codes	7,043	9,891	9,469

Panel B: Estimates by Initial Conditions

	ΔLog(FD)			
	1-Year	5-Year	10-Year	15-Year
	(1)	(2)	(3)	(4)
ΔLog(Dev) × I(Initial Low Conditions)	-0.152	0.082	0.255	0.354
	(0.376)	(0.551)	(0.478)	(0.451)
ΔLog(Dev) × I(Initial Medium Conditions)	-0.784	-0.023	1.289**	1.307**
	(0.635)	(0.599)	(0.562)	(0.531)
ΔLog(Dev) × I(Initial High Conditions)	-0.088	1.197	2.054**	3.330***
	(0.768)	(0.753)	(0.867)	(0.812)
Zip Code Covariates	Yes	Yes	Yes	Yes
State × Year FE	Yes	Yes	Yes	Yes
N	475,254	369,642	237,627	105,612
adj. R-sq	0.138	0.186	0.179	0.122

Notes: This table reports the impact of developed land on flood damage by different initial development conditions. We classify all zip codes into low initial development conditions as those with initial development level below the median and developed land density above the median, high initial development conditions as those with initial development level above the median and developed land density above the median), and medium initial development conditions as those not in the high or low conditions group. Panel A reports the summary of flood damage and land development sourced from other land covers over 1-year, 5-year, 10-year and 15-year horizon. Panel B reports the regression of $\Delta Log(FD)$ on $\Delta Log(Dev)$ over the 15-year horizon. We control for state by cohort year fixed effects and zip code covariates include the logarithmic changes in zip code by year level population, housing units, housing value, income, and NFIP CRS. Regressions are based on the stacked LD samples containing differences from four cohort years: 2001 to 2016, 2002 to 2017, 2003 to 2018 and 2004 to 2019. All regressions are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level. Asterisks denote significance levels (***=1%, **=5%, *=10%).

Table 5: Demographic Characteristics of Developed Land

Panel A: Demographic Characteristics in 2000

Dep Var	Share of Minority	Share of Low Income	Log of Housing Units	Log of Housing Value	Log of Median Income	Log of Population
	(1)	(2)	(3)	(4)	(5)	(6)
			Specification (1)			
ΔLog(Dev)	-0.397***	-0.439***	0.041	0.935***	1.017***	0.302**
	(0.041)	(0.020)	(0.136)	(0.077)	(0.048)	(0.140)
			Specification (2)			
Initial Share: Cropland	-0.247***	0.010	-2.379***	-0.455***	-0.072***	-2.249***
	(0.019)	(0.010)	(0.056)	(0.044)	(0.027)	(0.063)
Initial Share: Treeland	-0.348***	0.058***	-2.189***	-0.397***	-0.155***	-2.435***
	(0.017)	(0.011)	(0.054)	(0.043)	(0.030)	(0.061)
Initial Share: Other Non-Dev	-0.229***	0.061***	-2.196***	-0.508***	-0.169***	-2.462***
	(0.021)	(0.013)	(0.070)	(0.046)	(0.034)	(0.074)
			Specification (3)			
ΔLog(Dev: Cropland)	-0.290***	-0.371***	-0.485**	0.538***	0.860***	-0.030
	(0.095)	(0.047)	(0.228)	(0.170)	(0.113)	(0.234)
ΔLog(Dev: Treeland)	-1.333***	-0.735***	-2.109***	1.277***	1.757***	-1.978***
	(0.138)	(0.064)	(0.270)	(0.198)	(0.156)	(0.286)
ΔLog(Dev: Other)	-0.066	-0.331***	1.268***	1.037***	0.746***	1.388***
	(0.068)	(0.034)	(0.259)	(0.118)	(0.083)	(0.265)
			Specification (4)			
I(Initial Medium Conditions)	0.041***	-0.007***	0.979***	0.107***	0.031***	1.027***
	(0.005)	(0.002)	(0.023)	(0.012)	(0.006)	(0.024)
I(Initial High Conditions)	0.110***	-0.058***	2.396***	0.349***	0.174***	2.528***
	(0.009)	(0.004)	(0.027)	(0.018)	(0.011)	(0.028)
			Specification (5)			
ΔLog(Dev) × I(Initial Low)	-0.116*	-0.192***	-0.296	0.567***	0.420***	-0.130
	(0.060)	(0.022)	(0.353)	(0.097)	(0.057)	(0.354)
ΔLog(Dev) × I(Initial Medium)	-0.303***	-0.207***	-0.965***	0.402***	0.464***	-0.834***
	(0.050)	(0.023)	(0.152)	(0.113)	(0.059)	(0.159)
ΔLog(Dev) × I(Initial High)	-0.518***	-0.642***	0.765***	1.358***	1.503***	1.131***
	(0.054)	(0.028)	(0.194)	(0.100)	(0.065)	(0.205)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
N	25,270	25,270	25,270	25,270	25,270	25,270

Panel B: Change in Demographic Characteristics During 15-Year Span

Dep Var	ΔShare of Minority	ΔShare of Low-Income	ΔLog of Housing Units	ΔLog of Housing Value	ΔLog of Median Income	ΔLog of Population
	(1)	(2)	(3)	(4)	(5)	(6)
			Specification (1)			
ΔLog(Dev)	0.126***	0.076***	1.211***	0.126***	0.121***	1.324***
	(0.014)	(0.010)	(0.050)	(0.027)	(0.017)	(0.054)
			Specification (2)			
Initial Share: Cropland	-0.060***	-0.042***	0.145***	-0.050**	0.062***	0.138***
	(0.005)	(0.004)	(0.012)	(0.021)	(0.010)	(0.015)
Initial Share: Treeland	-0.076***	-0.032***	0.090***	-0.069**	0.036***	0.036**
	(0.005)	(0.005)	(0.014)	(0.029)	(0.012)	(0.018)
Initial Share: Other Non-Dev	-0.053***	-0.051***	0.142***	0.001	0.061***	0.116***
	(0.009)	(0.008)	(0.022)	(0.026)	(0.021)	(0.025)
			Specification (3)			
ΔLog(Dev: Cropland)	0.133***	0.033*	1.425***	0.053	0.174***	1.587***
	(0.024)	(0.019)	(0.065)	(0.052)	(0.032)	(0.070)
ΔLog(Dev: Treeland)	0.146***	0.058*	1.694***	0.274***	0.311***	1.819***
	(0.048)	(0.032)	(0.096)	(0.079)	(0.057)	(0.101)
ΔLog(Dev: Other)	0.103***	0.111***	0.753***	0.121***	-0.007	0.808***
	(0.023)	(0.021)	(0.099)	(0.039)	(0.037)	(0.108)
			Specification (4)			
I(Initial Medium Conditions)	0.017***	0.005***	0.014**	0.017*	-0.001	0.048***
	(0.001)	(0.001)	(0.007)	(0.009)	(0.006)	(0.008)
I(Initial High Conditions)	0.054***	0.036***	0.003	-0.004	-0.057***	0.057***
	(0.002)	(0.002)	(0.007)	(0.010)	(0.006)	(0.009)
			Specification (5)			
ΔLog(Dev) × I(Initial Low)	0.039**	0.027*	0.444***	0.166***	0.127***	0.512***
	(0.018)	(0.015)	(0.071)	(0.035)	(0.029)	(0.073)
ΔLog(Dev) × I(Initial Medium)	0.027*	0.026**	0.918***	0.141***	0.150***	0.971***
	(0.016)	(0.013)	(0.073)	(0.031)	(0.023)	(0.081)
ΔLog(Dev) × I(Initial High)	0.210***	0.118***	1.566***	0.107***	0.101***	1.727***
	(0.019)	(0.013)	(0.049)	(0.037)	(0.024)	(0.052)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
N	25,270	25,270	25,270	25,270	25,270	25,270

Notes: This table reports the demographic characteristics of developed land. The explanatory variables are the change in developed land in Specification (1), initial shares of different land covers that measure the natural endowment in Specification (2), the change in developed land from different sources in Specification (3), the initial conditions categories in Specification (4), and the interaction of the change in developed land and the initial conditions categories in Specification (5). The dependent variables are the proportion of minority, the proportion of low-income population, the logarithm of housing units, the logarithm of housing value, the logarithm of median income, and the logarithm of population in 2001 in Panel A and changes in these variables from 2001 to 2016 in Panel B. Minority is defined as the population of Black, American Indian, Asian, Pacific Islander, and Hispanics. The low-income population is defined as households whose incomes are below two-thirds of the national median. All regressions in Panel B are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level. Asterisks denote significance levels (***=1%, **=5%, *=10%).

Table 6: Prediction Model

Dep Var	ΔLog(FD)$_{i,t_0+15}$		
	Initial Conditions		
	Low	Medium	High
ΔLog(Dev: Cropland)$_{i,t_0+15}$	1.512	2.595**	5.892***
	(1.011)	(1.097)	(1.434)
ΔLog(Dev: Treeland)$_{i,t_0+15}$	0.031	0.385	6.170***
	(1.754)	(1.912)	(1.674)
ΔLog(Dev: Other)$_{i,t_0+15}$	-0.414	0.696	-0.961
	(0.503)	(0.724)	(1.420)
Zip Code Covariates		Yes	
State × Year FE		Yes	
N		105,612	
adj. R-sq		0.123	

Notes: This table reports the impact of developed land from different sources and initial development conditions on flood damage. The dependent and main explanatory variables are $\Delta Log(FD)$ on $\Delta Log(Dev)$ over the 15-year horizon, respectively. We control for state by cohort year fixed effects and zip code covariates include changes in zip code by year level population, housing units, housing value, income, and NFIP CRS. Regressions are based on the stacked LD samples containing differences from four cohort years: 2001 to 2016, 2002 to 2017, 2003 to 2018 and 2004 to 2019. All regressions are weighted by zip code level population share (as a percent of national population) at the year level. Standard errors are clustered at the county level. Asterisks denote significance levels (***=1%, **=5%, *=10%).