NBER WORKING PAPER SERIES

TIME CONSISTENT INFRASTRUCTURE INVESTMENTS: OPTIMAL FLOOD PROTECTION POLICIES IN SPATIAL EQUILIBRIUM

Amine Ouazad
Matthew E. Kahn

Working Paper 33333
http://www.nber.org/papers/w33333

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2025

We thank Robert Huang and Nicholas Trottier-Lacourse for outstanding research assistance. We thank the audiences of the seminar at the Wharton School of Business, University of Pennsylvania, the MIT Center for Real Estate's First Climate Summit in 2022, the University of Wisconsin, the Urban Economics Association, for feedback on early versions of the paper. We thank Hafedh Bouakez, Chetan Dave, Gilles Duranton, Fernando Ferreira, Ambika Gandhi, Jesse Handbury, Jeffrey Lin, Luis Quintero, Nancy Wallace, Pang Xinle, Siqi Zheng for insightful comments in the early stages of the project. The usual disclaimers apply. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by Amine Ouazad and Matthew E. Kahn. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Time Consistent Infrastructure Investments: Optimal Flood Protection Policies in Spatial Equilibrium
Amine Ouazad and Matthew E. Kahn
NBER Working Paper No. 33333
January 2025
JEL No. H43, Q54, R31

ABSTRACT

Place-based investments can have unintended general equilibrium effects and face challenges of time inconsistency. This paper simulates the granular impact of alternative spatial and temporal designs of such investments, using Quantitative Spatial Models where the strategy of the policymaker is endogenized, with time-consistent policy analysis or policies with commitment. It can apply to sunk, fixed costs investments in transportation infrastructure, levees, and other location-based investments. Applying this framework to the 1936 Flood Control Act, the largest investment in flood control infrastructure in US history protecting 5% of land, the study examines the general equilibrium effects of levee investments on housing prices, population density, and racial demographics over eight decades. Protected neighborhoods initially had lower property values, higher minority shares, and greater flood risk, but experienced sustained property appreciation and changes in population density. Structural analysis reveals that optimal levee designs prioritize high-density areas, reduce price capitalization, and minimize urban sprawl. Policymakers who cannot commit to long-term plans tend to overbuild and maintain larger systems compared to those with time-consistent strategies.

Amine Ouazad
HEC Montreal
Department of Economics
Montreal, H3T 2A7, QC
Canada
amine.ouazad@hec.ca

Matthew E. Kahn
Department of Economics
University of Southern California
3620 South Vermont Ave.
Kaprielian (KAP) Hall, 300
Los Angeles, CA 90089-0253
and NBER
kahnme@usc.edu

Suppose the socially desirable outcome is not to have houses built in a particular flood plain but, given that they are there, to take certain costly flood-control measures. If the government's policy were not to build the dams and levees needed for flood protection and agents knew this was the case, even if houses were built there, rational agents would not live in the flood plains. But the rational agent knows that, if he and others build houses there, the government will take the necessary flood-control measures. Consequently, in the absence of a law prohibiting the construction of houses in the flood plain, houses are built there, and the army corps of engineers subsequently builds the dams and levees.

Kydland, F.E. and Prescott, E.C., 1977. Rules rather than discretion: The inconsistency of optimal plans. Journal of political economy, 85(3), pp.473-491.

1 Introduction

Public investments in infrastructure, such as levees, highways, or public transit, yield spatial and temporal effects that challenge policymakers to optimize policies to maximize long-term welfare given their dynamic constraints. These investments not only reshape economic activity across space but also affect private decisions about location, construction, and migration in ways that persist for decades. This paper develops a general framework for modeling and optimizing these investments in settings where agents' behaviors and future policymakers' actions depend on the timeline, credibility, commitment and spatial scope of public goods provision.

Levees represent a key example of spatial investments that trades off immediate benefits with long-term impacts on prices, construction, and flood exposure. First signed into law by Franklin Delano Roosevelt in 1936, the Flood Control Act of 1936 led to the construction of a federal levee system protecting an estimated 100 million acres (Arnold 1988), a unique experiment in U.S. history. These investments are central to urban economics because cities enable agglomeration economies (Rosenthal & Strange 2004, Glaeser 2010), are predominantly located along rivers and coasts (Rappaport & Sachs 2003, Bleakley & Lin 2012, Albouy 2016), making them vulnerable to sea level rise and floods.[1]

Policymakers face important decisions about whether to invest in levees – whether to maintain existing structures, build new ones, change their design, or proceed with managed retreat (Hino, Field & Mach 2017, Mach & Siders 2021). These decisions affect not only the welfare of current residents but also the long-term spatial equilibrium of cities, including patterns of population density, racial composition, and housing prices. Existing frameworks could benefit from incorporating the dynamic interplay between public investments, private location choices, and the policymaker's ability to commit to long-term plans.

[1] Figure 3.5, page 394, Shared Socioeconomic Pathways 1–5. See NOAA's Sea Level Rise and Coastal Flooding Impacts. River flow has increased by 10 to 20% along the Mississippi river, in California, along the Pacific Coast, and in South Florida. Figure 4.7(a), page 574, Figure 4.8, page 576.

This paper bridges this gap by developing a granular quantitative spatial model (Kleinman, Liu & Redding 2023*a*, Kleinman, Liu & Redding 2023*b*) that accounts for the time-consistency of policymaking and its implications for urban welfare. In each time period, the policymaker chooses optimally to invest in a subset of locations, under an intertemporal budget constraint, given her expectations of the city's response to such investment. The policymaker chooses both the spatial boundaries of the investment and its time profile. The policymaker chooses a subset of investment locations to maximize welfare given households' location choices, developers' decision to build in each location, and given the future policymaker's own investment plan. Heterogeneous households move to locations based on their amenities, affected by the policymaker's investment, based on prior population density, racial composition, house prices, and future welfare. Place-based investments have long run effects as there is path dependence (Allen & Donaldson 2022): place-based investments attract households that form population clusters, yet the investment depreciates and the policymaker faces a trade-off. The policymaker who can commit anticipates the impact of future investment on current location choices; when the policymaker cannot commit, future welfare-maximizing infrastructure investment may lead to lower welfare today, and this is the time inconsistency described by Kydland & Prescott (1977) in their seminal work. The game is solved using Blanchard & Kahn's (1980) approach that provides the longitudinal dynamics of households' location choices, house prices, new construction given the policymaker's investment strategy. The policymaker's strategy is determined by choosing the sequence of spatially located investments that maximizes welfare. In the commitment case, the policymaker chooses a spatial design and a time profile that maximizes welfare at $t = 0$; in the no-commitment, time consistent case, the policymaker chooses such time profile to maximize each time period's aggregate welfare.

This paper applies this approach to the long-run effects of the largest set of natural experiments on flood control infrastructure in the United States. This was sparked by the 1936 Flood Control Act, signed into law on June 22, 1936. This Act appropriated funds for the US Army Corps of Engineers to build embankments and floodwalls protecting 5% of US land in the three decades following the Act. While it was followed by Acts in 1938, 1946, 1954, 1974, and 1986, this initial 1936 Flood Control Act led to the largest expansion of floodwalls and embankments in the history of the nation.[2] [3] This paper builds a longitudinal panel of Census tracts over eight decades from 1940 onwards,[4] and assesses the long-run general equilibrium impacts of such Act. Such general equilibrium dynamics affect neighborhoods' population density, demographic composition, the price of housing, and the evolution of flood risk taking. The paper combines a regression disconti-

[2]This was a significant legal reinterpretation of the scope of the federal government's ability to fund local flood control projects (Arnold 1988), as previous acts funded projects facilitating navigation.

[3]The 1936 Act follows the 1927 Mississippi flood (Hornbeck & Naidu 2014), after which private flood insurance all but dried up, and predates the 1968 National Flood Insurance Program by 32 years (Knowles & Kunreuther 2014). Coverage started in 1969 on a voluntary basis, and it is only with the Flood Protection Act of 1973 that properties with federally backed mortgages were required to be covered by federal flood insurance.

[4]We build Census tract relationship files for 1940–2010, 1950–2010, 1960–2010. We thus extend the 1970-2010 sample of neighborhoods with consistent boundaries of Card, Mas & Rothstein (2008) by an additional three decades.

nuity design at the boundary of leveed areas with a difference-in-differences setup (RD-DiD) where the evolution of discontinuities in prices, rents, population density and racial composition inform us about the evolution of neighborhood attractiveness over the short-run and the long-run. Evidence suggests that neighborhoods subsequently protected by a levee system had lower initial house values, higher initial shares of minorities, and lower population densities. RD-DiD results suggest that following the 1936 Act, house prices appreciated in leveed areas relative to non-leveed areas. The fraction black declined across specifications. Such gentrification is not significantly driven by the Great Migration of African Americans towards northern cities. Depending on the specifications, population density increased or declined.

We then use this longitudinal panel to estimate the model's structural parameters, and simulate counterfactual levee designs and optimal investment strategies. Estimation of the structural parameters suggests that population inflows and price appreciation are driven by path dependence after 1960. Using a contraction mapping akin to Berry, Levinsohn & Pakes (1995) on tract-level flows of movers and stayers, we identify intertemporal utilities, continuation values, and mobility costs. The city's local amenity fixed effects, preference for population density, social preferences, are estimated using multiway fixed effect regressions.

The paper's numerical approach based on the log-linearization of the dynamic equilibrium as in Blanchard & Kahn (1980), Uhlig (1999), Sims (2002), DeJong & Dave (2012), provides impulse response *maps* in response to any arbitrary flood protection investment strategy that targets a different subset of locations in each period. These impulse response maps provide the dynamic evolution of the city in response to place-based investments. Such numerical simulation of impulse response maps confirm the paper's non-structural findings: a one-period investment in a fully-depreciating levee system leads to long term price appreciation, sustained population inflows, and construction in the leveed area. Even absent subsequent flood protection, prior flood control infrastructure can change neighborhood dynamics for more than five decades. The impulse response maps thus provide a quantitative characterization of a form of Spatial Peltzman (1975) Effect. While the place-based investment leads to immediate welfare gains, this turns to subsequent welfare losses as the system depreciates, floor surface prices offset the benefit of the levee system, and the levee system leads to lower average citywide population density.

The paper then solves for the policymaker's optimal levee design: the spatial design of the levee and the timing of the place-based investment. We provide a simple, low-dimensional approach to parameterization of the time-varying design. The USACE decision is an implementation problem (Maskin & Sjöström 2002) akin to the choice of a Heckman (1979) selection equation that defines the boundaries of the leveed area anticipating subsequent location choices and price change in each neighborhood. The estimation of a dynamic quantitative spatial model with endogenous place-based policies enables us to model the dynamics of *counterfactual cities* where flood control infrastructure is positioned to maximize intertemporal welfare at given flood risk exposure. In the extreme case where the supply of floor surface is perfectly inelastic, flood control infrastructure leads to increases in prices and rents, the resorting of households without corresponding increases in population density. This mitigates the benefits of agglomeration economies and offsets the beneficial impacts of flood protection by fully pricing them in. In the other extreme case where the supply of floor surface is perfectly elastic, flood control infrastructure may lead to lower population density and thus declines in the benefits of agglomeration economies.

The paper starts with the optimal spatial design of a one-period investment. A welfare-maximizing design leads to higher welfare gains than the observed system, as it prioritizes the flood protection of areas with higher population density and causes less capitalization of flood protection into house prices. Applied to Saint Louis, Missouri, the optimal design protects areas close to the downtown core, rather than expanding the supply of housing in areas more distant from the Central Business District (CBD). Investment in embankments protecting lower density areas causes urban sprawl and welfare losses given the mechanism of Glaeser & Gyourko (2005) featured in the model.

The paper then solves for the policymaker's optimal spatial design and time profile of investment. In each period (corresponding to each decade in the data), the policymaker chooses a subset of locations for investment. The numerical simulation based on structural estimates reveals the key role of commitment: a policymaker who can commit builds a levee system but lets the system gradually depreciate. With commitment, the policymaker anticipates that announcements to maintain and upgrade the levee system in future periods will induce more population flows into the floodplain. The policymaker who cannot commit tends to build a substantially larger levee system. This leads to substantially (3 to 4 times) larger inflows of population into the floodplain and greater price appreciation in the leveed area. Households move closer to streams and rivers. This highlights the key role of beliefs and policy announcements when building flood protection and other types of infrastructure, including bridges, roads, and public transport: as private decisions and public investment are complementary, and as decisions have long term consequences due to path dependence, this ties the urban economics literature to the important literature in macroeconomics on policy announcements.

This paper contributes to at least four literatures. This paper designs and estimates a game between a welfare-maximizing policymaker and households in dynamic Quantitative Spatial Model (Ahlfeldt, Redding, Sturm & Wolf 2015, Redding & Rossi-Hansberg 2017, Heblich, Redding & Sturm 2020, Kleinman et al. 2023*a*). An early literature has highlighted the importance of the general equilibrium responses of place-based investments (Sieg, Smith, Banzhaf & Walsh 2000) and this paper designs welfare-maximizing time-varying and spatially heterogeneous policies that take such dynamic heterogeneous general equilibrium responses into account. The QSM's results yield impulse response maps. These maps are estimated using the workhorse methods of DeJong & Dave (2012) and Kleinman et al. (2023*b*),[5] allowing for changes in heterogeneous demographic composition, housing supply, and allowing for a dynamic of long-lasting construction. The novelty here is that the impulse response approach is used to endogenize the policymaker's optimal investment, by choosing a sequence of place-based shocks under an intertemporal budget constraint. Hence the paper's counterfactual-city approach allows us to simulate the shape

[5]We thank Chetan Dave for his suggestions at preliminary stages of this paper.

of cities with different flood control infrastructure designs; and allows us to include households' responses to, and beliefs about, the policymaker's strategy. This is a dynamic game (Osborne & Rubinstein 1994) in which households behave according to a quantitative spatial model, and where the policymaker finds investments to maximize a sequence of welfare gains. This paper makes the flood control literature meet the literature on the optimal design of infrastructure (Fajgelbaum & Schaal 2020). By modeling time-varying amenity investments by a welfare-maximizing policymaker, we provide strategic mechanisms and structural methods for the time-varying process of endogenous amenities estimated in Almagro & Domínguez-Iino (2024), allowing for counterfactual analysis.

While this paper focuses on flood control infrastructure, the findings are common to a range of issues in the optimal provision of public goods. For instance, the welfare-maximizing strategy for levee design prioritizes dense urban cores, which are typical policy prescriptions when analyzing investments in public transit, when designing urban plans, and when investing in long term infrastructure. A policymaker that can commit is likely to plan a city with less urban sprawl, internalizing the impact of today's decisions on future incentives to choose locations and transportation modes.

An exciting and growing literature studies the unintended consequences of flood control infrastructure. Hsiao (2023) documents the moral hazard caused by Jakarta's flood control infrastructure in the face of rising sea levels, with high resolution data. The current paper, focused on the United States, builds on this literature by presenting a counterfactual general equilibrium view, where parameters are estimated with data over seven decades. Thus this paper's model could be used to simulate "counterfactual Jakartas" under different flood protection policies. There is a similar number of years between the 1936 Flood Control Act and 2020 as there is between now and 2100, the time horizon of NOAA's sea level rise forecasts. The lessons of this great experiment of U.S. history, protecting 5% of US land, inform us about the likely path of neighborhood demographics over the timeline of the IPCC's forecast. Fan (2024) estimates the impact of levee reliability and maintenance, affecting income inequality. Our work shows how strategic investments can use such estimates to maximize long-term welfare and minimize unintended effects like urban sprawl. A key mechanism is the complementarity of public infrastructure and private investment choices, a mechanism illustrated in the theory of Kousky, Luttmer & Zeckhauser (2006). Henkel, Kwon & Magontier (2022) uses the exogenous relative timing of hurricanes and presidential elections to get exogenous variation in post-disaster funding and its unintended consequences.

This paper also contributes to the literature on the hedonic pricing of flood risk. Using a careful design with microdata and regression discontinuity designs, Bradt & Aldy (2022) estimates the impact of contemporary USACE infrastructure investments on house prices and finds a statistically and economically significant impact of such investments. Using contemporary data has significant benefits, as it enables a granular identification at the boundaries using deeds transaction data. This paper contributes to this literature by estimating and simulating dynamic general equilibrium effects. Such dynamic general equilibrium effects are key to policy design as they enable a simulation of the aggregate welfare impacts of flood control infrastructure, accounting for neighborhood gentrification, spatial Peltzman effects, and price appreciation.

By documenting the impact of flood control infrastructure on the racial demographics of protected neighborhoods, the paper suggests that neighborhood tipping (Schelling 1971, Card et al. 2008, Guerrieri, Hartley & Hurst 2013) may be induced by public investments. As the USACE builds embankments and floodwalls, areas protected by such infrastructure become more white and black households move towards other areas, including areas on higher elevation and lower flood risk. These gentrification effects of public infrastructure have been documented in the case of transportation infrastructure investments (Kahn 2007, Zheng & Kahn 2013) as households respond to lower mobility costs by relocating towards such infrastructure, bidding prices up and inducing a resorting by income or race. The dynamic QSM built in this paper features household heterogeneity and social preferences, enabling predictions about the evolution of population distribution across neighborhoods within metropolitan areas.[6]

The paper is structured as follows. Section 2 presents the dynamic quantitative spatial model with endogenous dynamic place based investment. The solution method to the dynamic game is presented by finding households' best responses using log-linearization techniques as in Blanchard & Kahn (1980). The section presents the set of first-order conditions when the policymaker can commit and when the policymaker is time consistent. This is then applied to the case of a place-based investment, levee systems built by the US Army Corps. Section 3 presents the paper's digitization of the Flood Control Act of 1936 and the subsequent acts, and estimates their impact on miles of embankments and floodwalls, with an econometric approach aiming at teasing out windfall effects from causal impacts. The 1936 Act is the key piece of legislation that spurred the construction of today's levee systems. Section 4 builds a Census tract panel from 1940 to 2010, extending the 1970-2010 panel of Card et al. (2008) to the 1940-1960 period, and using the US Army Corps's delineated leveed areas with hydrological flood risk measures. It estimates the impact of levees on discontinuities in house prices, rents, demographics, and population density. Section 5 uses this longitudinal panel to estimate the structural parameters. The structural analysis provides new insights: path dependence is key in explaining population flows post 1950. It estimates the impulse response maps due to a one-time fully depreciating levee investment. Section 5.5 presents the optimal design of the levee system, in the case of St Louis, Missouri, from 1940 to 2010. Section 5.6 presents the optimal time profile of levee investment, showing that a time consistent policymaker who can commit builds a smaller system. Section 6.1 discusses the hydrological literature on the long-run effectiveness of levee systems.

2 Optimal Place-Based Investments: General Equilibrium Effects and Time Consistency

This paper's model aims at answering two key questions:

[6] See also the recent literature on non-homothetic utilities (Couture, Gaubert, Handbury & Hurst 2019, Handbury 2021).

1. What is the **welfare-maximizing spatial design** of the place-based investment, e.g. embankments and floodwalls, that can achieve welfare gains, accounting for the dynamic general equilibrium response of populations and developers?

2. What is the **welfare-maximizing timing** of place-based investments, both for new infrastructure and for the maintenance or upgrade of existing infrastructure? Does the optimal timing of investment depend on the ability of the policymaker to commit to a plan at $t = 0$?

The next sections present a model based on the quantitative spatial modeling literature but where the policymaker's investment decisions are endogenized, and part of the state vector.[7] This paper's quantitative spatial model features heterogeneous locations, path dependence as in (Allen & Donaldson 2020, Allen & Donaldson 2022) due to the preference for prior population density and due to mobility costs (Krugman 1991), preferences for neighbors' demographics generating heterogeneous distributions of demographic characteristics (Kahn 2007), L-shaped housing supply with slowly depreciating housing (Glaeser & Gyourko 2005). Path dependence suggests that even one-period fully depreciating place-based investments have long term impacts on population distribution and prices. As flood risk rises, households face a trade-off between the current preference for locations, the attractiveness of cheap housing stimulated by past protection, and the current risk of flooding. A key ingredient of our model is that, in contrast to self-protection efforts (Ehrlich & Becker 1972), funding and planning decisions for the place-based investment are centralized at the policymaker's level and then households resort, developers build, in a way similar to the approach of the US Congress and the US Army Corps of Engineers.[8]

2.1 The Spatial Economy

The city has $j = 1, 2, \ldots, J$ locations and a population of households with measure N choosing across these locations in each time period $t = 1, 2, 3, \ldots$.

The initial distribution of population in the city is $\{L^g_{j0}\}_{j=1,2,\ldots,J}$. The population is heterogeneous with a discrete number $g = 1, 2$ of demographic groups. A location is characterized by its amenity A_{jt}, the price q_{jt} of its floor surface, prior population level L_{jt-1}, the share x_{jt} of group $g = 1$, and the continuation value, i.e. next period's welfare Π^g_{jt+1}. Welfare in the next period depends on the location in the previous period due to mobility costs μ_j. Households have rational expectations about the city's future population and price distributions as well as the policymaker's investment decisions.[9]

[7] This is related to the macroeconomics literature, e.g. Woodford (2003) where the optimal policy rule is endogenized. In Woodford (2003) the policymaker minimizes the present discounted value of a loss function. Here we maximize households' welfare when unobservables have a Fréchet distribution.

[8] Federal funding for levees follows series of failures of levee boards as described by Arnold (1988).

[9] This is akin to the assumption of macroeconomic models (Sargent & Vilmunen 2013). Such assumption can be relaxed by, e.g., using adaptive expectations as in Evans, Honkapohja & Mitra (2022). Here this would be about flood protection investments. This can account for the fact that, as in Fairweather, Kahn, Metcalfe & Olascoaga (2024), households have imperfect information about flood risk.

Household Utility, Consumption, and Housing Consumption

The household chooses a location j, a consumption of floor surface h_{jt}, a consumption c_{jt} of the numeraire that maximizes its utility:

$$V_{jt}^g \equiv \max_{c_{jt},h_{jt}} \quad (c_{jt}^\sigma + A_{jt}^g h_{jt}^\sigma)^{\alpha/\sigma} L_{jt-1}^\xi x_{jt-1}^{\gamma^g} \Pi_{jt+1}^\delta z(\varepsilon_{jt}),$$
$$s.t. \quad c_{jt} + q_{jt} h_{jt} \leq y_{jt}^g \tag{1}$$

where $z(\varepsilon_{jt})$ is a Fréchet distributed unobservable preference parameter with dispersion θ. u_{jt}^g is the one period flow utility. ξ is the preference for population density. δ is the time discount factor. Denote by $V(A_{jt}^g, q_{jt}, y_{jt}^g, L_{jt-1}, x_{jt-1}, \Pi_{jt+1}^g) = V_{jt}^g$ the indirect utility of the household when choosing location (j,t). The $(c_{jt}^\sigma + A_{jt}^g h_{jt}^\sigma)^{\alpha/\sigma}$ allows for an endogenous response of the demand h_{jt} for housing, which can be complement or substitute w.r.t. local amenities. Denote by U_{jt}^g the deterministic part of V_{jt}^g.

Welfare

The equilibrium is characterized by an aggregate welfare, derived from revealed preferences arguments. The welfare of a resident of demographic group g in j depends on the utility levels in each location k and on the mobility costs μ:

$$\Pi_{jt}^g = \Gamma \left\{ \sum_{k=1}^J \left(\frac{U_{kt}^g}{\mu_j^{\mathbf{1}(j \neq k)}} \right)^\theta \right\}^{\frac{1}{\theta}}, \tag{2}$$

which makes it clear that mobility costs $\{\mu_j\}$ are a mechanism of spatial frictions, as residents may not move towards the utility-maximizing location in the next period t.

The city's aggregate welfare at t follows the usual approach with Fréchet residuals.

$$\Pi_t = \sum_g \Pi_t^g, \qquad \Pi_t^g = \Gamma \sum_{j=1}^J L_{jt}^g \left\{ \sum_{k=1}^J \left(\frac{U_{kt}^g}{\mu_j^{\mathbf{1}(j \neq k)}} \right)^\theta \right\}^{\frac{1}{\theta}}, \tag{3}$$

This defines a dynamic sequence of welfares $(\Pi_t^1, \Pi_t^2)_{t=1,2,\ldots,T} \in \mathbb{R}^{2T}$.

Dynamic Location Choices

Households choose the location that maximizes $U_{kt}/\mu_j^{\mathbf{1}(j \neq k)}$ when starting in location j.

$$j_g^*(k,t) = \text{argmax}_j \left\{ \frac{U_{jt}^g}{\mu_k^{\mathbf{1}(j \neq k)}} \right\}. \tag{4}$$

This pins down population flows between t and $t+1$ towards each location j:

$$L^g_{jt+1} = \sum_{k=1}^{J} L^g_{kt} \frac{\left(U^g_{jt}/\mu_k^{\mathbf{1}(j\neq k)}\right)^\theta}{\sum_{l=1}^J \left(U^g_{lt}/\mu_k^{\mathbf{1}(l\neq k)}\right)^\theta} \tag{5}$$

The Bellman equation of the household in location j pins down the relationship between her welfare in t and her welfare in $t+1$.

$$\Pi^g_{jt} = \Gamma \left\{ \sum_{k=1}^{J} \left(u_{kt}\Pi^\delta_{kt+1}/\mu_j\right)^\theta \right\}^{\frac{1}{\theta}} \tag{6}$$

Housing Supply

Housing supply is L-shaped in each location, as in Glaeser & Gyourko (2005), which can potentially generate hysteresis in prices and housing supply. Developers supply housing competitively with elasticity η. Such housing depreciates at rate ζ. Housing supply in $t+1$ only exceeds the depreciated total stock of t if demand L_{jt} exceeds such depreciated housing stock $(1-\zeta)H_{jt}$:

$$H_{jt+1} = \max\left\{(1-\zeta)H_{jt}, D_{jt}\right\}, \tag{7}$$

where D_{jt} is the demand for housing coming from all groups $D_{jt} = \sum_{g=1}^G h^g_{jt} L^g_{jt}$.

The price of floor surface is thus either the marginal cost mc_{jt} or the market-clearing price.

$$\begin{cases} q_{jt} = \frac{1}{s_j} D_{jt}^{1/\eta_j}, & \text{if } D_{jt} > (1-\delta)H_{jt-1} \\ q_{jt} = mc_{jt} \text{ otherwise} \end{cases} \tag{8}$$

The equilibrium concept that this yields is similar to the established quantitative spatial literature, with the addition of a depreciating housing stock.

A Sequence of Place-Based Investments: Spatial Designs and Timeline

In each period t, the policymaker chooses to invest in a subset of locations of the city to maximize aggregate welfare.[10] The place-based investment multiplies amenity values (A_{jt}) for each location $j \in \mathcal{J}$ so that:

$$A_{jt}(I_{jt}) = (1 + \zeta_{jt} \cdot I_{jt})A_{jt} \tag{9}$$

we denote $\tilde{A}_{jt} \equiv \zeta_{jt} \cdot I_{jt}$.[11]

We represent the set of possible place-based investments as a sequence in each period t of subsets $\mathcal{J}_1, \mathcal{J}_2, \ldots$ of locations $\{1, 2, \ldots, J\}$ of the city. A subset is parameterized by a selection

[10] The difference between the commitment and time consistent cases is described below.

[11] Throughout the paper we use the convention of the literature of denoting log-linearized deviations from the steady state by a tilde.

equation that determines the probability that the policymaker invests in location j at time t:

$$P(I_{jt}=1)=\Lambda(\mathbf{x}_{jt}\boldsymbol{\beta})\in(0,1) \tag{10}$$

$\mathbf{x}_{jt}$ can include measures such as flood risk as well as measures of household income, existing housing supply, house values. The coefficients $\boldsymbol{\beta}_t \in \mathcal{B}$ are constrained to reflect technical requirements, such as higher probabilities along the floodplain. The covariates $\mathbf{x}_{jt}\boldsymbol{\beta}_t$ are time-varying to reflect the dynamic nature of the place-based investment: maintenance and reinvestment. The sequence of coefficients $\boldsymbol{\beta}_t$ may be chosen by the policymaker at different time periods (see definitions 3 and 4 below).[12]

With this parameterization, the probability that the place-based investment benefits the set $\mathcal{J}_t$ of locations at t is:

$$P(\mathcal{J}_t)=\prod_{j=1}^{J}P(I_{jt}=1)^{\mathbf{1}(j\in\mathcal{J}_t)}(1-P(I_{jt}=1))^{1-\mathbf{1}(j\in\mathcal{J}_t)} \tag{11}$$

The policymaker's budget constraint is that the total intertemporal cost of protecting locations $j=1,2,\ldots,J$ at each $t=1,2,\ldots$ is less than the budget C:

$$\sum_{t=1}^{\infty}\frac{1}{(1+r)^t}\sum_{j=1}^{J}P(I_{jt}=1)c_{jt}\leq C \tag{12}$$

where $\sum_{t=1}^{\infty}\sum_{j=1}^{J}\frac{1}{(1+r)^t}c_{jt}<\infty$. r is the policymaker's time discount factor.

Equilibrium concept

Households make location and consumption choices based on current and future investments, while the policymaker chooses an investment path based on households' location choices.

Definition 1. (Dynamic Equilibrium Path with Policymaker's Commitment) *Given an initial distribution of population (L_{j0}), an initial stock of housing (H_{j0}) across locations, an equilibrium is a sequence $(L^g_{jt}, q_{jt}, c^g_{jt}, h^g_{jt}, P^g_{jkt}, \Pi^g_{jt}, H_{jt}, I_{jt})$ of population numbers L^g_{jt}, floor surface prices q_{jt}, numeraire consumption c^g_{jt}, per capita floor surface consumption h^g_{jt}, probabilities of mobility P^g_{jkt}, continuation values Π^g_{jt}, housing stock H_{jt} and place-based investments I_{jt} in each location for each $j=1,2,\ldots,J$, $t=1,2,\ldots$, and $g=1,2$, such that:*

1. *Population L^g_{jt+1} evolves according to the number of households of group g moving to location j. $L^g_{jt+1}=\sum_{i=1}^{J}L^g_{it}P^g_{ijt}$, equation (5), and probabilities of mobility are pinned down by utilities and mobility costs.*

[12] In particular, in the case with commitment, the policymaker chooses the sequence of vectors $\boldsymbol{\beta}_t$ at $t=0$. In the time consistent case, the policymaker chooses each vector $\boldsymbol{\beta}_t$ at each time $t>0$. This gradually alters the design of the place-based investment.

2. *Floor surface prices q_{jt} evolve according to the elasticity of supply η_j when the demand for floor surface exceeds the existing stock, equation 8.*

3. *Conditional on a location choice j, households of group g choose consumption c^g_{jt} and per capita floor surface consumption h^g_{jt} optimally to maximize (1).*

4. *The continuation values Π^g_{jt} for $j = 1, 2, \ldots, J$ and $t = 1, 2, \ldots$ satisfy the Bellman equation, equation (6).*

5. *Housing stock H_{jt} depreciates at rate ζ, and the stock increases when demand exceeds the stock $(1-\zeta)H_{jt}$ (equation (7)).*

6. *The policymaker chooses a sequence of place-based designs $\boldsymbol{\beta}_{jt}$ such that, at each time period t, she invests $I_{jt} = 1$ in location j with probability $P(I_{jt} = 1) = \Lambda(\mathbf{x}_{jt}\boldsymbol{\beta}_t)$. The sequence $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots\}$ is chosen at $t = 0$ to maximize households' aggregate welfare at $t = 0$ given the intertemporal budget constraint (12).*

When the policymaker does not or cannot commit, the policymaker reoptimizes at each t.

Definition 2. (Dynamic Equilibrium Path without Policymaker's Commitment) *The dynamic equilibrium without commitment is identical to the case with commitment, except that the policymaker's behavior is as follows:*

6. *The policymaker chooses a sequence of place-based designs $\boldsymbol{\beta}_{jt}$ such that, at each time period t, she invests $I_{jt} = 1$ in location j with probability $P(I_{jt} = 1) = \Lambda(\mathbf{x}_{jt}\boldsymbol{\beta}_t)$. The vector $\boldsymbol{\beta}_t$ is chosen at each t to maximize households' aggregate welfare at $t > 0$ given the intertemporal budget constraint (12).*

We solve both equilibria in the next sections.

2.2 The Policymaker's Place-Based Investment Strategy

2.2.1 Optimal Design with Commitment

The goal is to maximize the aggregate welfare at $t = 0$ given an expected sequence $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_T\}$ of place-based investments affecting the subset $\mathcal{J}_t$ of locations at time t:

$$\Pi_{t=0}(\boldsymbol{\beta}) = \sum_{g=1}^{G}\sum_{j=1}^{J} L^g_{j,-1}\Pi^g_{jt=0}(\boldsymbol{\beta}), \tag{13}$$

which is the weighted average of the welfare of the demographic groups, given the initial population distribution $L^w_{j,-1}$, $L^b_{j,-1}$ at $t = -1$.

Definition 3. (Policymaker's Investment Strategy with Commitment) *At time $t = 0$, the policymaker finds the sequence of vectors $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots\}$ that maximizes the city's aggregate welfare gain at $t = 0$ given the intertemporal budget constraint:*

$$\max_{\{\boldsymbol{\beta}_t\}} \quad \Pi_{t=0}(\{\boldsymbol{\beta}_t\}_{t>0}) \tag{14}$$

$$\text{s.t.} \quad \sum_{t=1}^{\infty} \frac{1}{(1+r)^t} \sum_{j=1}^{J} P(I_{jt} = 1) c_{jt} \leq C \tag{15}$$

$$\text{and} \quad P(I_{jt} = 1) = \Lambda(\mathbf{x}_{jt} \boldsymbol{\beta}_t) \tag{16}$$

This optimal selection framework is performed in general equilibrium. For each potential design $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots\}$, we simulate the path of prices, populations, housing supply over the seven decades.

The general equilibrium impact functions through a shift in the distribution of prices q_{jt} (in the short run, appreciation due to the improvement in amenities, and in the long-run either a decline to the expansion of supply, or a continued increase due to agglomeration economies or social preferences), a change in population distributions, a change in the demand for floor surface (urban sprawl vs increases in density).

The optimal sequence $\boldsymbol{\beta}^*$ satisfies the sequence of first-order conditions for each time period:

$$\underbrace{\frac{d\Pi_0}{d\boldsymbol{\beta}_t}(\boldsymbol{\beta}^*)}_{\text{Marginal Impact on Aggregate Welfare in General Equilibrium}} = \underbrace{\lambda \sum_j c_{jt} \left. \frac{dP(I_{j0} = 1)}{d\boldsymbol{\beta}_t} \right|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{**}}}_{\text{Marginal Impact on Cost}}, \qquad t = 1, 2, \dots \tag{17}$$

The solution with commitment accounts for the response of households' location choices in periods t to the expectation of future policies in periods $t + k$. This is the point made by Kydland & Prescott (1977).

2.2.2 Without Commitment: The Time Consistent Case

The previous program is not time consistent as it does not maximize welfare at each period t.

Definition 4. (Policymaker's Investment Strategy without Commitment) *At each time t, the policymaker chooses a vector $\boldsymbol{\beta}_t$ that maximizes the city's aggregate welfare gain at $t > 0$ given the intertemporal budget constraint. For each t,*

$$\max_{\boldsymbol{\beta}_t} \quad \Pi_t(\boldsymbol{\beta}_t, \{\boldsymbol{\beta}_{t+k}\}_{k>0}) \tag{18}$$

$$\text{s.t.} \quad \sum_{t=1}^{\infty} \frac{1}{(1+r)^t} \sum_{j=1}^{J} P(I_{jt} = 1) c_{jt} \leq C \tag{19}$$

$$\text{and} \quad P(I_{jt} = 1) = \Lambda(\mathbf{x}_{jt} \boldsymbol{\beta}_t) \tag{20}$$

In the maximization program, the policymaker takes her future self's choice $\{\boldsymbol{\beta}_{t+k}\}_{k>0}$ as given.Her future self at $t+k$ will maximize Π_{t+k}. The budget constraint is intertemporal: the policymaker at t is constrained by the budget used by the previous selves at $t-k$.

The time consistent optimal sequence $\boldsymbol{\beta}^{**}$ satisfies the sequence of first-order conditions:

$$\underbrace{\frac{d\Pi_t}{d\boldsymbol{\beta}_t}(\boldsymbol{\beta}^{**})}_{\text{Marginal Impact on Aggregate Welfare in General Equilibrium}} = \underbrace{\lambda \sum_j c_{jt} \left.\frac{dP(I_{j0}=1)}{d\boldsymbol{\beta}_t}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{**}}}_{\text{Marginal Impact on Cost}}, \qquad t = 1, 2, ... \tag{21}$$

As in Kydland & Prescott's (1977) example, this strategy will not typically maximize households' welfare as the policymaker does not account for the impact of future investments on households' past location choices.

2.2.3 Solution with a Sequence of Place-Based Investments

This section provides a simple and computationally efficient way to solve this game. The key issue tackled by this Blanchard & Kahn (1980) approach is that population flows, prices and welfare levels depend on expectations of their future path.

Denote by $\mathbf{z}_t$ the state vector $(q_{jt}, L^w_{jt}, L^b_{jt}, h^w_{jt}, h^b_{jt}, \Pi^w_{jt}, \Pi^b_{jt}, H_{jt})$ at t (see Definitions 1 and 2). We solve for $\mathbf{z}_t$ given the policymaker's investment strategy $(I_{jt})_{t=1,2,...,T}$ across locations $j = 1, 2, \ldots, J$.

The city's dynamic equilibrium conditions 1–6 can be summarized by a single function $\boldsymbol{\psi}$:

$$\boldsymbol{\psi}(\mathbf{z}_{t+1}, \mathbf{z}_t, \mathbf{z}_{t-1}, \mathbf{A}_t) = 0 \tag{22}$$

and denote by $\mathbf{z}^*$ the steady-state vector of tract-level values and by $\mathbf{A}^*$ the vector of exogenous amenities without place-based investments. The vector-valued function $\boldsymbol{\psi}$ takes in parameters from $(J+2JG) \times (J+2JG) \times J$ and is a set of $J+2JG$ relationships. Denote by $\tilde{z}_t$ log values of the difference between the value at t and the steady state:

$$\tilde{\mathbf{z}}_t = \log \mathbf{z}_t - \log \mathbf{z}^* \tag{23}$$

and by $\tilde{\mathbf{A}}_t = (\zeta_{jt} I_{jt})_{j=1,2,...,J}$ the vector of place-based investments. Taking the log-linearized version of (22),[13] we obtain the multivariate linear relationship:

$$H\tilde{\mathbf{z}}_{t+1} = F\tilde{\mathbf{z}}_t + G\tilde{\mathbf{z}}_{t-1} + E\tilde{\mathbf{A}}_t \tag{24}$$

Equation (22) is a relationship with two lags. This is rewritten as an equation with one lag by stacking $\tilde{\mathbf{w}}_{t+1} = (\tilde{\mathbf{z}}_{t+1}\, \tilde{\mathbf{z}}_t)'$, and writing $B\tilde{\mathbf{w}}_{t+1} = C\tilde{\mathbf{w}}_t + E\tilde{\mathbf{A}}_t$. Denote by $\tilde{\mathbf{z}}_t = P\tilde{\mathbf{w}}_t$ the projection of $\tilde{\mathbf{w}}_t$ onto $\tilde{\mathbf{z}}_t$.

[13] The kink in the L-shaped housing supply function means that such supply is not differentiable when the demand for housing is exactly equal to the depreciated stock, $D_{jt} = (1-\zeta)H_{jt}$. We address this using a common technique, the smooth maximum, see e.g Biswas, Kumar, Banerjee & Pandey (2021).

The typical issue in these models is that endogenous variables, e.g. prices, depend on the forward-looking sequence $\{\mathbf{z}_{t+k}\}_{k>0}$. Following DeJong & Dave (2012), we perform a Generalized Schur decomposition of B, separating the explosive and non-explosive parts of the model, using the position of the eigenvalues relative to 1. We then solve for the explosive part by forward induction. In simpler terms, we find linear combinations of prices and quantities such that each variable can be expressed as the present value of this linear combination.

The impact of the investment $\tilde{\mathbf{A}}_t$ in amenities is then simulated. The technique is related to Kleinman et al.'s (2023*b*) approach. The impulse response thus generated are *maps* that portray the evolution of the city's neighborhood prices, demographics, housing supply, and flood risk exposure in response to shocks to value of locating in a subset j of locations.

We solve for the dynamics of population, prices, quantities and welfare in each decade t in response to a shock to each location's amenities. We use an approach close to Blanchard & Kahn (1980), for now in an environment of perfect foresight of the levee investment's fading impact. This generates more conservative predictions about the levee's welfare impact, as households foresee that the levee depreciates. A further discussion of this assumption about beliefs will be presented in the next version of the paper.

We start with Equation (24):

$$B\tilde{\mathbf{w}}_{t+1} = C\tilde{\mathbf{w}}_t + E\tilde{\mathbf{A}}_t \tag{25}$$

The key issue here is that some variables depend on expectation of future prices, welfares, population levels, racial composition, and continuation values.

Perform the generalized Schur (QZ) decomposition of B and C:

$$\begin{cases} B & = Q'\Lambda Z' \\ C & = Q'\Omega Z' \end{cases} \tag{26}$$

where $QQ' = \text{Id}$ and $ZZ' = \text{Id}$. Multiplying (25) by Q, we get:

$$\Lambda Z'\mathbf{w}_{t+1} = \Omega Z'\mathbf{w}_t + QE\tilde{\mathbf{A}}_t \tag{27}$$

and define $\mathbf{y}_t = Z'\mathbf{w}_t$. Then partition Λ and Ω into a part with eigenvalues above 1 (explosive part) and a part with eigenvalues below 1. This is:

$$\begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ 0 & \Lambda_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y}_{1t} \\ \mathbf{y}_{2t} \end{bmatrix} = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ 0 & \Omega_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y}_{1t-1} \\ \mathbf{y}_{2t-1} \end{bmatrix} + \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} E\tilde{\mathbf{A}}_t \tag{28}$$

The lower block can then be solved forward as $\mathbf{y}_{2t} = (\Omega_{22})^{-1}\Lambda_{22}\mathbf{y}_{2t+1} - (\Omega_{22})^{-1}Q_2E\tilde{\mathbf{A}}_{t+1}$. This lower part $\mathbf{y}_{2t}$ is the present discounted value of the impact of future place-based investments. The upper part $\mathbf{y}_{1t}$ is the iteration of the autoregressive identity that depends on $\mathbf{y}_{1t-1}$, $\mathbf{y}_{2t}$, and

$\mathbf{y}_{2t-1}$. We get the general solution:

$$\mathbf{y}_{2t} = -\sum_{i=0}^{\infty} M^i (\Omega_{22})^{-1} E \tilde{\mathbf{A}}_{t+1+i} \tag{29}$$

$$\mathbf{y}_{1t} = (\Lambda_{11})^{-1}\Omega_{11}\mathbf{y}_{1t-1} - (\Lambda_{11})^{-1}\Lambda_{12}\mathbf{y}_{2t} + (\Lambda_{11})^{-1}\Omega_{12}\mathbf{y}_{2t-1} + (\Lambda_{11})^{-1}Q_1 E\tilde{\mathbf{A}}_t \tag{30}$$

This allows the urban economist to simulate the response of the city to a sequence of amenity improvements $\left\{\tilde{\mathbf{A}}_t\right\}_{t>0}$ due to place-based investments I_{jt} that increase $\tilde{A}_{jt} = \zeta_{jt} I_{jt}$.

Given $\mathbf{w}_0 = Z\mathbf{y}_0$ and that welfare gains $\tilde{\Pi}_{t=0}$ are a linear combination of elements of $\mathbf{w}_0 = (\tilde{\mathbf{z}}_0\, 0)'$ (see definition of $\tilde{\mathbf{z}}_t$ at the beginning of this section), this proves the proposition:

Proposition 5. (Welfare Gain and Sequence of Place-Based Investments) *The welfare change can be written in linear form as a function of the sequence of J amenity increases $\tilde{\mathbf{A}}_t = \left(\tilde{A}_{jt}\right)$.*

$$\tilde{\Pi}_{t=0} = \underbrace{\mathbf{L}'_{-1}}_{\substack{\textit{Initial}\\ \textit{population}\\ \textit{distribution}}} \;\Theta\; \underbrace{\left(\tilde{\mathbf{A}}_1\, \tilde{\mathbf{A}}_2\, \ldots\, \tilde{\mathbf{A}}_T\right)}_{\substack{\textit{Change in}\\ \textit{Amenity due to sequence of}\\ \textit{place-based investments } I_{jt}}} \quad . \tag{31}$$

where $\mathbf{L}_{-1}$ is the vector of initial population distribution, and Θ is a square matrix of size J that converts a sequence of investments into a vector of welfare gains from each locations. The parameter λ is the welfare cost of the budget constraint. The optimal coefficient $\boldsymbol{\beta}^*$ internalizes the general equilibrium impact of flood protection on prices, demographics, and the demand for land.

The sequence of place-based investments is driven by the impact of investments on amenities times the probability of investing in each location:

$$\left(\tilde{\mathbf{A}}_1\, \tilde{\mathbf{A}}_2\, \ldots\, \tilde{\mathbf{A}}_T\right) = \begin{bmatrix} \zeta_{11} P(1 \in \mathcal{J}_1) & \cdots & \zeta_{1T} P(1 \in \mathcal{J}_T) \\ \vdots & & \vdots \\ \zeta_{J1} P(J \in \mathcal{J}_1) & \cdots & \zeta_{JT} P(1 \in \mathcal{J}_T) \end{bmatrix}. \tag{32}$$

This linear relationship makes the choice of optimal place-based investment policies computationally feasible.

The Simple Case of the One-Period Investment

A simple case is that of the one-period investment. When there is a single one-time spatially heterogeneous shock $\tilde{\mathbf{A}}_1$, the impact at $t = 0$ on the lower part of $\mathbf{y}_{20}$ is:

$$\mathbf{y}_{20} = -M(\Omega_{22})^{-1} Q_2 E \tilde{\mathbf{A}}_1, \tag{33}$$

given that the infrastructure is a one time investment. The non-explosive part $\mathbf{y}_{10}$ of (28) is $\mathbf{y}_{10} = -(\Lambda_{11})^{-1}(\Lambda_{12})\mathbf{y}_{20}$ at $t = 0$. We've obtained $\mathbf{y}_0 = (\mathbf{y}_{10}\, \mathbf{y}_{20})'$ and thus $\mathbf{w}_0 = Z\mathbf{y}_0$, and $\mathbf{z}_0 = P\mathbf{w}_0$, and the welfare gain at $t = 0$, denoted $\tilde{\mathbf{\Pi}}_0 = D\mathbf{z}_0$.

Corollary 6. (Welfare Gain and One-Period Investment) *When the investment lasts only one period, welfare gains can be written explicitly in linear form as:*

$$\tilde{\Pi}_{t=0} = \mathbf{L}'_{-1} \underbrace{DPZ \begin{bmatrix} -M(\Omega_{22})^{-1}Q_2E \\ (\Lambda_{11})^{-1}(\Lambda_{12})M(\Omega_{22})^{-1}Q_2E \end{bmatrix}}_{\Theta} \tilde{\mathbf{A}}_1, \tag{34}$$

This general equilibrium impact on the J locations of the city accounts for the discounted future changes in welfare, populations, prices.

The researcher who has the matrix Θ can thus compute general equilibrium changes efficiently for any set of spatially-located shocks $\tilde{\mathbf{A}}_1$, a vector of size J. This in particular allows us to simulate the impact of a large number of different one-period investments.

2.3 Application to Levee Systems

The framework of this section has broad application to a range of place-based policies that affect the spatial distribution of amenities. We apply it in this paper to the case of investments in flood protection infrastructure such as levee systems. These levee systems, through a set of embankments, floodwalls, protect a set of locations called the leveed area. The next sections design the leveed area to maximize welfare. Levees lead to general equilibrium price shifts and heterogeneous population flows. The depreciation of levees, combined with the path dependence of investments, leads to a trade-off, an optimal timeline of levee construction and maintenance that differs when the policymaker commits, and when the policymaker is time consistent.

We implement this approach using a longitudinal panel of Census tracts from 1940 to 2010. This panel, built for this paper using newly created relationship files, includes population by race, average house prices, tract-level flood risk measures, housing units, and mobility flows. Section 4 describes the data set. It shows that the Flood Control Act of 1936 led to significant increases in prices, population density, and lead to changes in demographics that persist at least until 2010. Section 5.2 estimates the model's structural parameters. Section 5 implements the structural approach of this section by finding the sequence of levee investment that maximizes welfare. Section 5.5 provides the welfare-maximizing levee design (Figure 9) by maximizing initial welfare in Corollary 6. Section 5.6 provides the timeline of optimal investment, with (resp. without commitment), by solving the first order conditions (17) (resp. (21)). The welfare-maximizing design differs substantially from the observed one.

3 The Flood Control Acts' Impact on Flood Control Infrastructure

The 1927 Mississippi Flood inundated an estimated 18,000 square miles, took 246 lives, and displaced 700,000.[14] Hornbeck & Naidu (2014) describes its impact on displaced labor and capital investment. The problem of flood control is, at the time, described as beyond the responsibility of the federal budget:

> Under the present law the United States says to the threatened ones, "No pay, no protection." [...] Some say it is not the affair of the United States Government to do this work. [...] Under the present law, and similar proposed laws, money for flood – control works must come from the levee districts along the Mississippi River. These levee districts, while authorized by State law, are in no way connected with the State. They get no State funds and they are not permitted to use the credit of the State.[15]

While the Gibbons v. Ogden, 22 US 1 (1824), Supreme Court decision did provide a legal basis for federal funding of local levee projects, this was not the interpretation until 1936.[16] Two events change this dynamic. First, the Great Depression brought about a series of legislation, including the July 1932 to 1935 Emergency Relief Acts. Hearings ignited debates about the potential appropriation of funds towards flood control projects, but the final bill did not include such projects.[17] H.R. 6803, "A Bill to Authorize Funds for the Prosecution of Works for Flood Control and Protection Against Flood Disasters" was discussed but did not pass. The second event was the significant 17–19 March 1936 floods affecting the Northeast, including Washington D.C.. It is described by the National Weather Service as one of the region's worst natural disasters.[18]

The 1936 Flood Control Act protects an estimated 100 million acres, or 4 to 5% of U.S. land. Additional Flood Control Acts include the 1938, 1946, 1954, 1976, 1986 Acts. Yet, Figure 1 suggests that the 1936 Act was a watershed moment in the history of federal flood control funding. Visual inspection of these graphs suggest a significant uptick in – if not simply the start of – the construction of embankments and floodwalls after the 1936 Act. The figures suggest that, with a few exceptions, the US Army Corps started building levees after the 1936 Flood Control Act. Such US Army Corps levees were either federally constructed and operated (a minority of the projects) or federally constructed and turned over to a public sponsor. The next subsection investigates whether this evidence can be understood as causal or whether this is simply coincidental, a windfall of public funding whereby flood control infrastructure would have been constructed regardless of such federal funding.

Figure 2 presents the geocoded location of projects described in the digitized copies of the

[14] "*Flood Control in the Mississippi Valley, Report Submitted by Hon.Frank R. Reid of Illinois, Chairman, from the Committee on Flood Control (to Accompany H.R. 8219).*" U.S House of Representatives, 70th Congress, 1st Session, Report No. 1072.

[15] Op. Cit., Preface and Chapter III, page 21.

[16] Congressional Record, House, 1936, page 7741,

[17] *Final Statistical Report of the Federal Emergency Relief Administration,* Federal Works Agency, 1942.

[18] 1936 Flood Retrospective, National Weather Service, https://www.weather.gov/lwx/1936Flood.

successive Flood Control Acts.[19] From 1936 onwards, projects extended from the West to the East coast, and including the Los Angeles and Seattle levee systems. Section 4 uses the precise boundaries of embankments, floodwalls, and leveed areas. This section investigates the potential causal impact of federal funding on levee construction.

The 1936 Act had the largest individual number of funded projects (202) until the 1986 Act, for a total of 279 million USD funded in 1936 dollars. This, adjusted for the consumer price index between June 1936 and April 2023, corresponds to a total funding of 6.1 billion USD in 2023 dollars. Further adjustment may be performed to account for the rising cost of infrastructure construction (Brooks & Liscow 2023). The median project received 210,750 USD (in 1936 dollars) or 4.6 million USD in 2023 dollars.

Table 1: Funding Appropriated by Flood Control Acts – Summary Statistics

			Funding earmarked for specific locations		
Act	# of Projects	Total, Current USD	Median, Current USD	Total, 2023 USD	Median, 2023 USD
1936	202	279,774,180	210,750	6,127,902,342	4,616,063
1938	19	130,021,000	1,608,000	2,794,663,493	34,562,254
1946	53	623,788,069	2,100,000	10,120,488,852	34,070,909
1954	163	502,745,025	436,000	5,649,025,917	4,899,054
1974	23	42,367,000	400,000	261,648,321	2,470,303
1986	215	12,820,610,000	18,400,000	35,470,354,333	50,906,666

Source: Digitization of the Congressional Record by the authors. Total is the total of the geocoded individual appropriations mentioned in each document.

The correlation between the location of projects and subsequent embankment and floodwall construction may not be reflective of a causal relationship. Indeed, federal funding could have been a windfall that did not cause the construction of levees; but rather levees might have been built in areas most at risk of flooding, regardless of the federal funding. Such windfall may have been driven by the willingness to hire workers during the Great Depression. It could also have been driven by political considerations, as the choice of specific projects might have benefitted influential congressional districts.

To address this question, we build a propensity score by collecting four sets of data, each illustrated on Appendix Figure A15. First, we collect historical data on precipitation at monthly frequency between 1895 and 1986, the last Act considered in this paper. For each Act, we compute the average, 90th percentile precipitation (inches to 100ths) by county (i) between 1896 and the Act (e.g. June 1936), and (ii) in the year preceding the Act. This 90th percentile captures, for instance, the salience of the March 1936 flood event in the months leading up to the vote of the law. The data is described in Vose, Applequist, Squires, Durre, Menne, Williams Jr, Fenimore, Gleason & Arndt (2014). Data on a century of natural disasters, including past floods and hurricanes at the county-level is from Boustan, Kahn, Rhode & Yanguas (2020).

The second set of data is the share unemployed in the decennial Census year preceding the act, from the *Historical, Demographic, Economic, and Social Data: The United States, 1790-1970*.

[19] The 1974 and 1986 Acts are called Water Resources Development Act.

This is depicted on Figure A15d. This map suggests that areas around the Mississippi river, the State of Alabama, and the Seattle area (where the 1936 Act funds projects), also experienced high unemployment. Given that the 1936 Flood Control Act was signed only a year after the New Deal's 1935 Emergency Relief Act, this may be a driver of investment. In these historical data, we consider the % unemployed as a fraction of total population.[20]

The third and fourth data sets pertain to the political economy of federal funding. For the 1936, 1938, and 1946 Acts, a Flood Control Committee of the House of Representatives examined the appropriations. For the 1954 Act, it is a Committee on Public Works; and for subsequent Acts, a Committee on Public Works and Transportation. We use digitized reports of these committees to list the members (chairman and ranking members).[21] The final source of data is the elected representatives for each congressional district. We use Jeffrey Lewis' Complete Congressional District Boundaries (Lewis, DeVine, Pitcher & Martis. 2013), based on Kenneth C. Martis' The Historical Atlas of United States Congressional Districts: 1789-1983.[22]

These four data sets are used to build a propensity score, where the left-hand side is equal to 1 when a project has been funded in the county, and the right-hand side is the covariates from these four data sets.

$$\begin{aligned}\text{Project Funded}^*_{ca} \quad = \quad & a_1\text{P90 of Precipitation from 1895 to Year } a \text{ of Flood Control Act}_{ca} \\ & +a_2\text{Frac. Unemployed}_{ca} \\ & +a_3\text{Flood Control Committee Member}_{ca} \\ & +a_4\text{House Representative}_{ca} + \varepsilon_{ca} \qquad (35)\end{aligned}$$

where c is the county, a is the year of the Flood Control Act (from 1936 to 1986). The weights are $1/p_{ca}$ for the treatment group and $1/(1-p_{ca})$ for the control group. The distribution of propensity scores in the treatment (funded) and control (no project funded) suggests reasonable overlap: a substantial number of counties did not receive funding even though the political economy, the weather, or the unemployment suggest a high probability of appropriation.

The specification regresses the miles of embankments and floodwalls on an event-study set of indicator variables, controlling for county fixed effects, and weighting for the inverse propensity score as in Hirano, Imbens & Ridder (2003), to estimate the average treatment effect. For the left-hand side, we use the asinh of the length of embankments and floodwalls (Bellemare & Wichman 2020), which allows for zeros in the dependent variable. Coefficients δ_k are then interpreted as

[20] Tables 937TOTALLY UNEMPLOYED, 930TTL MALES, and 930TTL FEMALES.

[21] For example, the "Hearings before the Committee on Flood Control, House of Representatives, Seventy-Fourth Congress, First Session, on H.R. 6803, A Bill to Authorize of Works for Flood Control and Protection against Flood Disasters," March 22 and 23, and April 2nd, 1935. https://hdl.handle.net/2027/mdp.39015067181993.

[22] The Farmer-Labor party is coded as Democrat. The Progressives are coded as Republican.

semi-elasticities.

$$asinh(\text{Length of Infrastructure}_{ajt}) = \sum_{k=-T, k\neq -1}^{T} \delta_k \mathbf{1}(\text{Treated}_{jt+k}) + \text{County}_j + \text{Year}_{y(t,a)} + \mathrm{Act}_a + \varepsilon_{ajt}, \tag{36}$$

and standard errors are double-clustered at the county and year levels. a indexes the act from 1936 to 1986, j indexes counties, and t indexes years relative to the natural disaster; y is the year. $\delta_{-1} \equiv 0$ is the reference point. The three figures illustrate the results. The upper panel considers all miles of embankments and floodwalls, and all flood control acts and simply conditions for fixed effects without weighting. These present *flows* of new construction, and thus the cumulative impact on miles today is the sum of the coefficients. The upper panel suggests that the construction peaks approximately two decades after the Act, consistent with the timeline of the US Army Corps' studies. Panel (b) presents coefficients adjusted for the inverse probability weighting as in Hirano et al. (2003). This has the benefit of making the control and the treatment groups more comparable. The point estimates are substantially larger, with wider standard errors, but the effects remain significant at 95%. The lower-left panel considers embankments only, for all acts. The effects are driven by the construction of embankments.

4 City Dynamics over 80 Years: A Neighborhood-Level Panel

This section estimates the impact of flood protection infrastructure on population flows, prices, and construction. It provides empirical evidence that levees have economically and statistically significant impacts on a city's dynamics. These observed flows play a key role in our subsequent analysis determining the optimal flood construction infrastructure with or without commitment. Structural estimation of the model's parameters is presented in Section 5.

Estimating the impact of levees on long-run neighborhood dynamics requires a longitudinal panel of neighborhoods with consistent geographic boundaries, matched to neighborhood-level flood risk measures and to leveed areas boundaries pinned down by flood control infrastructure. Section 4.3 estimates the selection regression accounting for the US Army Corps' choice of areas protected by levees. Section 4.4 presents the identification strategies. Section 4.5 presents the econometric specifications and the results.

4.1 A 1940–2010 Longitudinal Panel with Consistent Tract Boundaries

We build a longitudinal panel with neighborhood demographics and population density, house values, rents, and mobility flows, matched with the boundaries of leveed areas and flood risk measures to measure the magnitude of the mechanisms described in the previous section: capitalization, social interactions, mobility. Such longitudinal panel is built using two sets of sources. For 1940, 1950, 1960, we use tract-level data from the National Historical Geographic Information

System (NHGIS) of the University of Missouri. These provide both tract-level data and tract shapefiles. For 1970-2010, we use Geolytics' Neighborhood Change Database (NCDB), which provides tract-level data in consistent 2010 tract shapefiles. This second data set is Card et al.'s (2008). For the first data set we build Census tract relationship files for 1940 to 2010, 1950 to 2010, and 1960 to 2010, by intersecting tracts and building counts in the same process as Geolytics. This is illustrated on Figure A16 for a section of the city of St Louis. We keep a longitudinal data set of tracts for the metropolitan areas that are followed continuously from 1940 to 2010, and check that results are robust to the use of a balanced longitudinal panel vs. a panel which includes all tracts.

Table 7 presents the 20 metropolitan areas of our 1940-2010 longitudinal tract-level sample with the largest 1940 population within the leveed area. The metro area with the largest such 1940 population is Los Angeles, with 800,427 residents, 783 million dollars (in 1940 dollars) of aggregate value.

4.2 The Height Above Nearest Drainage

In flow routing models such as David, Maidment, Niu, Yang, Habets & Eijkhout (2011), flood risk depends on the probability that the stage (water depth) reaches each parcel of land's height above nearest drainage (HAND). Here for this analysis, we use a measure of flood risk at approximately 8 meter resolution, or 1/3 arcsecond[23]. The essential idea of this approach is to measure the height of a given cell relative to the nearest stream. Given a volume of water flowing through a section of river, also known as the reach, such volume can be converted into a water depth, also known as the stage, and a cell is flooded if its height relative to the nearest reach is less than the stage. Simply,

$$P(\text{Flooded}_{jt} = 1) = P(\text{HAND}_{jt} \leq \text{Stage}_{jt}), \tag{37}$$

which is the essential principle of recent developments in hydrological papers leading to inundation maps (Nobre, Cuartas, Hodnett, Rennó, Rodrigues, Silveira & Saleska 2011, Maidment 2017). Thus a higher height above nearest drainage is negatively correlated with flood risk probabilities. The strength of this HAND approach is its ability to provide flood risk proxies at high resolution using a Digital Elevation Model (DEM), a systematic nationwide method at high resolution. We use 10 meter resolution (1/3 of an arcsecond) measures of the HAND from the National Flood Interoperability Experiment (NFIE),[24] for the universe of catchments of our longitudinal panel. Here we use the HAND raster layers for each of the 6-digit Hydrological Units of the United States, and extract the average HAND by Census tract. This yields a measure of HAND for the 13,567 census tracts of the sample across 59 metropolitan areas.

[23]At a latitude of 37 degrees. At the equator, the resolution is 10 meters.

[24]HAND rasters are available from `https://web.corral.tacc.utexas.edu/nfiedata/`, sorted by hydrological unit. The HAND can also be calculated for an arbitrary DEM using the US Army Corps' GRASS `https://grass.osgeo.org/`.

4.3 Within-City Spatial Design of Leveed Areas: the Historical Selection Equation

Which demographics and places were initially protected by levees? Table 8 presents a selection regression using the tracts of the metropolitan areas of the balanced panel 1940–2010. These specifications correlate 1940 demographics and house values with the subsequent protection by a levee system, the leveed areas boundaries delineated by the US Army Corps. It regresses an indicator variable for whether the Census tract is within such leveed area, on 1940 racial composition (fraction black), average house value, measured as the ratio of the aggregate value of the housing stock in 1940 dollars over the number of housing units, and the Height Above Nearest Drainage. This is the observed counterpart of the structural investment selection equation (10).

$$\begin{aligned}\mathbf{1}(\text{Leveed After 1940})_j = \beta_0 &+ \beta_{\text{HAND}}\text{HAND}_j + \beta_1\text{Frac. Black}_{j,1940}\\ &+ \beta_2\text{Black Population Flow}_{1910-1940,j}\\ &+ \beta_3\text{Population Outflow}_{1910-1940,j}\\ &+ \beta_4\log(\text{Average Housing Value})_{j,1940}\\ &+ \beta_5\log(\text{Distance to Center})_j + \text{MSA}_{s(j)} + \varepsilon_j \end{aligned} \tag{38}$$

The covariates are 1940 tract-level measures, including the flood risk proxy HAND, the fraction black, the log of the aggregate housing value, as well as interactions between the HAND and these covariates. The HAND is inversely related to flood risk. Table 8 presents the results of such analysis. Results suggest a strong correlation between levee area protection and flood risk, which survives the inclusion of interaction terms and fixed effects: a 10 meter increase in the HAND reduces the probability of levee protection by 1.2 to 1.4 percentage points. Results also suggest that areas protected by levees had initially a higher fraction of black residents. A 10 percentage point increase in the fraction black is correlated with a 8.2 and 32 percentage points higher probability of levee protection. This correlation is statistically significant in specifications (4) of the upper panel, and (2) of the lower panel. The point estimates are positive in the 6 specifications that include this variable. The results also suggest that aggregate house values were lower (negative coefficient in the 4 regressions of the lower panel), although the effect is not statistically significant when clustering at the state level. Interaction terms do not suggest evidence that areas with higher fractions of black residents were less likely to be in the leveed area conditional on the HAND flood risk measure.

The selection of areas for flood protection is illustrated by the maps of St Louis and Los Angeles presented on Figure 4. Areas of the river initially settled are significantly exposed to flood risk and the city's population moved uphill to areas on higher ground. This is visible in the upper panel, Figure 4(a), as there is an up to 50 meter difference in elevation between river banks and downtown St Louis. Figure 4(b) presents aggregate housing values by tracts, suggesting that the highest housing values were observed in 1940 in areas with higher elevation.

Los Angeles was known as the city of a thousand rivers (Orsi 2004) as an overlay of the city's

1896 USGS Topographic map with the current layout of leveed areas displays the historical presence of numerous small streams, creeks, and rivers that once flowed through the region. Many of these waterways were channeled or buried as Los Angeles urbanized. Areas from downtown to Long Beach include the neighborhoods of Lynwood and Compton. The HAND of Figure 4(b) suggests that these areas had lower elevation relative to the nearest stream, as affluent neighborhoods were initially on higher ground and thus significantly less exposed to flood risk. Such areas had higher house values than areas protected by the levee system in 1940.

4.4 Identification Strategies for the Impact of Levees on City Dynamics

Identifying the causal impact of levees is challenging for at least three reasons. First, hydrological flood risk is a smooth decreasing function of the distance to the river banks and the coastline is thus confounded by the preference for riverine and coastal amenities. Correlations of levee protection with house prices and population partially capture such preferences as well as levee protection. Second, levees protect areas that, in 1940, had lower house values, lower population density, and higher shares of minority residents, which could be a source of downward bias for the impact of levees. The construction of levees could have coincided with expectations of complementary investments or pre-existing population trends, which could bias our estimated effects upward. This would be the case if, for instance, levees were built in areas experiencing gentrification. Third,

This paper's identification strategies address these concerns by combining controls for time-varying confounders and by using a regression discontinuity design at the boundary of leveed areas. This public information about leveed areas differs from hydrological flood protection. Indeed, the probability of flooding described on equation 37 is continuous rather than binary: the distribution of the stage is a continuous function of the streamflow (such relationship is called the rating curve (Maidment et al. 1993)), and the HAND does not experience systematic discontinuities at the boundary of leveed areas. In contrast, the communication of flood protection is binary and discontinuous. The border of leveed areas was typically delineated by projecting the crest of the embankment as a natural line on surrounding elevations.

Panel Fixed Effect Regressions

Three identification strategies address each of these empirical concerns. We start with a descriptive approach that performs a simple panel analysis. This panel regression controls for year-, tract- and metropolitan-area specific confounders. This is a useful benchmark before embarking on more sophisticated approaches. We estimate regressions at the Census tract $\times$ year level.

$$\begin{aligned}\text{Outcome}_{jt} &= \sum_{\tau=1950,\ldots,2010} \delta_t \text{In Levee}_j \times 1(\text{Year}_t = \tau) \\ &\quad + \text{Census Tract}_j + \text{County}_{c(j)} \times \text{Year}_t + \text{Residual}_{jt} \qquad (39)\end{aligned}$$

where Outcome$_{jt}$ is the log average value of housing units, the fraction Black in the neighborhood, the log number of housing units, and the log aggregate value of housing in the neighborhood. $\text{County} \times \text{Year}$ fixed effects account for the evolution of county-level trends in population and house values, and thus this specification identifies the impact of leveed areas within metropolitan area and within each decade. The inclusion of Census Tract fixed effects means that the impact of levees is identified by looking at relative population, house value, rent growth. Standard errors are double-clustered at the tract and year levels. Together this accounts for common tract-level non-time varying unobservable confounders and for location-specific time trends.

Regression Discontinuity Designs 1940–2010:
Cross Sectional RDs and Combined Difference-in-Differences RD Designs

Our next identification strategy combines the controls of the previous specification with tract and county $\times$ year fixed effects, with a regression discontinuity design at the boundary of leveed areas. The identification assumption here is that the discontinuity in the labelling of the leveed area changes the information set of households considering purchases and mobility towards leveed areas, while it does not affect the physical flood risk in the area.

To perform this, we consider two versions of the specification (i) repeated regression discontinuity designs every decade from 1940 to 2010, controlling for county $\times$ year fixed effects, which provides a regression discontinuity estimate for each decade over time; this would for instance allow us to observe if the discontinuity has flipped over the eight decades of the analysis; or (ii) a regression discontinuity design where all eight previous regression discontinuity designs of (i) are pooled, and a tract fixed effect is added. This controls for tract, county$\times$year, fixed effects and performs the analysis at the boundary of the leveed area.

The main advantage of approach (i) is that it provides an initial, in 1940, RD estimate at the boundary. Such initial RD is absorbed by the tract fixed effect in approach (ii).

Approach (i) amounts to performing regressions separately for each decade $t = 1940, 1950, \ldots, 2010$

$$\text{Outcome}_{jt} = \delta_t \text{In Levee}_j + \text{County}_{c(j)} \times \text{Year}_t + \text{Residual}_{jt} \tag{40}$$

while weighing observations by a function of their distance to the leveed area boundary. With a rectangular kernel, this amounts to limiting observations to tracts at a maximum distance of $h \in \{0.25, 0.5, 1\}$ mile of the boundary. With a Gaussian kernel, this amounts to weighing observations by $f\left(\frac{d_j}{h}\right)$, where d_j is the distance of tract j to the leveed area boundary, f is the density of the normal distribution, and h is the bandwidth (Imbens & Lemieux 2008). We consider bandwidths $h \in \{0.25, 0.5, 1\}$, as for the rectangular kernel.

Approach (ii) amounts to estimating specification 39:

$$\begin{aligned} \text{Outcome}_{jt} &= \sum_{\tau=1950,\ldots,2010} \delta_t \text{In Levee}_j \times 1(\text{Year}_t = \tau) \\ &\quad + \text{Census Tract}_j + \text{County}_{c(j)} \times \text{Year}_t + \text{Residual}_{jt} \end{aligned} \tag{41}$$

while weighing for $w_j = f\left(\frac{d_j}{h}\right)$ as before.

Regression Discontinuities without Flood Risk Discontinuities

In (41), tract fixed effect controls for preexisting differences in flood risk across the leveed area boundary. We conduct a robustness check to assess whether results are robust to comparing tracts on parts of the leveed area where there is no statistically significant discontinuity in physical flood risk. We compare census tracts with similar flood risk – i.e. with no statistically significant difference in the Height Above Nearest Drainage defined in Section 4.2 – but with different public information about levee protection.

We consider the set of 2,171 5-digit ZIP codes containing the Census tracts of the 1940–2010 balanced panel. 339 5-digit ZIP codes straddle both sides of a leveed area's boundary, containing 3,838 tracts. 261 ZIP codes have sufficient data to estimate a regression using both sides of the boundary, containing 2,184 tracts. For each ZIP code j, we run the Census tract-level regression:

$$\text{HAND}_i = \alpha_j + \beta_j \text{In Levee}_i + \varepsilon_i, \qquad \text{for all } i \text{ s.t. } J(i) = j,$$

where $J(i)$ is the ZIP code of tract i. This provides an estimate of the ZIP-level discontinuity $\hat{\beta}_j$ and an associated t-statistic t_j. We consider the set of tracts $j$such that $\hat{\beta}_j$ is not significant at 95%:

$$\text{Condition :} \qquad \left|\frac{\hat{\beta}_j}{t_j}\right| < 1.96 \tag{42}$$

This set includes 179 ZIP codes do not exhibit a statistically significant discontinuity in the Height Above Nearest Drainage at 95%, containing 1,446 tracts. The t-statistic in this sample is on average -0.51. Table A31 suggests that such no-discontinuity sample includes the largest metropolitan areas of the balanced panel of 1940–2010: Los Angeles has 71 such 5-digit ZIP codes including 659 tracts, Indianapolis 11 ZIP codes including 76 tracts, St Louis includes 7 such ZIP codes with 45 tracts. Overall this sample covers 31 metropolitan areas, and 1,440 Census tracts.

Timing of Congressional Appropriations and Regression Discontinuity

Finally, we consider a regression that (a) has the features of specification (41) but focuses only on those discontinuities in spatial proximity with places funded in the text of the 1936 Flood Control Act. The printed version of the Flood Control Act was digitized and geocoded as described in Section 41. Place names were given a latitude and longitude according to the US Census Bureau's place names. The panel is then limited to the set of Census tracts within 10 or 20 miles of an appropriated project. Table 15 provides the number of tracts thus considered. In 1940, 383 tracts are within 10 miles of an appropriated project and within a leveed area. 748 tracts are within 20 miles of an appropriated project and within a leveed area.

The regression discontinuity is performed using a rectangular kernel with a bandwidth of 1 mile, and the regression controls for tract, county $\times$ year fixed effects. Standard errors are double-clustered by tract and county$\times$ year, at the same level as the fixed effects.

4.5 Results

This section presents the results of the four specifications: specification (39) on Table 10, specification (40) on Tables 11 to A25, specification (41) on Tables 12 to 12, and the specifications with no HAND discontinuity on Table 14, and within a project funded by the Flood Control Act on Table 16 and A30.

Housing Values Average housing values are the dependent variables in column (1) of each of the tables of results.

A common set of patterns emerges from these regressions. In each regression, average housing values increase, with point estimates significant at 5 and 1%. The different approaches differ in the timing and the specific magnitude, but suggest overall economically significant positive impacts on housing values. In the panel f.e. regression with tract and county $\times$ year fixed effects, housing values increase by 14.58% in 1960 before tapering off and increasing 8.9% relative to non-leveed tracts in the same county-year. This effect is also observed when combining the RD and the panel f.e. approach.

The repeated cross sectional RD estimates suggest that this estimate is mostly due to a *catch-up effect*. Indeed, Table 11 and other tables suggest that tracts within leveed areas had lower housing values in 1940, between 15 and 21% lower, before catching up, either with no significant difference or with positive differences.

Neighborhood Demographics

The fraction black in the tract is the dependent variable in column (2) of each table.

Here again the range of estimates suggests that leveed areas become significantly less Black over time. The panel f.e. regression with tract and county $\times$ year fixed effects suggest that such neighborhoods were 2.6 percentage points less Black in 2010, with the change in racial composition mostly occurring in 1960 and 1970. Regression discontinuity estimates suggest significantly larger impacts. On Table 11, the fraction Black is higher in leveed areas in 1940, by about 1.5 percentage points. In 1950 there is no statistically significant difference, and the difference becomes negative in 1960, statistically significant at 1% from 1990 onwards. In 2010, leveed areas are 2.3 percentage points less Black. Significance levels improve markedly when expanding the bandwidth to 0.5 mile and 1 mile. These results are robust to the inclusion of tract fixed effects in the pooled RD and panel f.e. regression. Effects reach up to 5 to 6 percentage points in the long run.

Population Density and Housing Density: Do Levees Lead to Density or Urban Sprawl? On each table, column (3) is for log(tract population), column (4) for log(housing units) and column (5) for log(aggregate tract housing value). Our preferred estimate using regression discontinuity designs suggests increasing density and construction in leveed areas.

Urban economic theory suggests that improvements in the quality of land (through a reduction in flood risk) can cause at least two potential different scenarios. In the first scenario, the desirability of protected land leads to population inflows, reductions in the consumption of land per capita, and higher house prices. This occurs when amenities and the consumption of land are substitutes. In the second scenario, the desirability of such land leads to greater consumption of land, higher house prices, and less density. This would occur when amenities and the consumption of land are complements.

Results using the panel f.e. approach with both tract and county $\times$ year f.e.s, suggest that neighborhoods become more populated and have more housing in the four decades following the 1936 Flood Control Act. Yet the estimates on log population and log housing units become negative in 2000 and 2010. This suggests that as areas protected by levees may become more dense and experience more construction activity, these areas become less built up and less populated in 2000 and 2010. The impact on the log aggregate value of housing is not significant, while the impact on the average value is strongly positive, suggesting that these neighborhoods are experiencing higher long-run house prices but fewer housing units. This is consistent with a more constrained long-run housing supply in leveed areas.

Regression discontinuity designs are line with the panel f.e. results when looking at the impact on house prices and fraction black. On population density, the results depend on the specification. They provide results that suggest that leveed areas became more dense over the 1950–2010 time period when using Gaussian or rectangular kernels, with buffers of 0.25, 0.5, 1 mile; this is true in the cross sectional RDs or the pooled RD in the longitudinal dimension with tract f.e.s, and/or areas close to 1936 Flood Control Act appropriations. The design and timing of infrastructure investment depends but does not hinge on the estimated results thus obtained.

5 Structural Analysis of Optimal Levee Designs

5.1 Estimation of Utilities and Mobility Costs

The first step is to estimate tract-level structural parameters: utilities U_{jt}^g, mobility costs μ_{jt}^g, and continuation values Π_{jt}^g. Intertemporal valuations U_{jt}^g and mobility costs μ_j are estimated by matching the flow of movers to each location, and matching the probability of outward mobility from each location. Conditional on mobility, the destination of the move identifies the intertemporal values. The mobility cost is estimated by showing that, conditional on valuations $\{U_{jt}^g\}_{j=1,2,\dots,J}$, the probability of leaving neighborhood j is a one-to-one function of μ_j. For the share of movers in each location, we use mobility measures from the 1940–2010 Censuses. For the numbers leaving a neighborhood, we notice that this can be estimated by taking the difference between the change in population and inward mobility flow.[25]

Utilities U_{jt}^g are estimated using data on the location choices of movers. The share of movers choosing each location is observable given tract-level data on migration. Mobility costs are estimated using data on stayers. This is computable given data on (i) inflows of movers to each neighborhood combined with (ii) the evolution of population between each decade. An accounting equation provides us the relationship between the change in population in j between t and $t+1$ and the inflow of movers to the location.

$$\underbrace{L_{jt+1} - L_{jt}}_{\text{Population change, observed in Census}} = \underbrace{L_{jt+1}^{+}}_{\text{Inflow of movers}} - \underbrace{L_{jt}(1 - P(\text{staying}|j,t))}_{\text{Outflow of leavers}}$$

The quantities L_{jt+1}, L_{jt}, and L_{jt+1}^{+} are observable in Census data, and provide us with an estimate of $P(\text{staying}|j,t)$.

For the identification of utilities, we rely on the probability of moving to j conditional on mobility:

$$P(j,t+1|\text{mobility}) = \frac{U_{jt+1}^{\theta}}{\sum_{k=1,k\neq i}^{J} U_{kt+1}^{\theta}} \tag{43}$$

This provides a vector U_{jt} for each period that explains the inflows. The probability of staying is explained by finding the μ_j in each location j.

$$P(\text{staying}|j,t) = \frac{(U_{jt})^{\theta}}{(U_{jt})^{\theta} + \sum_{k=1,k\neq j}^{J}(U_{kt}/\mu_j)^{\theta}} \tag{44}$$

The continuation values Π_{jt} are the sum of utilities divided by the mobility costs.

$$\Pi_{jt}^{g} = \Gamma \left\{ \sum_{k=1}^{J} \left(\frac{U_{kt}^{g}}{\mu_j^{\mathbf{1}(k\neq j)}} \right)^{\theta} \right\}^{\frac{1}{\theta}}.$$

Continuation values Π_{jt+1}^{g} are right-hand side covariates in the regression where utilities U_{jt}^{g} are the dependent variables (specification (45)).

5.2 Parameter Estimation

For each tract j in each decade t, the valuation U_{jt}^{g} depend on local amenities, house prices q_{jt}, population L_{jt-1} in the previous decade, racial composition x_{jt-1} in the previous decade, the

[25]A robustness check can adjust for the natural change in population due to births and deaths. If those natural rates do not differ across locations, the identification of the mobility costs can proceed as described.

future value Π_{jt+1} in $t+1$ starting from location j. This leads to the following specification:

$$\log(U^g_{jt}) = \underbrace{\frac{\alpha}{\sigma}\log A^g_{jt}}_{\text{Tract}\times\text{Race f.e.}} - \frac{\alpha}{\sigma}\log q_{jt} + \frac{\alpha}{\sigma}\log y^g_t + \xi\log L_{jt-1} + \gamma^g \log x_{jt-1} + \delta\log\Pi_{jt+1} + \varepsilon_{jt} \quad (45)$$

Where U^g_{jt} is identified using movers, Π_{jt+1} is the continuation value, γ is the preference for density, γ^g is the racial preference parameter. The continuation value is built before running the regression using the intertemporal utilities and the mobility costs estimated by the previously mentioned contraction mapping.

$$\widehat{\Pi_{jt+1}} = \Gamma\left\{\sum_{k=1}^{J}\left[\frac{\hat{U}^g_{kt+1}}{\hat{\mu}_j^{\mathbf{1}(j\neq k)}}\right]^{\theta}\right\}^{1/\theta} \quad (46)$$

This is described in Appendix Section 5.1. We parameterize the amenity fixed effects by a location- and race-specific amenity fixed effect, a year fixed effect, and the impact of flood control infrastructure In Levee$_j = 0, 1$ on amenities.

$$\log(A^g_{jt}) = \log A^g_j + \log \text{Year}_t + \sum_{\tau=1950,\ldots,2010} \zeta_\tau \text{In Levee}_j \times \text{Year}_t + \text{Residual}^g_{jt} \quad (47)$$

We estimate $\widehat{\log A^g_j}$, $\widehat{\log \text{Year}}_t$, $\widehat{\zeta}_\tau$, $\widehat{\xi}$, $\widehat{\gamma^g}$ and, following Ahlfeldt et al. (2015), we calibrate α to the share of housing in income. We calibrate r to the average rate of the 10-year Treasury (FRED time series).

The estimates are $\widehat{\xi} = 0.6961^{***}(0.045)$ for the agglomeration economy parameter, suggesting path dependence, $\widehat{\gamma}^b = 0.5090^{***}(0.0397)$ for the racial preference of black residents for black neighbors. The coefficient $\widehat{\gamma^w} = 0.1$ (0.200) is not significant at 10%. The coefficients $\widehat{\zeta}_\tau$ measure the impact of levees on local amenities.

5.3 The role of path dependence in the measured impact of levees

Coefficients ζ_τ in specification (45) measure the impact of levees on values conditional on prior population distributions. It is useful to compare them to estimates reduced-form estimates $\widehat{\phi_\tau}$ without conditioning on prior population to estimate the role of path dependence in the attractiveness of leveed areas. This is done in this specification as:

$$\log(U^g_{jt}) = \log A^g_j + \log \text{Year}_t + \sum_{\tau=1950,\ldots,2010} \phi_\tau \text{In Levee}_j \times \text{Year}_t + e_{jt} \quad (48)$$

which do not control for future values, prices, and demographics. Reduced-form estimates suggest that values are significantly higher in neighborhoods protected by levees, increasing from 1940 to 1950 ($+1.102^{***}(0.467)$) and from 1950 to 1960 ($+1.765^{***}(0.4171)$), before tapering off, consistent with descriptions of investments by the US Army Corps of Engineers.

Structural estimates $\widehat{\zeta}_\tau$ controlling for agglomeration economies γ and γ^g, future values, and prices suggest that higher values are due to this investment in *one* decade, which persists over multiple decades due to path dependence implied by the preference for density and racial preferences. Indeed, when controlling, only $\widehat{\zeta_\tau}$ for $\tau = 1960$ is positive and significant at 1% $(+1.137^{*}(0.5572))$. This is a precursor for the results of the impulse response analysis of flood control investment presented next. These results for $\widehat{\zeta_\tau}$ suggest a significant role of γ and γ^g in the dynamic of neighborhoods post-levee investment.

The other parameters of the model are calibrated as follows. Housing supply elasticity is that of Saiz (2010) for St Louis. The housing depreciation parameter ζ in equation (7) is estimated using the autoregressive parameter of log housing units.

5.4 Empirical Results: General Equilibrium Impacts and Path Dependence with the Actual Design

The quantitative spatial approach allows us to (a) isolate the impact of the levee separately from the impact of other drivers of city dynamics, assessing the contribution of each mechanism to the long-run impacts of levee investment; what share is due to the preference for population density, due to racial preferences, due to beliefs about future welfare, and due to mobility costs? By shutting off each channel, we can estimate the contribution of these mechanisms. (b) it also allows us simulate counterfactual levee investments: either by simulating different timings, or by simulating different choices of places protected against floods.

This section describes (a) the impact of the observed levee system and its mechanisms. It highlights the importance of path dependence in generating impacts for multiple decades for a single fully depreciating investment.

Impulse response functions are presented on Figure 6 and impulse response maps on Figure 7. Figure 7a shows that the areas in red experience an increase in the amenity value $\tilde{A}_{jt}$ for one decade. These are the Census tracts within the leveed area. Figure 7b shows that these areas capitalize flood protection, with price increases ranging between 20 and 30%. This is in response to population flows, depicted on Figure 7c. The last panel, on Figure 7d shows that population flows are heterogeneous by race, as the fraction white changes differently across locations. White households tend to leave higher-elevation areas to move to the floodplain, which experiences up to a 15 percentage point increase in the fraction White. These effects are in line with the estimated effects in the non-structural part of this paper.

On each impulse response function presented on Figure 6, the dotted line is the impulse response when protecting the leveed area for one decade, and the dotted line is the impulse response function when protecting the leveed area for two decades. In each case, the levee fully depreciates after either one decade or two decades. This is in line with the estimates presented on Figure 8. The levee system causes welfare gains, as intertemporal welfare increases by 7% (one-decade flood protection policy) and by 12% (two-decade flood protection policy) for Whites. It causes a welfare gain of 1.3% for Blacks (one decade of flood protection) and 2.5% (two decades of flood protection). The depreciation of the levee causes welfare declines after 3 (resp.

4) decades. This is due to two phenomena. First, households that have moved to the leveed area are less likely to move out as they face mobility costs. Figure 6c shows that prices remain elevated in the leveed area even after the full depreciation of the levee system. This is in part due to the path dependence caused by the preference for population density. Figure 6a shows that white households experience welfare losses in decades 3 to 7 (one decade flood protection policy) and in decades 4 to 7 (two-decade flood protection policy). The second phenomena that explains the welfare losses is that, for households already in the floodplain, flood protection is capitalized, and prices rise in response to inflows from households of other parts of the metro area, while mobility costs prevent perfect mobility towards other parts of the metro area. Thus Figure 6b shows that black households experience welfare losses in decades 2 to 7 (one-decade flood protection policy) and in decades 3 to 7 (two-decade flood protection policy).

5.5 Welfare-Maximizing Spatial Design, One-Period Investment

We start by focusing on the spatial design in the case of a one-period investment. The levee is built at $t = 1$ by choosing a subset of locations for investment, and it fully depreciates from $t = 2$ onwards.

We find the welfare-maximizing spatial design using the approach of Section 2.2.3. The boundaries of the leveed area are pinned down by three covariates of the selection equation: the Height Above Nearest Drainage, the log price of housing and the prior log population. This ties the probability of flood protection of equation (10) to covariates as follows:

$$P(\text{Investment}_{j1} = 1) = \Lambda(\beta_0 + \beta_{\text{HAND}}\text{HAND}_j + \beta_q \log \text{Price}_{jt} + \beta_L \log \text{Population}_{jt}) \tag{49}$$

and $P(\text{Investment}_{jt} = 1) = 0$ for any $t > 1$.

This is the counterfactual counterpart of the observed levee system as designed by the US Army Corps, and estimated in Section 4.3's regression (38). In this sense the model allows us to simulate the city's path with a different of investment selection equations.

We find the optimal $(\beta_0^* \, \beta_{\text{HAND}}^* \, \beta_q^* \, \beta_L^*)$ by maximizing households' aggregate welfare with respect to $\boldsymbol{\beta}$. The vector $\boldsymbol{\beta}^*$ is as follows: the constant is -3.64, the coefficient of the height above nearest drainage is -18.65, the coefficient of the log house price is 8.70, the coefficient of log population is 0.028. Compared to the observed design (Figure 9), this vector $\boldsymbol{\beta}^* = (-3.64 \; -18.65 \; 8.70 \; 0.028)$ reduces flood protection more aggressively as the height above nearest drainage increases, and, instead of protecting more sparsely populated neighborhoods, as the US Army Corps did, the optimal design protects those areas with larger population.

The map of the optimal levee design is on Figure 9. The green lines are for the actual leveed area, and the light blue lines for the area protected under the optimal design. The map makes it clear that those areas more likely to be protected under the optimal design are those close to the city's core. The optimal design does not protect more sparsely populated areas of the northwest and the south of the city.

Figure 10 presents the impulse response functions when investing in the optimal levee design vs investing in the actual design. The charts are obtained as in Section 2.2.3, by the Blanchard & Kahn (1980) method described in DeJong & Dave (2012). The optimal design leads to a larger welfare gain for the first two decades and a smaller welfare loss than the actual design (panels (a) and (b)). This is true for both Whites and Blacks. With the optimal design, Blacks experience welfare gains for two decades, while they experience welfare losses with the actual design. This is largely for two reasons. First, the optimal design leads to a substantially lower price increase, as flood protection leads to an increase in prices of only 7% while the observed design leads to a 19% increase in prices. This causes smaller welfare declines for households either already in the floodplain (Blacks) or moving to the floodplain (Whites). This also causes smaller capital gains for land owners in the floodplain. The second reason for the substantially higher welfare increase with the optimal design is that it leads to less urban sprawl. Panel (d) shows the level of population concentration, measured as the evolution (log-linearized) of the average population density for an average household. This is akin to a Herfindahl index. Households benefit from population concentration thanks to the preference ξ for density. With the observed design there is a decline in population concentration, stretching the existing population by providing incentives to move to lower density floodplains. With the optimal design the level of population concentration increases as households move to already populated areas.

5.6 The Optimal Timeline of Infrastructure Investment: Commitment vs. Time Consistency

We then turn to the case of a sequence of investments over time. The policymaker designs a levee system at $t = 0$ and then alters this levee system over time either by upgrading it, maintaining it, or letting it depreciate.

This is modeled as a time-varying probability of investment in each location:

$$P(\text{Investment}_{j1} = 1) = \Lambda(\beta_{\text{HAND}}\text{HAND}_j + \beta_q \log \text{Price}_{j0} + \beta_L \log \text{Population}_{j0} + \sum_{\tau=1}^{T} \beta_\tau \text{Year}_{\tau t}) \quad (50)$$

where $(\beta_{\text{HAND}}, \beta_q, \beta_L)$ determine the spatial design of the levee, and the sequence of $\{\beta_1, \beta_2, \ldots, \beta_T\}$ determines the investment in maintenance and upgrading of the levee. With a fully depreciating levee $\beta_\tau = 0$ at $\tau = 1$ and $\beta_\tau = -\infty$ thereafter.

When the policymaker can commit to a specific levee design, the choice of the design $(\beta_{\text{HAND}}, \beta_q, \beta_L)$ and the choice of the sequence $\{\beta_1, \beta_2, \ldots, \beta_T\}$ is made at $t = 0$. The first-order conditions are such that the marginal impact of a shift in β_k on welfare at $t = 0$ equates the Lagrange multiplier of the cost constraint times the marginal impact of β_k on the cost.

When the policymaker cannot commit to maintenance and upgrades, the choice of the spatial design $(\beta_{\text{HAND}}, \beta_q, \beta_L)$ is made at $t = 0$ but the choice of $\beta_2, \ldots, \beta_T$ is made at $t = 2$, $t = 3$, ..., $t = T$ respectively, to maximize each period t's welfare. The policymaker is competing with its future self: today's policymaker selects coefficients $\boldsymbol{\beta}_t$ given the choice of the future policymakers at $t+k$, $k>0$, and given the choices $t-k$ of past policymakers. Households' location choices are determined in part by expectations of future investment. This is Kydland & Prescott's (1977)'s insight.

We solve for the optimal investment strategy in the commitment and the no-commitment (time consistent) case. The estimates of the betas of the selection equation (50) are on Table 19. General equilibrium responses of prices and flood risk exposure are Figure 11.

The main findings are as follows. First, the no-commitment policymaker invests substantially more in levee systems throughout the time periods than the time consistent policymaker who follows the investment strategy at $t=0$. The probability that a neighborhood is within the leveed area increases 3 to 4 fold from 0.46% to 1.342%. Second, in the long run, the policymaker increases its investment in the levee system in the no-commitment case, while the policymaker lowers its investment in flood protection in the commitment case. In this case, committing to lower investment in flood protection leads to lower inflows into leveed area at the dynamic equilibrium and less flood risk exposure. Third, Figure 11 panel (a) shows that households move substantially closer to the rivers when the policymaker cannot commit (solid line) as opposed to the case where the policymaker can commit (dashed line). Fourth, panel (b) suggests that there is substantially more capitalization of flood protection in the no-commitment case.

6 Discussion

6.1 Floods and Levees: Historical Gauge Records, Flow Routing, and Neighborhood Attractiveness

Structural and non-structural estimates suggest that (i) 1936 was the key Flood Control Act that had the most economically significant impact on flood protection infrastructure (Table 3), and that (ii) population flows after 1950 were driven by path dependence, preference for population density, and racial preferences rather than by the causal effect of the levee system on neighborhood attractiveness (Figure 8). This is consistent with the model's mechanisms.

This section discusses whether such decline in the attractiveness of leveed areas is driven by a decline in the effectiveness of the levee system. While hydrological analysis is beyond the scope of an economics paper, a range of empirical evidence suggests frequent flooding within leveed areas.

The U.S. Geological Survey provides water gauge data across at 18,767 locations across the United States, provided as the *USGS Surface-Water Historical Instantaneous Data for the Nation*. We collect it for the set of water gages across our 59 metropolitan areas to build a time series panel between 1950 and 2010. Water gages provide an invaluable historical record sometimes starting as early as the end of 19th century. Figure 13 presents a timeline that focuses on the 6 water gauges within leveed areas of the Los Angeles, with the first recorded peak streamflows in

February 1937. This type of chart has the benefit of being a more comprehensive record when textual archives do not provide details. The Figure displays familiar events. The *Times-News* reports on February 22, 1980 that "caskets floated out of rain-sodden graves as levees crumbled and dams overflowed in Southern California, forcing thousands of people to flee before the rain subsided today. [...] Six California Counties – from Ventura just north of Los Angeles to San Diego on the Mexican Border – were declared national disaster areas by President Carter. [...] Los Angeles has received 12.75 inches of rain in a nine-day period." One of the 6 water gages indeed recorded a corresponding peak streamflow in February 1980, as reported on Figure 13.

Table A32 presents systematic statistics on USGS water gages, within and outside leveed areas, and the number of peak streamflow events recorded at these stations. Streamflow is key in understanding risk as there is a one-to-one relationship between streamflow and water levels (the 'stage') through the rating curve (David et al. 2011, Maidment 2017).

The 56 metropolitan areas of our sample have up to 9 water gages within leveed areas (Louisville/Jefferson County), 8 (Kansas City), 6 (Los Angeles-Long Beach-Anaheim). These metropolitan areas have between 185 and 5 water gages within their boundaries. These record up to more than 6,000 peak streamflows, and up to 309 peak streamflows within leveed areas. Focusing on those peaks in the 75th percentile of the overall distribution (across water gages) of streamflow yields up to 89 peaks in Kansas City, 9 in Los Angeles, 18 in Riverside-San Bernardino-Ontario, CA.

The second piece of descriptive evidence comes from presidential disaster declarations. The Disaster Relief and Emergency Assistance Act, 42 U.S.C. §§ 5121-5206, also known as the Stafford Act of 1988, provides a path for the Governor of a state affected by a major disaster to request a presidential disaster declaration. This request is filed with one of the 10 regional FEMA offices. Disaster declarations lead to assistance, repair, restoration, and replacement of damaged facilities, debris removal, unemployment assistance, relocation assistance, community disaster loans, emergency grants, hazard mitigation, and other measures. Presidential disaster declarations are recorded and disseminated by FEMA in the *OpenFEMA Dataset: Disaster Declarations Summaries - v2*.

Figure 14 displays descriptive statistics on the number of disaster declarations per county per year. Charts suggest that from the early 1960s onwards, the number of disaster declarations in leveed counties (more than 60% of the county labelled protected) is higher than the average number of disaster declarations in non-leveed counties. This is also true when focusing on counties where 30% or more of the surface area is protected by a levee system.

7 Conclusion

As the risk of flooding rises in the 21st Century, governments are increasingly investing in infrastructure projects to protect coastal metropolitan areas. Notable examples include the storm surge barriers proposed in the NY & NJ Harbor & Tributaries (HATS) study, such as the West Side floodwall for New York City,[26] and the Peninsula Perimeter Protection Project for Charleston. Engineering forecasts for Charleston's 2082 flood scenarios illustrate potential impacts without changes in housing, population, or demographics.[27] However, economic agents inevitably respond to these risks, increasing demand for prime locations, while policymakers react to these shifts by updating benefit assessments for protective measures.

Modern quantitative spatial modeling approaches can be adapted to provide us with the economist's view of the general equilibrium impacts of such levee design. This paper's approach complements engineering solutions as it captures the dynamic general equilibrium response of populations' location choices, prices, and the supply of floor surface to infrastructure investments. This paper's approach also complements quantitative spatial models by endogenizing the behavior of the policymaker, who competes with their own future self. The policymaker's response depends on the impact of climate shocks on exposed households' welfare.[28] Climate change makes flooding an increasingly non-stationary process (Byun & Hamlet 2020), and policymakers face a dynamic problem of funding and maintaining levees and updating their design in response to 'new news.' Choosing the optimal design of place-based investments is a problem of implementation theory (Maskin & Sjöström 2002, Jackson 2001) whereby the policy anticipates the general equilibrium response of the city's spatial distribution of economic activity; in turn, the designer of the system should expect to face dynamic incentives to maintain or upgrade the infrastructure.

The paper underscores the critical role of credibility, time consistency, and expectations in policy effectiveness. Policymakers' ability to commit depends significantly on political institutions (Besley & Coate 1998, Barseghyan & Coate 2014, Azzimonti 2015). Because private investments hinge on expectations of the timing and location of public infrastructure, the tools developed here help policymakers design and implement consistent, forward-looking investment strategies.

[26] https://www.nan.usace.army.mil/Missions/Civil-Works/Projects-in-New-York/New-York-New-Jersey-Harbor-Tributaries-Focus-Area-Feasibility-Study/

[27] https://coastalconservationleague.org/projects/charleston-peninsula-coastal-flood-risk-management-study-by-the-us-army-corps-of-engineers/

[28] See Kahn (2016), Barreca, Clay, Deschenes, Greenstone & Shapiro (2016), Young & Hsiang (2024).

References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. & Wolf, N. (2015), 'The economics of density: Evidence from the berlin wall', *Econometrica* **83**(6), 2127–2189.

Albouy, D. (2016), 'What are cities worth? land rents, local productivity, and the total value of amenities', *Review of Economics and Statistics* **98**(3), 477–487.

Allen, T. & Donaldson, D. (2020), Persistence and path dependence in the spatial economy, Technical report, National Bureau of Economic Research.

Allen, T. & Donaldson, D. (2022), 'Persistence and path dependence: A primer', *Regional Science and Urban Economics* **94**, 103724.

Almagro, M. & Domínguez-Iino, T. (2024), Location sorting and endogenous amenities: Evidence from amsterdam, Technical report, National Bureau of Economic Research.

Arnold, J. L. (1988), *The evolution of the 1936 flood control act*, Office of History, US Army Corps of Engineers.

Azzimonti, M. (2015), 'The dynamics of public investment under persistent electoral advantage', *Review of Economic Dynamics* **18**(3), 653–678.

Barreca, A., Clay, K., Deschenes, O., Greenstone, M. & Shapiro, J. S. (2016), 'Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century', *Journal of Political Economy* **124**(1), 105–159.

Barseghyan, L. & Coate, S. (2014), 'Bureaucrats, voters, and public investment', *Journal of Public Economics* **119**, 35–48.

Bellemare, M. F. & Wichman, C. J. (2020), 'Elasticities and the inverse hyperbolic sine transformation', *Oxford Bulletin of Economics and Statistics* **82**(1), 50–61.

Berry, S., Levinsohn, J. & Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica: Journal of the Econometric Society* pp. 841–890.

Besley, T. & Coate, S. (1998), 'Sources of inefficiency in a representative democracy: a dynamic analysis', *American Economic Review* pp. 139–156.

Biswas, K., Kumar, S., Banerjee, S. & Pandey, A. K. (2021), 'Smu: smooth activation function for deep networks using smoothing maximum technique', *arXiv preprint arXiv:2111.04682* .

Blanchard, O. J. & Kahn, C. M. (1980), 'The solution of linear difference models under rational expectations', *Econometrica: Journal of the Econometric Society* pp. 1305–1311.

Bleakley, H. & Lin, J. (2012), 'Portage and path dependence', *The quarterly journal of economics* **127**(2), 587–644.

Boustan, L. P., Kahn, M. E., Rhode, P. W. & Yanguas, M. L. (2020), 'The effect of natural disasters on economic activity in us counties: A century of data', *Journal of Urban Economics* **118**, 103257.

Bradt, J. T. & Aldy, J. E. (2022), Private benefits from public investment in climate adaptation and resilience, Technical report, Working paper.

Brooks, L. & Liscow, Z. (2023), 'Infrastructure costs', *American Economic Journal: Applied Economics* **15**(2), 1–30.

Byun, K. & Hamlet, A. F. (2020), 'A risk-based analytical framework for quantifying non-stationary flood risks and establishing infrastructure design standards in a changing environment', *Journal of Hydrology* **584**, 124575.

Card, D., Mas, A. & Rothstein, J. (2008), 'Tipping and the dynamics of segregation', *The Quarterly Journal of Economics* **123**(1), 177–218.

Couture, V., Gaubert, C., Handbury, J. & Hurst, E. (2019), Income growth and the distributional effects of urban spatial sorting, Technical report, National Bureau of Economic Research.

David, C. H., Maidment, D. R., Niu, G.-Y., Yang, Z.-L., Habets, F. & Eijkhout, V. (2011), 'River network routing on the nhdplus dataset', *Journal of Hydrometeorology* **12**(5), 913–934.

DeJong, D. N. & Dave, C. (2012), *Structural macroeconometrics*, Princeton University Press.

Ehrlich, I. & Becker, G. S. (1972), 'Market insurance, self-insurance, and self-protection', *Journal of political Economy* **80**(4), 623–648.

Evans, G. W., Honkapohja, S. & Mitra, K. (2022), 'Expectations, stagnation, and fiscal policy: a nonlinear analysis', *International Economic Review* **63**(3), 1397–1425.

Fairweather, D., Kahn, M. E., Metcalfe, R. D. & Olascoaga, S. S. (2024), Expecting climate change: A nationwide field experiment in the housing market, Technical report, National Bureau of Economic Research.

Fajgelbaum, P. D. & Schaal, E. (2020), 'Optimal transport networks in spatial equilibrium', *Econometrica* **88**(4), 1411–1452.

Fan, Y. (2024), Levees and levies: Local financing of climate infrastructure maintenance and housing market dynamics, Technical report.

Glaeser, E. L. (2010), *Agglomeration economics*, University of Chicago Press.

Glaeser, E. L. & Gyourko, J. (2005), 'Urban decline and durable housing', *Journal of political economy* **113**(2), 345–375.

Guerrieri, V., Hartley, D. & Hurst, E. (2013), 'Endogenous gentrification and housing price dynamics', *Journal of Public Economics* **100**, 45–60.

Handbury, J. (2021), 'Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities', *Econometrica* **89**(6), 2679–2715.

Heblich, S., Redding, S. J. & Sturm, D. M. (2020), 'The making of the modern metropolis: evidence from london', *The Quarterly Journal of Economics* **135**(4), 2059–2133.

Heckman, J. J. (1979), 'Sample selection bias as a specification error', *Econometrica: Journal of the econometric society* pp. 153–161.

Henkel, M., Kwon, E. & Magontier, P. (2022), 'The unintended consequences of post-disaster policies for spatial sorting', *MIT Center for Real Estate Research Paper* (22/08).

Hino, M., Field, C. B. & Mach, K. J. (2017), 'Managed retreat as a response to natural hazard risk', *Nature Climate Change* **7**(5), 364–370.

Hirano, K., Imbens, G. W. & Ridder, G. (2003), 'Efficient estimation of average treatment effects using the estimated propensity score', *Econometrica* **71**(4), 1161–1189.

Hornbeck, R. & Naidu, S. (2014), 'When the levee breaks: black migration and economic development in the american south', *American Economic Review* **104**(3), 963–90.

Hsiao, A. (2023), 'Sea level rise and urban adaptation in jakarta', *Technical Report* .

Imbens, G. W. & Lemieux, T. (2008), 'Regression discontinuity designs: A guide to practice', *Journal of econometrics* **142**(2), 615–635.

Jackson, M. O. (2001), 'A crash course in implementation theory', *Social choice and welfare* **18**, 655–708.

Kahn, M. E. (2007), 'Gentrification trends in new transit-oriented communities: Evidence from 14 cities that expanded and built rail transit systems', *Real Estate Economics* **35**(2), 155–182.

Kahn, M. E. (2016), 'The climate change adaptation literature', *Review of Environmental Economics and Policy* .

Kleinman, B., Liu, E. & Redding, S. J. (2023*a*), 'Dynamic spatial general equilibrium', *Econometrica* **91**(2), 385–424.

Kleinman, B., Liu, E. & Redding, S. J. (2023*b*), The linear algebra of economic geography models, Technical report, National Bureau of Economic Research.

Knowles, S. G. & Kunreuther, H. C. (2014), 'Troubled waters: The national flood insurance program in historical perspective', *Journal of Policy History* **26**(3), 327–353.

Kousky, C., Luttmer, E. F. & Zeckhauser, R. J. (2006), 'Private investment and government protection', *Journal of Risk and uncertainty* **33**, 73–100.

Krugman, P. (1991), 'History versus expectations', *The Quarterly Journal of Economics* **106**(2), 651–667.

Kydland, F. E. & Prescott, E. C. (1977), 'Rules rather than discretion: The inconsistency of optimal plans', *Journal of political economy* **85**(3), 473–491.

Lewis, J. B., DeVine, B., Pitcher, L. & Martis., K. C. (2013), Digital boundary definitions of u.s. congressional districts, 1789-2012, Technical report, UCLA.

Mach, K. J. & Siders, A. (2021), 'Reframing strategic, managed retreat for transformative climate adaptation', *Science* **372**(6548), 1294–1299.

Maidment, D. R. (2017), 'Conceptual framework for the national flood interoperability experiment', *JAWRA Journal of the American Water Resources Association* **53**(2), 245–257.

Maidment, D. R. et al. (1993), *Handbook of hydrology*, Vol. 9780070, McGraw-Hill New York.

Maskin, E. & Sjöström, T. (2002), 'Implementation theory', *Handbook of social Choice and Welfare* **1**, 237–288.

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A. & Saleska, S. (2011), 'Height above the nearest drainage–a hydrologically relevant new terrain model', *Journal of Hydrology* **404**(1-2), 13–29.

Orsi, J. (2004), *Hazardous metropolis: Flooding and urban ecology in Los Angeles*, Univ of California Press.

Osborne, M. J. & Rubinstein, A. (1994), *A course in game theory*, MIT press.

Peltzman, S. (1975), 'The effects of automobile safety regulation', *Journal of political Economy* **83**(4), 677–725.

Rappaport, J. & Sachs, J. D. (2003), 'The united states as a coastal nation', *Journal of Economic growth* **8**, 5–46.

Redding, S. J. & Rossi-Hansberg, E. (2017), 'Quantitative spatial economics', *Annual Review of Economics* **9**, 21–58.

Rosenthal, S. S. & Strange, W. C. (2004), Evidence on the nature and sources of agglomeration economies, *in* 'Handbook of regional and urban economics', Vol. 4, Elsevier, pp. 2119–2171.

Saiz, A. (2010), 'The geographic determinants of housing supply', *The Quarterly Journal of Economics* **125**(3), 1253–1296.

Sargent, T. J. & Vilmunen, J. (2013), *Macroeconomics at the service of public policy*, Oxford University Press.

Schelling, T. C. (1971), 'Dynamic models of segregation', *Journal of mathematical sociology* **1**(2), 143–186.

Sieg, H., Smith, V. K., Banzhaf, H. S. & Walsh, R. J. (2000), 'Estimating the general equilibrium benefits of large policy changes: the clean air act revisited'.

Sims, C. A. (2002), 'Solving linear rational expectations models', *Computational economics* **20**(1-2), 1.

Uhlig, H. (1999), A toolkit for analyzing nonlinear dynamic stochastic models easily, *in* 'Computational Methods for the Study of Dynamic Economies', Oxford University Press, pp. 30–61.

Vose, R. S., Applequist, S., Squires, M., Durre, I., Menne, M. J., Williams Jr, C. N., Fenimore, C., Gleason, K. & Arndt, D. (2014), 'Improved historical temperature and precipitation time series for us climate divisions', *Journal of Applied Meteorology and Climatology* **53**(5), 1232–1251.

Woodford, M. (2003), 'Optimal interest-rate smoothing', *The Review of Economic Studies* **70**(4), 861–886.

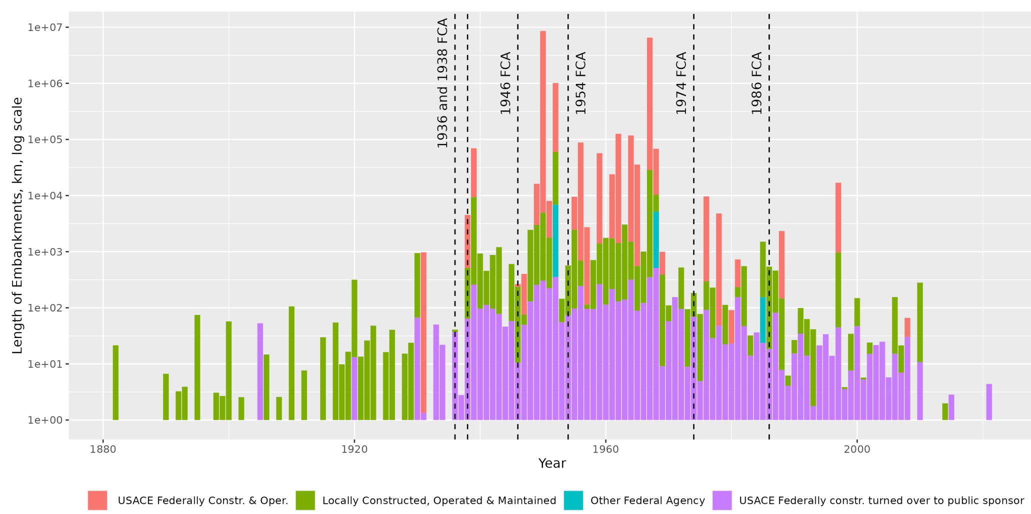
Young, R. & Hsiang, S. (2024), 'Mortality caused by tropical cyclones in the united states', *Nature* pp. 1–8.

Zheng, S. & Kahn, M. E. (2013), 'Does government investment in local public goods spur gentrification? evidence from beijing', *Real Estate Economics* **41**(1), 1–28.

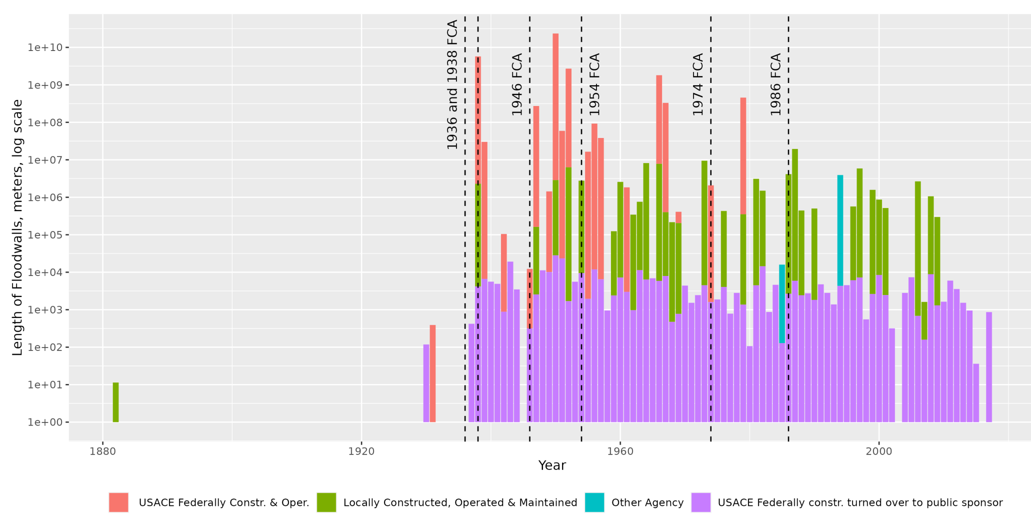
Figure 1: Flood Control Infrastructure and Flood Control Acts – Timeline of Embankments and Floodwalls

These figures present the length of embankments (upper panel) and floodwalls (lower panel) by date of construction. The colors correspond to the federal and/or local construction and operation. The dotted vertical lines mark the dates Flood Control Acts were signed into law by successive presidents. The green color is for levees that are locally constructed, operated, and maintained. This figure is discussed in Section 3.

(a) Timeline of the Construction of Embankments



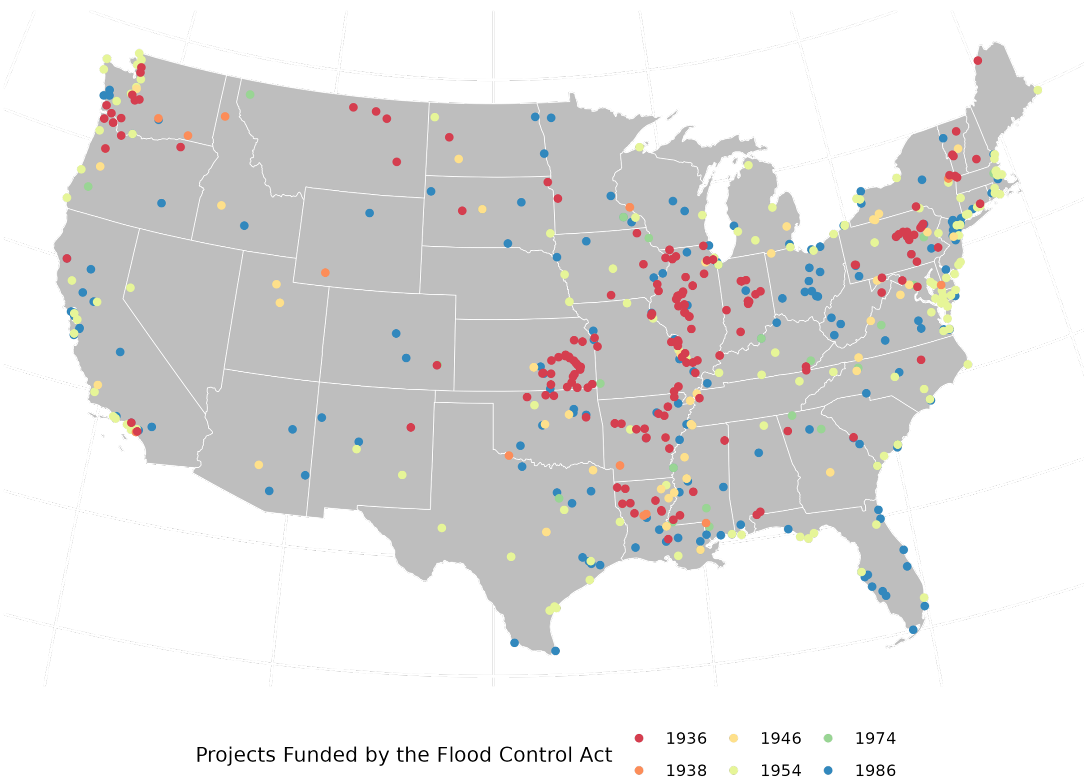
(b) Timeline of the Construction of Floodwalls



Source: National Levee Database, and calculations from the authors using the Coordinate Reference System of the U.S. National Atlas.

Figure 2: From 1936 to 1986: Projects Appropriated by Flood Control Acts

This map presents the location of projects funded by the 1936, 1938, 1946, 1954, 1974, 1986 Flood Control Acts. Each Act includes an approximate location of projects, with the nature of the project and an amount funded. The precise boundaries of embankments, floodwalls, and leveed areas are described in Section 4.3 and Figure 4.

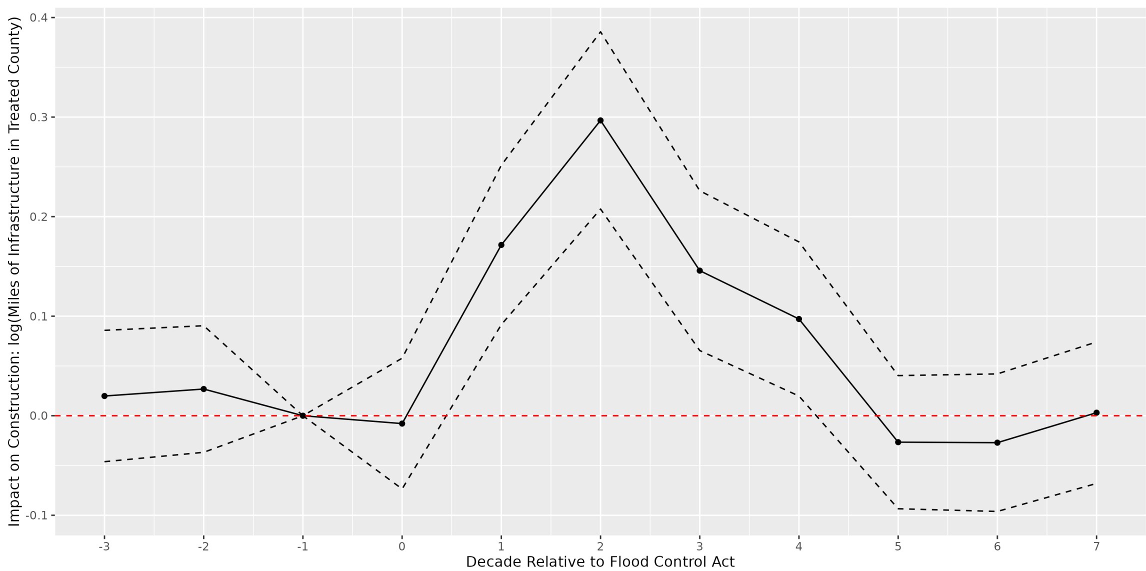


Source: Flood Control Act of 1936, HR 8455, June 22, 1936; Flood Control Act of 1938, HR 10618, June 28, 1968; Flood Control Act of 1946, HR 6597, July 24th, 1946; Flood Control Act of 1954, HR 9859, September 3rd, 1954; Flood Control Act of 1974, also known as the Water Resources and Development Act, HR 10203, March 7th, 1974; Flood Control Act of 1986, also known as the Water Resources Development Act, HR 6, November 17th, 1986.

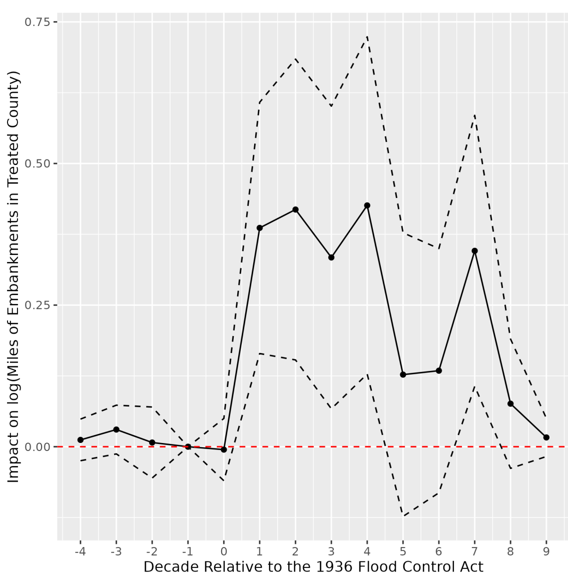
Figure 3: Event Study Analysis – Flood Control Act and Subsequent Miles of Embankments and Floodwalls

These figures present the coefficients of the event study regression, specification (36). Panel (a) presents the fixed effect regression with county, year, act fixed effects. Panel (b) weighs the regression by the inverse propensity score (Hirano et al. 2003), with variables presented on Tables 5 and 6 and on Figure A15. *Panel (c) focuses on embankments only. Dotted lines for the double-clustered 95% confidence intervals, where the clustering is at the year and county-levels.*

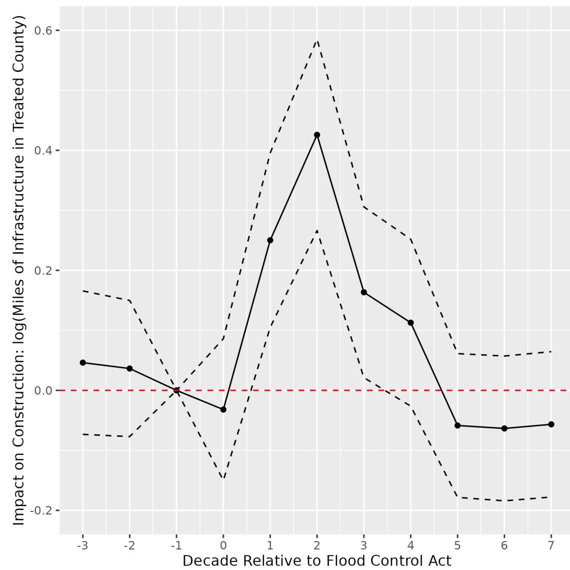
(a) All Infrastructure Pooled, All Acts



(b) Embankments, Inverse Propensity Score Weighting



(c) Embankments, Fixed Effect Panel Estimation

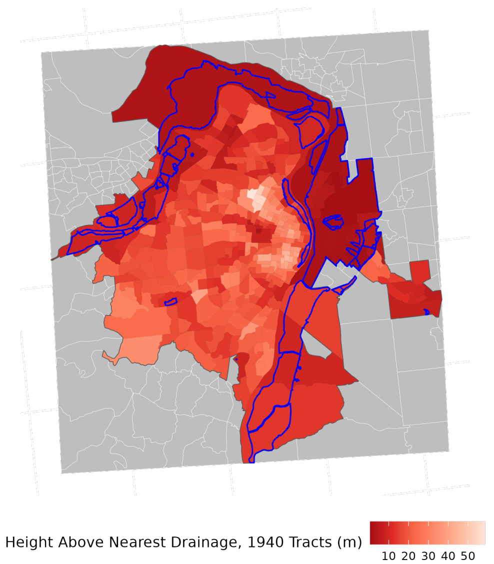


Source: length of embankments and floodwalls from the National Levee Database. Location of projects as in Figure 2, digitized from the Congressional Record.

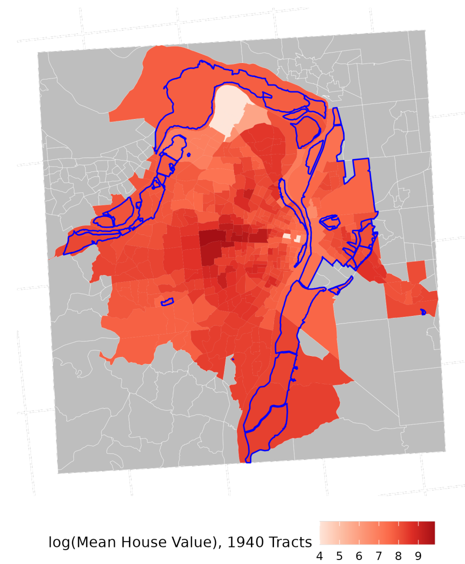
Figure 4: Spatial Selection of Leveed Areas

These maps present the boundaries of leveed areas, the Height Above Nearest Drainage, this paper's measure of flood risk, and the log Mean House Value in 1940, before levee construction. These maps provide illustration of the systematic results presented on Table 8.

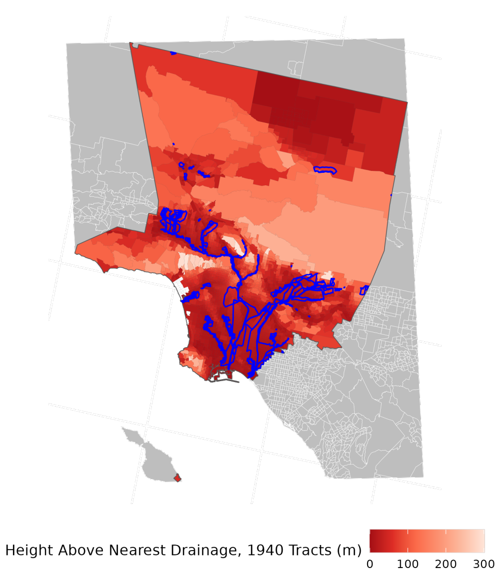
(a) St Louis, Height Above Nearest Drainage and Leveed Area



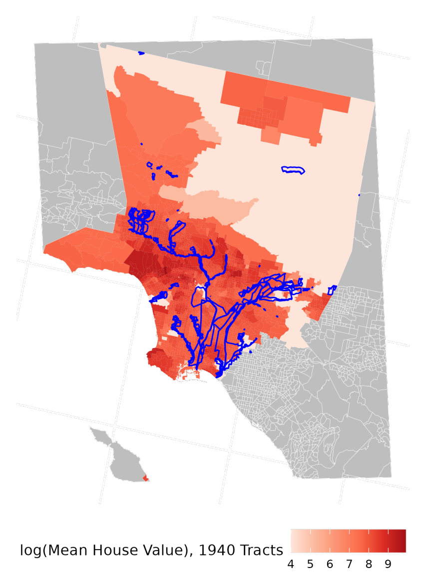
(b) St Louis, log(Mean House Value) and Leveed Area



(c) Los Angeles, Height Above Nearest Drainage and Leveed Area



(d) Los Angeles, log(Mean House Value) and Leveed Area

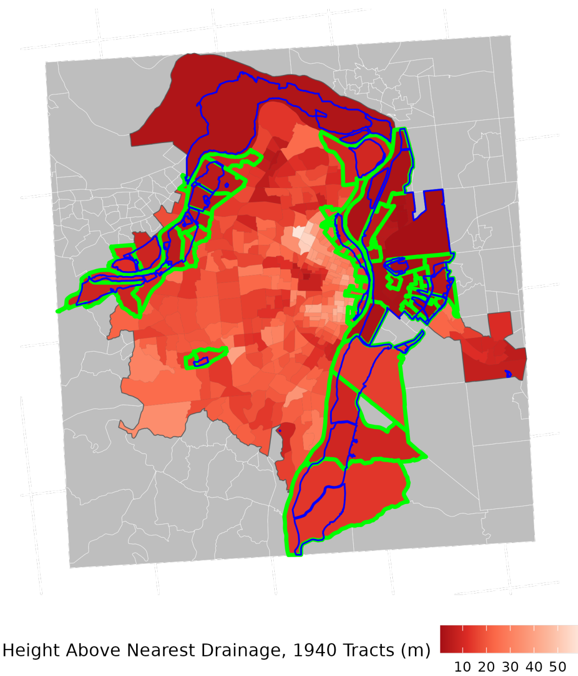


Sources: Mean House Value from the tract-level 1940 Census, National Historical GIS at the University of Minnesota. Height Above Nearest Drainage at 1/3 arcsecond resolution from the National Flood Interoperability Experiment.

Figure 5: Identification Strategies

These two figures illustrate the two identification strategies, beyond the panel tract f.e. approach of Table 10. Strategies #1a and #1b compare the evolution of the discontinuity in a distance buffer around the boundaries of leveed areas. Strategy #1a performs repeated cross-sectional regression discontinuities, while Strategy #1b pools the RDDs and controls for MSA×year fixed effects. Tracts within the buffer have green (—) boundaries. Leveed area boundaries are blue (—). Strategy #2 focuses on such Census tracts within the buffer for which there is no significant discontinuity in flood risk, as measured by the Height Above Nearest Drainage. In this case, flood risk is continuous at the boundary.

(a) Strategy #1 – Difference-in-Differences at the Regression Discontinuity



(b) Strategy #2 – Areas at Boundaries with no Discontinuity in Physical Flood Risk

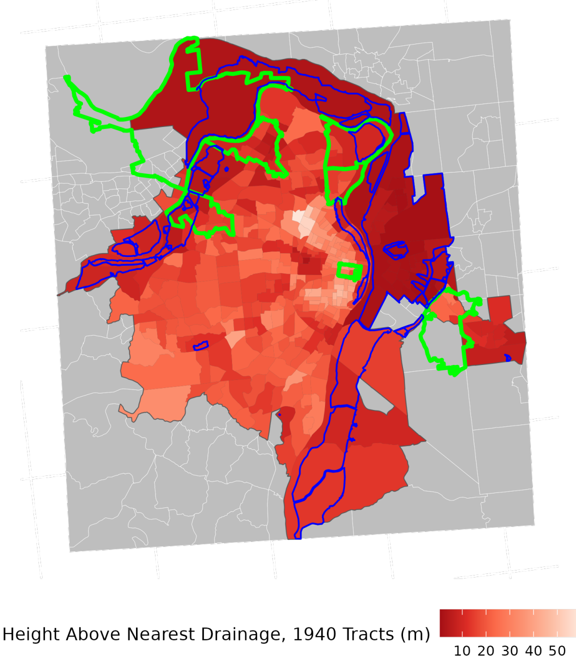


Figure 6: Structural Model – Impulse Response of Neighborhood Prices and Demographics to One or Two Decades of Levee Protection

These figures present the impulse response functions when the city's floodplain is protected by the levee for one decade (solid line), or for two decades (dashed line). The next figure presents neighborhood-level impulse response maps.

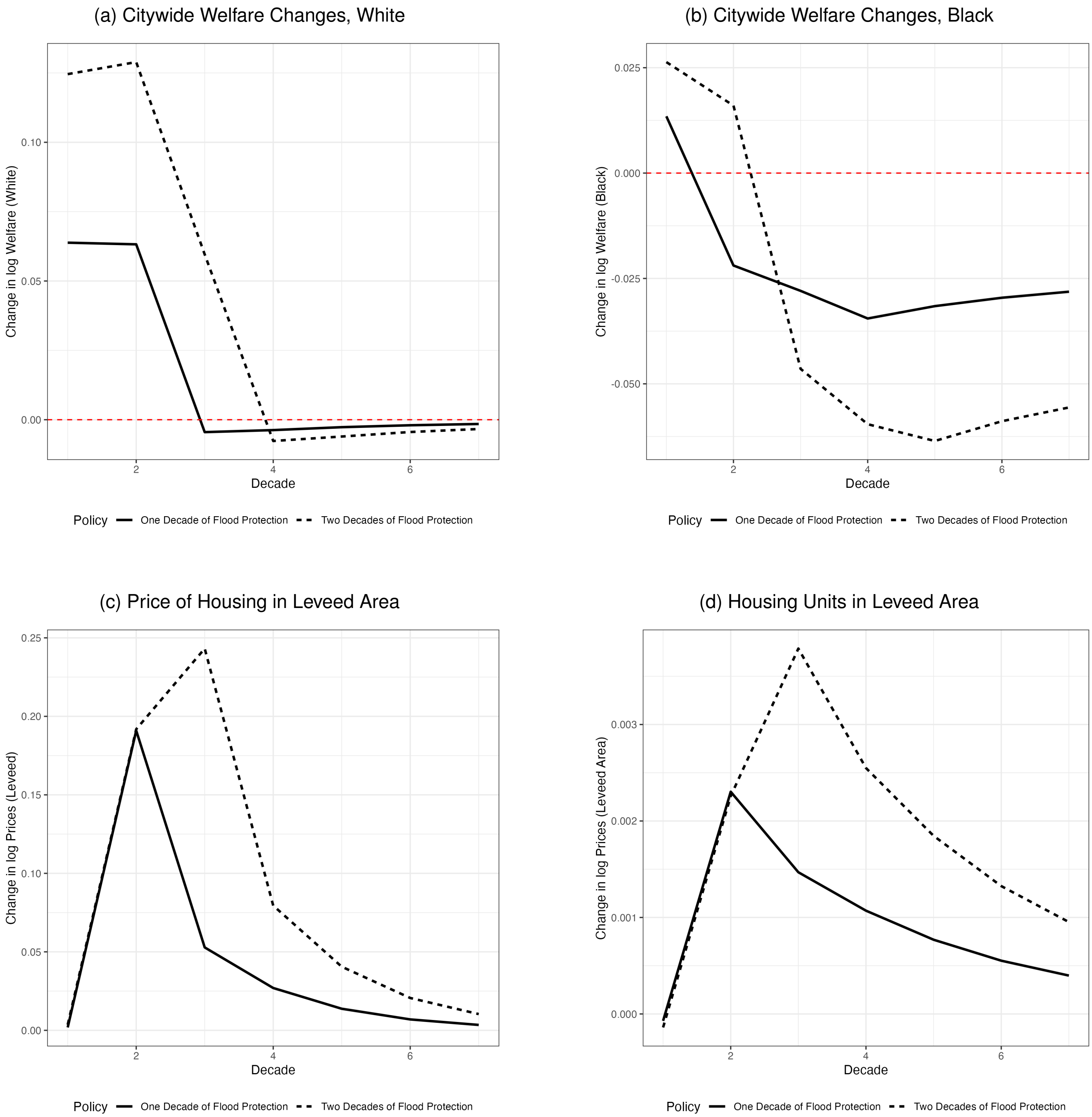
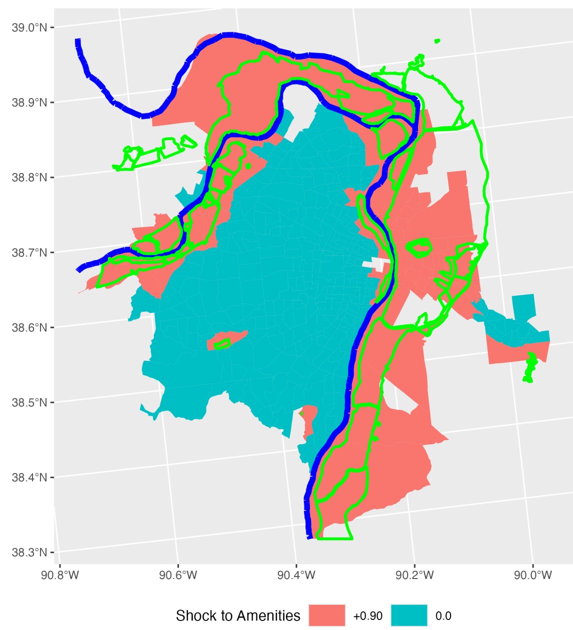


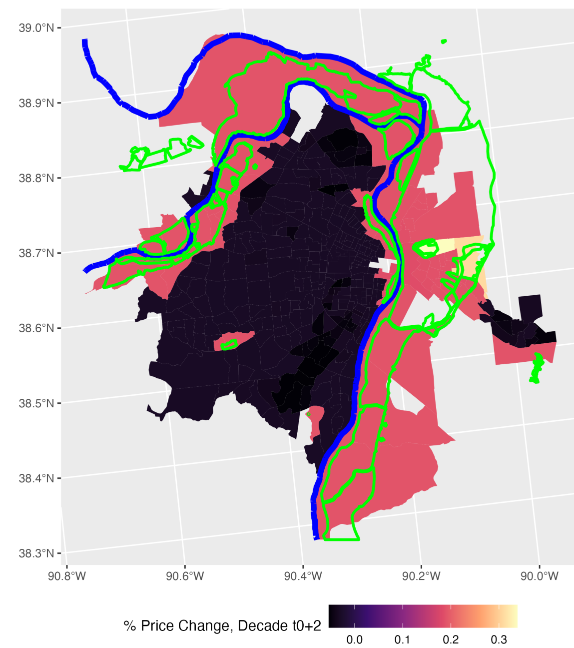
Figure 7: Structural Model – Impulse Response Maps of Neighborhood Prices and Demographics with One Decade of Levee Protection – Actual Levee Design

These figures present the impulse response maps after a one-period investment in flood protection (sub-figure (a)). Prices clear each local housing market and increase (sub-figure (b)) in response to population flows (sub-figure (c)). The leveed area attracts white households (sub-figure (d)).

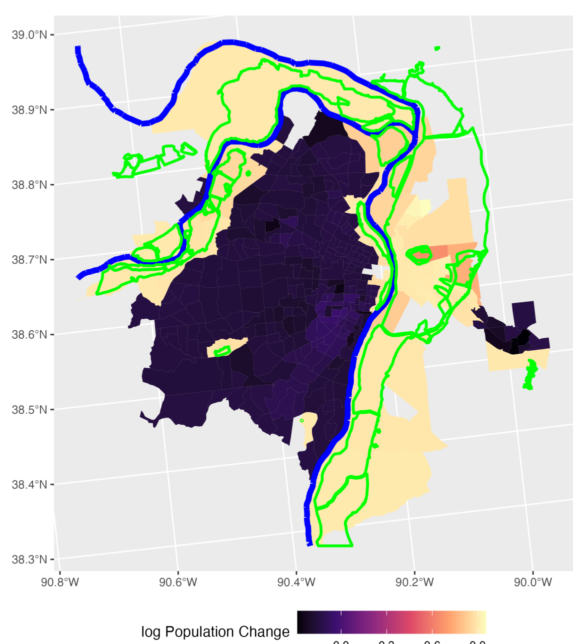
(a) Map of the Shock $\tilde{A}_{jt}$



(b) Price Changes



(c) Population Flows



(d) Change in Fraction Black

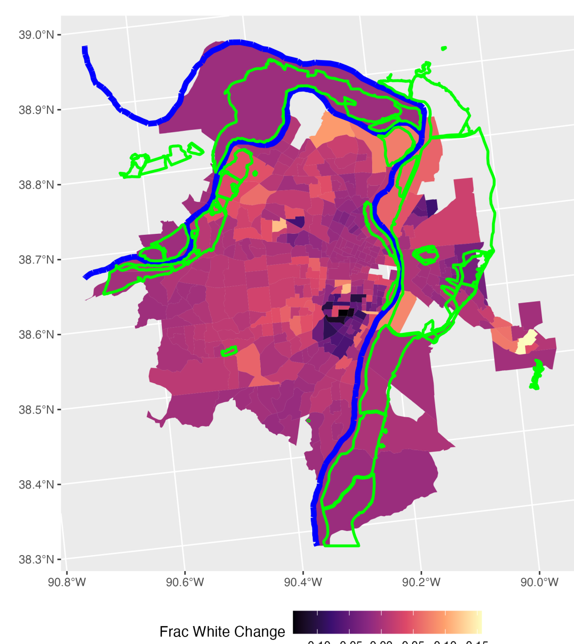
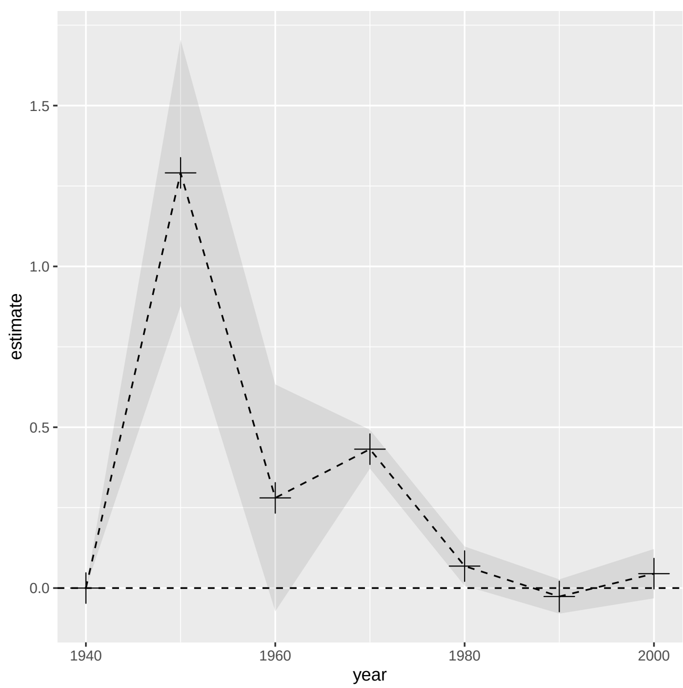


Figure 8: Estimation of the Impact of Levee Protection on Neighborhood-Level Intertemporal Values

These graphs present the estimates of the impact of a leveed area on intertemporal values, either with only tract $\times$ race and year fixed effects, or when controlling for log prior population, log prior fraction black, future values and log prices. Thus the left hand side graph is the sequence of coefficients ϕ_τ in specification (48). The right hand side graph is the sequence of coefficients ζ_τ in specification (47) after controlling.

(a) Without Controls for Path Dependence



(b) With Controls for Path Dependence

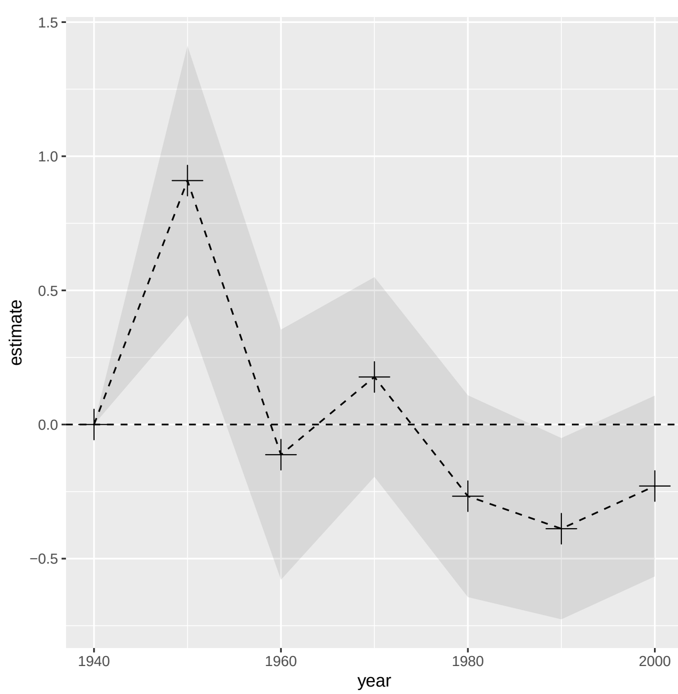


Figure 9: Welfare-Maximizing Levee Design vs Actual Design —

This figure maps the design that maximizes intertemporal welfare gains at $t = 0$, accounting for general equilibrium responses. The shades of blue represent the predicted values $P(\ell_{jt} = 1)$ from the selection equation. The green lines are for the boundaries of the actual leveed area. The blue thick lines are the Mississippi and the Missouri rivers. Grey boundaries are Census tracts.

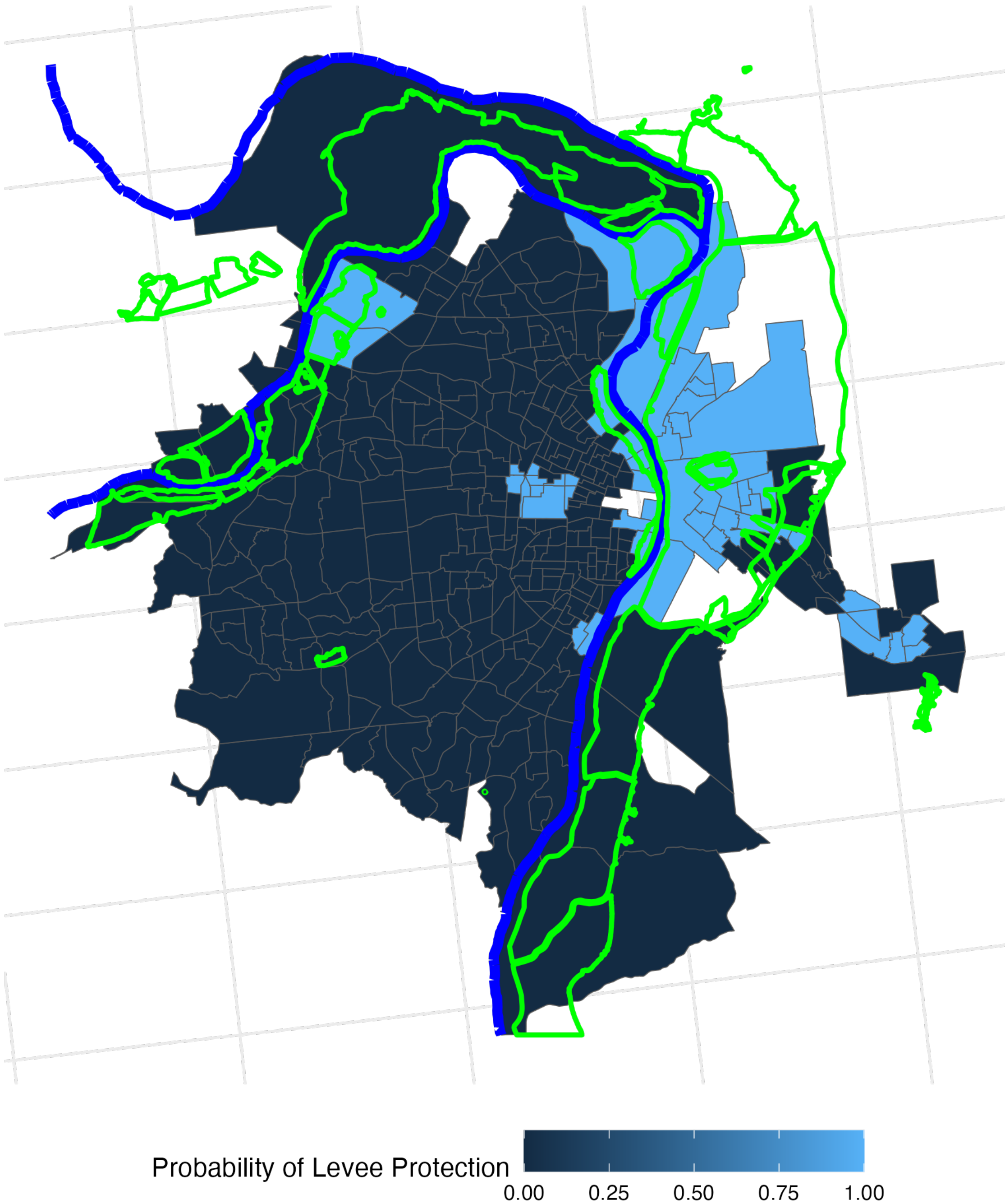


Figure 10: Comparing Impulse Responses of Welfare with the Optimal Design and with the Actual Design Chosen by the US Army Corps

These impulse response functions compare the welfare gains, the price changes, and the changes in population concentration, with the welfare-maximizing levee design (solid line) and with the actual levee design (dashed line).

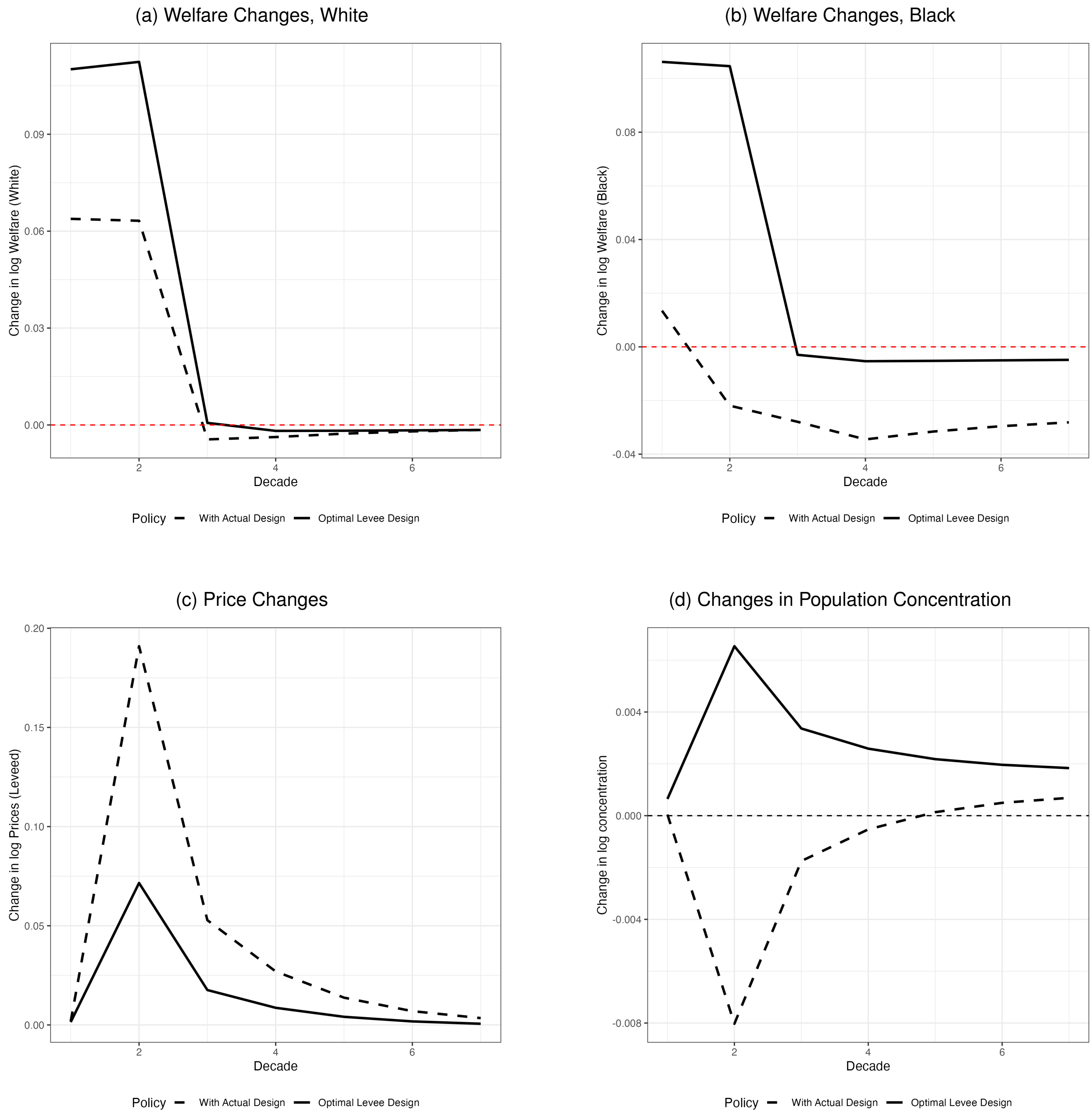
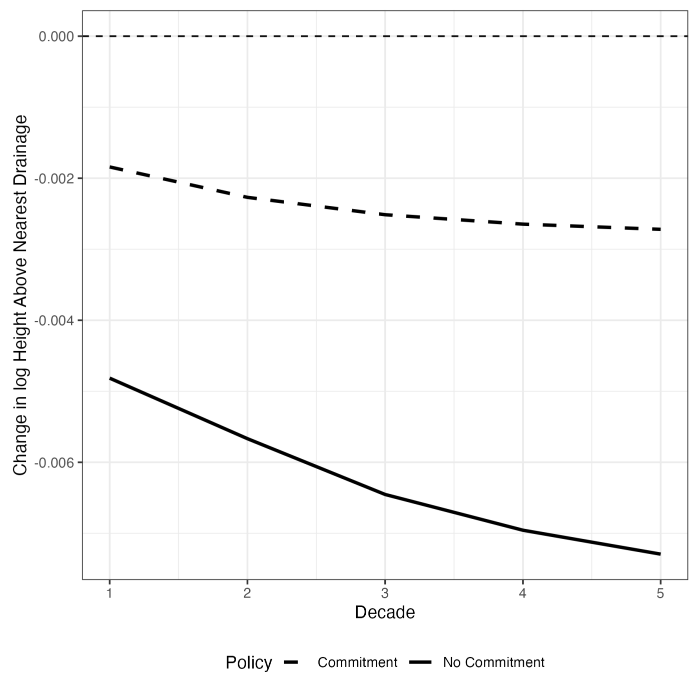


Figure 11: Commitment and Time Consistent Cases

These figures present the evolution of the average exposure to flood risk (top panel) and the evolution of the price of housing in leveed areas (bottom panel) in the case where the policymaker chooses the welfare-maximizing investment with commitment in the initial period (dashed line), and in the case where the policymaker chooses the welfare-maximizing investment in each period t (solid line).

(a) Exposure to Flood Risk – Log Change in Height Above Nearest Drainage



(b) Log Price increase in the Leveed Area

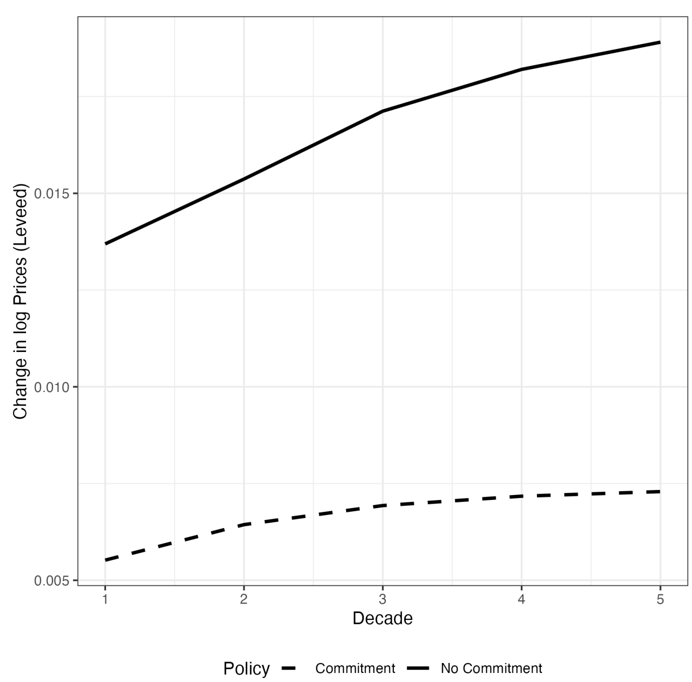
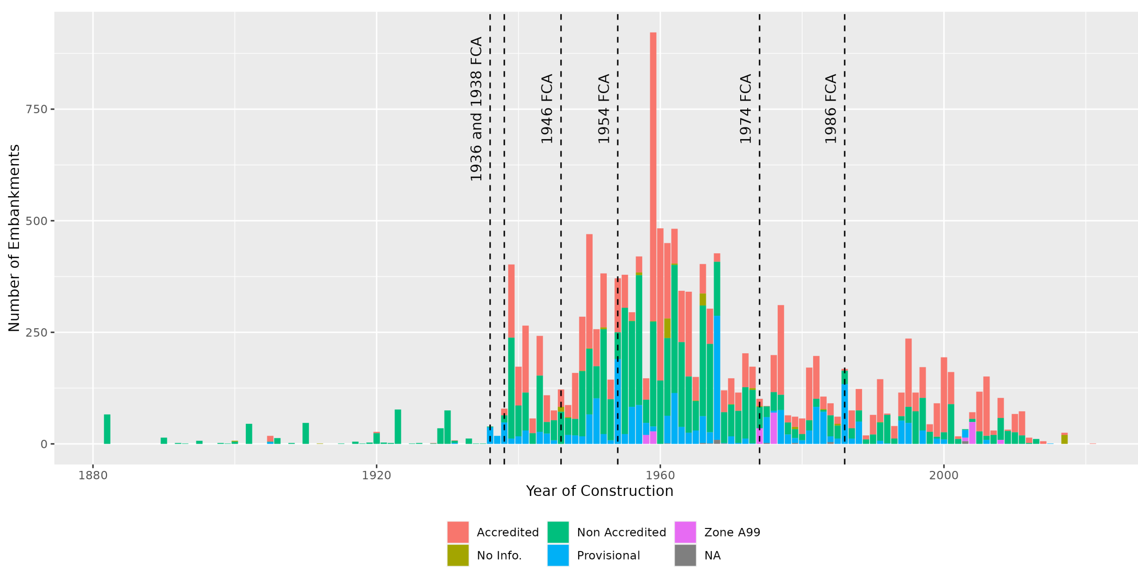


Figure 12: Current Accreditation of Past Infrastructure

These two timelines show the number of embankments and floodwalls by 2020 accreditation status. A non-accredited levee system is, according to the US Army Corps, "a levee system that does not meet the requirements in the NFIP regulations at Title 44, Chapter 1, Section 65.10 of the Code of Federal Regulations (44CFR§65.10), Mapping of Areas Protected by Levee Systems [...]." Zone A99: "areas subject to inundation by the 1-percent-annual-chance flood event, but which will ultimately be accredited upon completion of an under-construction levee system." Provisional: "FEMA is awaiting data and/or documentation that will demonstrate the levee system's compliance."

(a) Embankments



(b) Floodwalls

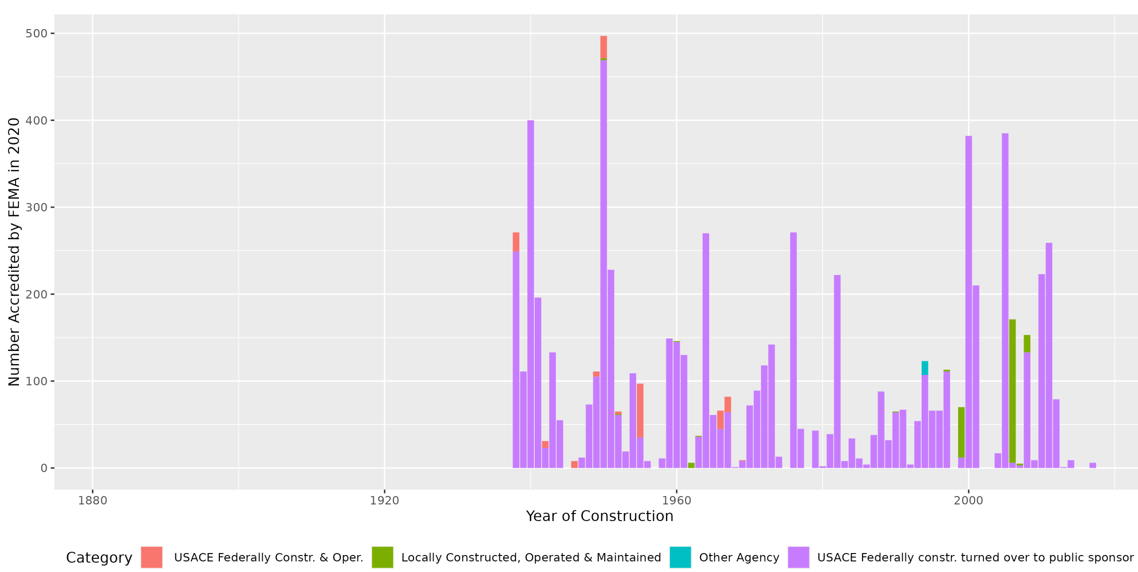


Figure 13: Peak Streamflows for Water Gages of the Los Angeles-Long Beach-Anaheim Metropolitan Area

This figure presents the sequence of peak streamflows of water gauges within leveed areas of the Los Angeles metro area. Source: Instantaneous water data for the Nation, USGS. Leveed areas from National Levee Database, US Army Corps of Engineers.

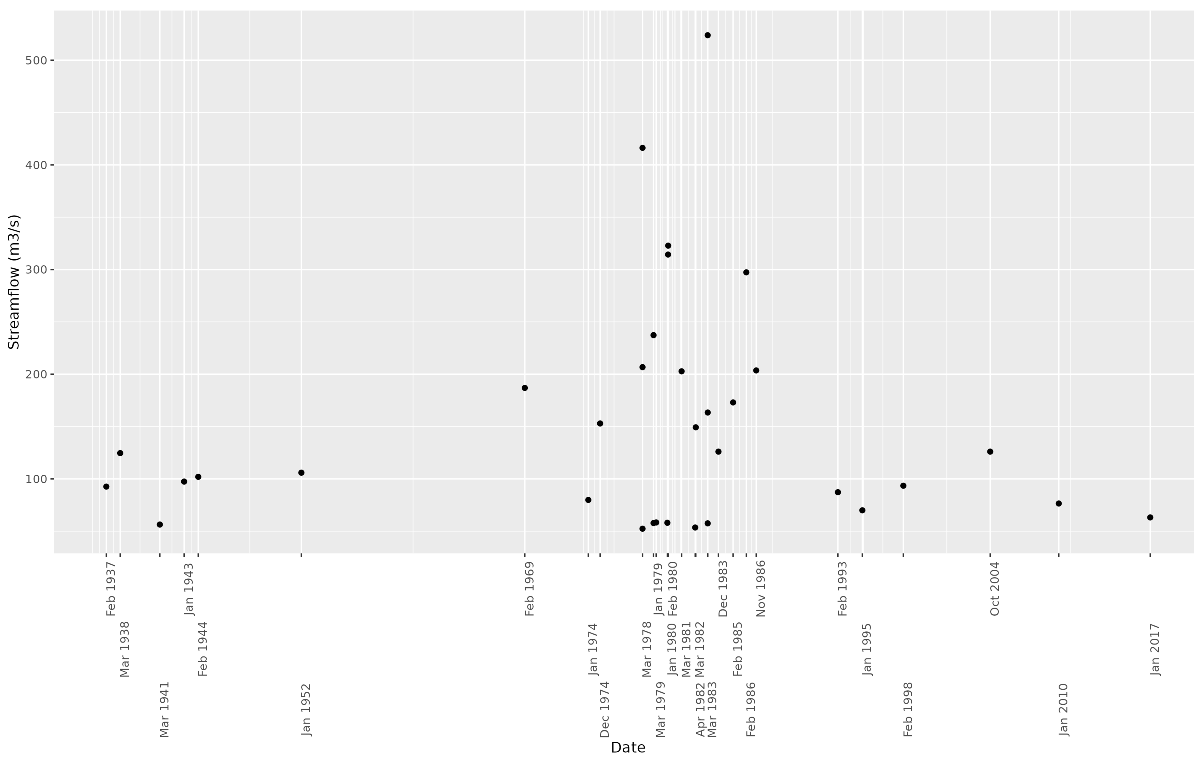
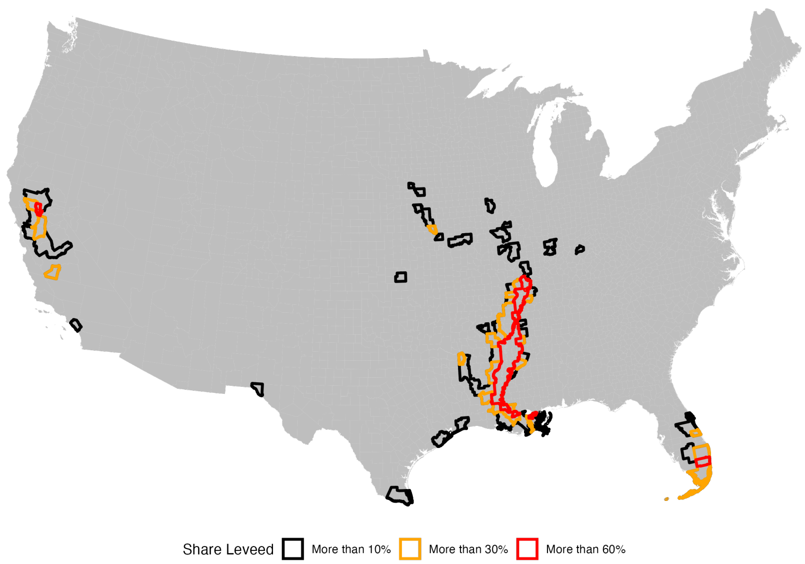


Figure 14: Presidential Declarations in Leveed Counties

The map of the upper panel presents the location of 'leveed' counties, i.e. counties for which the leveed area is more than 10, 30, or 60% of the surface area of the county. The chart of the lower panel presents descriptive statistics on the number of disaster declarations within leveed areas and overall.

(a) Surface Area Leveed, by County



(b) Disaster Declarations

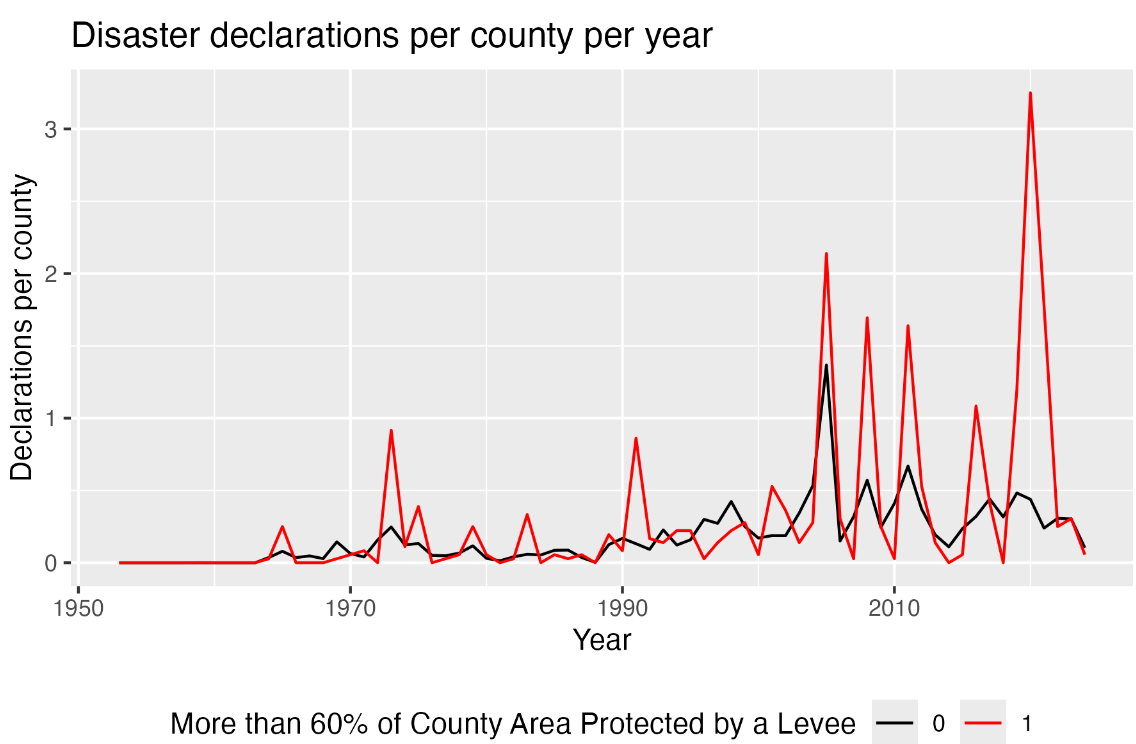


Table 2: The Impact of Congressional Appropriations - Flood Control Act and Subsequent Miles of Embankments and Floodwalls – All Acts

This table presents the impact of Flood Control Acts of 1936, 1938, 1946, 1954, 1974, 1986 on the length of infrastructure built in each particular decade. Regressions pool observations from each act. Column (1) is for embankments and floodwalls; column (2) for embankments, and column (3) for floodwalls. Coefficients are interpreted as semi-elasticities. These coefficients measure flows of newly built infrastructure.

Dependent Variable:	asinh(length_infra)		
Model:	(1)	(2)	(3)
Variables			
within_buffer_law × decade_m3	0.0198	0.0462	-0.0066
	(0.0337)	(0.0610)	(0.0285)
within_buffer_law × decade_m2	0.0268	0.0364	0.0172
	(0.0324)	(0.0579)	(0.0293)
within_buffer_law × decade_0	-0.0079	-0.0320	0.0162
	(0.0335)	(0.0602)	(0.0294)
within_buffer_law × decade_1	0.1716***	0.2502***	0.0929***
	(0.0408)	(0.0742)	(0.0342)
within_buffer_law × decade_2	0.2966***	0.4258***	0.1675***
	(0.0454)	(0.0814)	(0.0402)
within_buffer_law × decade_3	0.1457***	0.1635**	0.1280***
	(0.0410)	(0.0726)	(0.0379)
within_buffer_law × decade_4	0.0972**	0.1128	0.0816**
	(0.0395)	(0.0709)	(0.0347)
within_buffer_law × decade_5	-0.0266	-0.0585	0.0054
	(0.0341)	(0.0611)	(0.0304)
within_buffer_law × decade_6	-0.0271	-0.0634	0.0093
	(0.0352)	(0.0616)	(0.0344)
within_buffer_law × decade_7	0.0030	-0.0566	0.0627*
	(0.0363)	(0.0618)	(0.0380)
Fixed-effects			
county_fips_infra_act	Yes	Yes	Yes
decade_infra_act	Yes	Yes	Yes
Fit statistics			
Observations	294,428	147,214	147,214
R^2	0.19241	0.19269	0.16897
Within R^2	0.00116	0.00153	0.00111

Clustered (county_fips_infra_act) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3: The Impact of Congressional Appropriations - Flood Control Act and Subsequent Miles of Embankments and Floodwalls – 1936 Act

*These eight regressions perform the event study separately by Act and by type of infrastructure: embankments in columns (1) and (3), floodwalls in columns (2) and (4). We present the acts with the largest number of projects funded for the sake of clarity. Clustered (county) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	asinh(length_infra)			
Model:	(1)	(2)	(3)	(4)
Variables				
within_buffer_law × decade_m4	-0.0015	0.0156	-0.0015	0.0156
	(0.0015)	(0.0271)	(0.0015)	(0.0271)
within_buffer_law × decade_m3	0.000***	0.0776	0.000***	0.0776
	(0.000)	(0.0476)	(0.000)	(0.0476)
within_buffer_law × decade_m2	0.000***	0.0635	0.000***	0.0635
	(0.000)	(0.0461)	(0.000)	(0.0461)
within_buffer_law × decade_0	0.0126	0.0397	0.0126	0.0397
	(0.0126)	(0.0454)	(0.0126)	(0.0454)
within_buffer_law × decade_1	0.1710***	0.5864***	0.1710***	0.5864***
	(0.0555)	(0.1199)	(0.0555)	(0.1199)
within_buffer_law × decade_2	0.2524***	1.011***	0.2524***	1.011***
	(0.0746)	(0.1657)	(0.0746)	(0.1657)
within_buffer_law × decade_3	0.2387***	0.8006***	0.2387***	0.8006***
	(0.0699)	(0.1485)	(0.0699)	(0.1485)
within_buffer_law × decade_4	0.3344***	0.8821***	0.3344***	0.8821***
	(0.0769)	(0.1610)	(0.0769)	(0.1610)
within_buffer_law × decade_5	0.0910*	0.2601**	0.0910*	0.2601**
	(0.0481)	(0.1075)	(0.0481)	(0.1075)
within_buffer_law × decade_6	0.1109**	0.2421***	0.1109**	0.2421***
	(0.0539)	(0.0937)	(0.0539)	(0.0937)
within_buffer_law × decade_7	0.2472***	0.4035***	0.2472***	0.4035***
	(0.0691)	(0.1026)	(0.0691)	(0.1026)
within_buffer_law × decade_8	0.0636*	0.1110	0.0636*	0.1110
	(0.0359)	(0.0683)	(0.0359)	(0.0683)
within_buffer_law × decade_9	0.0271	0.0627	0.0271	0.0627
	(0.0223)	(0.0407)	(0.0223)	(0.0407)
Fixed-effects				
county_fips	Yes	Yes	Yes	Yes
decade	Yes	Yes	Yes	Yes
Fit statistics				
Observations	37,305	37,305	37,305	37,305
R^2	0.12573	0.15745	0.12573	0.15745
Within R^2	0.00678	0.01247	0.00678	0.01247

Clustered (county_fips) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4: The Impact of Congressional Appropriations - Flood Control Act and Subsequent Miles of Embankments and Floodwalls – 1974

*These eight regressions perform the event study separately by Act and by type of infrastructure: embankments in columns (1) and (3), floodwalls in columns (2) and (4). We present the acts with the largest number of projects funded for the sake of clarity. Clustered (county) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	asinh(length_infra)			
Model:	(1)	(2)	(3)	(4)
Variables				
within_buffer_law × decade_m8	-0.1124*	-0.1829	-0.1124*	-0.1829
	(0.0620)	(0.1206)	(0.0620)	(0.1206)
within_buffer_law × decade_m7	-0.1109*	-0.1877	-0.1109*	-0.1877
	(0.0619)	(0.1207)	(0.0619)	(0.1207)
within_buffer_law × decade_m6	-0.1109*	-0.1164	-0.1109*	-0.1164
	(0.0619)	(0.1469)	(0.0619)	(0.1469)
within_buffer_law × decade_m5	-0.1109*	0.0573	-0.1109*	0.0573
	(0.0619)	(0.1505)	(0.0619)	(0.1505)
within_buffer_law × decade_m4	-0.1109*	-0.1319	-0.1109*	-0.1319
	(0.0619)	(0.1495)	(0.0619)	(0.1495)
within_buffer_law × decade_m3	-0.0230	0.1473	-0.0230	0.1473
	(0.0978)	(0.2312)	(0.0978)	(0.2312)
within_buffer_law × decade_m2	-0.0268	-0.1084	-0.0268	-0.1084
	(0.1165)	(0.1910)	(0.1165)	(0.1910)
within_buffer_law × decade_0	0.0525	-0.1652	0.0525	-0.1652
	(0.1113)	(0.1809)	(0.1113)	(0.1809)
within_buffer_law × decade_1	0.0374	-0.0326	0.0374	-0.0326
	(0.1178)	(0.1986)	(0.1178)	(0.1986)
within_buffer_law × decade_2	0.1111	0.0377	0.1111	0.0377
	(0.1443)	(0.2123)	(0.1443)	(0.2123)
within_buffer_law × decade_3	0.2619	0.0834	0.2619	0.0834
	(0.1723)	(0.1988)	(0.1723)	(0.1988)
within_buffer_law × decade_4	0.0071	0.0223	0.0071	0.0223
	(0.0928)	(0.1728)	(0.0928)	(0.1728)
within_buffer_law × decade_5	-0.1148*	-0.1888	-0.1148*	-0.1888
	(0.0621)	(0.1205)	(0.0621)	(0.1205)
Fixed-effects				
county_fips	Yes	Yes	Yes	Yes
decade	Yes	Yes	Yes	Yes
Fit statistics				
Observations	32,670	32,670	32,670	32,670
R^2	0.11998	0.13741	0.11998	0.13741
Within R^2	0.00298	0.00064	0.00298	0.00064

Clustered (county_fips) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 5: Drivers of Congressional Appropriations - Flood Control Act Funding, Weather, and Political Connections

These two sets of regressions correlate 1936 Flood Control Act funding with measures of political connections and past weather. These regressions, together with the results of the next table, build a propensity score. Figure A15 provides a graphical illustration and maps of the data.

Dependent Variable:	within_buffer_law			
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	0.1984***	0.0051	0.0181	0.2127***
	(0.0219)	(0.0297)	(0.0275)	(0.0221)
democratic	-0.0267	-0.0625**	-0.0744***	-0.0436*
	(0.0251)	(0.0249)	(0.0250)	(0.0254)
committee	0.1501***	0.1282***	0.1449***	0.1463***
	(0.0412)	(0.0405)	(0.0403)	(0.0411)
p90_precip_before		0.0403***		
		(0.0043)		
mean_precip_before			0.0735***	
			(0.0071)	
mean_precip_year_of_demeaned				0.0267***
				(0.0064)
committee × mean_precip_year_of_demeaned				0.0094
				(0.0503)
Fit statistics				
Observations	2,378	2,377	2,377	2,377
R^2	0.00627	0.04226	0.04977	0.01384
Adjusted R^2	0.00544	0.04105	0.04856	0.01218

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6: Drivers of Congressional Appropriations - Flood Control Act Funding, Unemployment and Farmland Value

These regressions correlate 1936 Flood Control Act funding with measures of unemployment, the value of farmland, the share urban, and the share of Black population. These regressions, together with the results of the previous table, build a propensity score. Figure A15 provides a graphical illustration and maps of the data.

Dependent Variable:	within_buffer_law			
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	-0.4598***		-0.4002***	
	(0.1009)		(0.1053)	
asinh(total_value_farmland_1930)	0.0096	0.0007	0.0089	0.0034
	(0.0064)	(0.0076)	(0.0067)	(0.0077)
asinh(unemployed_1930)	0.0686***	0.0457**	0.0587***	0.0349**
	(0.0064)	(0.0170)	(0.0073)	(0.0156)
share_black_1930			-0.0221	0.0904
			(0.0462)	(0.1004)
share_urban_1930			0.1264***	0.0967**
			(0.0372)	(0.0400)
Fixed-effects				
statefips		Yes		Yes
Fit statistics				
Observations	2,307	2,307	2,307	2,307
R^2	0.05829	0.28928	0.06341	0.29225
Within R^2		0.02314		0.02722

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 7: Metropolitan Areas of the 1940–2010 Longitudinal Tract-Level Sample, by 1940 Population in Leveed Area

This table presents the list of metropolitan areas of the 1940–2010, ranked in decreasing order of the population within areas protected by levees. The aggregate house value is in thousands of current 1940 dollars.

		in Leveed Area		
Metro Area	1940 Population	% of Pop.	1940 Agg. House Value ('000)	% of Agg. House Value
Los Angeles-Long Beach-Anaheim, CA	800,427	28.8	783,276	21.3
New Orleans-Metairie, LA	494,395	100.0	367,249	100.0
Louisville/Jefferson County, KY-IN	310,319	92.8	246,974	87.8
St. Louis, MO-IL	270,151	21.8	175,984	12.8
Cincinnati, OH-KY-IN	150,980	24.3	157,587	19.6
Memphis, TN-MS-AR	128,958	42.2	64,187	28.4
Dayton, OH	123,448	56.5	117,389	48.5
Kansas City, MO-KS	117,573	29.5	89,581	25.4
Indianapolis-Carmel-Anderson, IN	110,496	24.7	94,148	21.4
Providence-Warwick, RI-MA	99,291	39.3	79,047	30.1
Washington-Arlington-Alexandria, DC-VA-MD-WV	88,033	13.3	119,625	8.3
Hartford-West Hartford-East Hartford, CT	86,003	34.7	83,255	28.7
Des Moines-West Des Moines, IA	72,464	45.3	51,675	39.2
Dallas-Fort Worth-Arlington, TX	68,888	21.6	48,740	14.8
New York-Newark-Jersey City, NY-NJ-PA	62,647	0.7	63,775	0.5
Augusta-Richmond County, GA-SC	51,558	75.4	24,799	58.4
Richmond, VA	46,987	24.3	25,270	12.5
Columbus, OH	42,512	13.9	37,273	11.0
Milwaukee-Waukesha-West Allis, WI	35,516	6.0	52,924	8.1
Minneapolis-St. Paul-Bloomington, MN-WI	32,358	3.8	27,235	3.1

Table 8: Within-City Spatial Selection of Projects – Flood Risk And Levee Protection

This table estimates the correlation (conditional on state fixed effects) between 1940 tract characteristics and subsequent flood protection by a levee. The HAND is the Height Above Nearest Drainage (Nobre et al. 2011), in meters, and is inversely related to flood risk. A 10 meter increase in the HAND reduces the probability of being in a leveed area by 1.2 to 2.4 percentage points. Standard errors are clustered at the state level.

Dependent Variable:			in_levee			
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	0.1486**			0.1932**		
	(0.0604)			(0.0872)		
HAND	-0.0012**	-0.0020***	-0.0021***			
	(0.0006)	(0.0006)	(0.0006)			
HAND_cut(8.13,17.3]				-0.0706	-0.0466**	-0.0623**
				(0.0572)	(0.0215)	(0.0263)
HAND_cut(17.3,33.9]				-0.1050	-0.1060**	-0.1329**
				(0.0663)	(0.0508)	(0.0572)
HAND_cut(33.9,308]				-0.1283*	-0.1739**	-0.2016**
				(0.0717)	(0.0832)	(0.0884)
Fixed-effects						
‘State f.e.‘		Yes			Yes	
CBSAFP			Yes			Yes
Fit statistics						
Observations	12,560	12,560	12,560	12,530	12,530	12,530
R^2	0.01127	0.29480	0.34931	0.02261	0.29780	0.35935
Within R^2		0.03612	0.04020		0.04197	0.05577

Clustered (CBSAFP) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:			in_levee			
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	0.1385**			0.1430**		
	(0.0611)			(0.0556)		
HAND	-0.0011**	-0.0020***	-0.0021***	-0.0012**	-0.0020***	-0.0021***
	(0.0005)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
‘Frac.Black‘	0.1444	0.0893**	0.0886**			
	(0.1361)	(0.0397)	(0.0415)			
‘BlackPopulationFlow19101940‘				0.0917	0.1542	0.0539
				(0.2428)	(0.2248)	(0.0896)
Fixed-effects						
‘State f.e.‘		Yes			Yes	
CBSAFP			Yes			Yes
Fit statistics						
Observations	12,558	12,558	12,558	12,220	12,220	12,220
R^2	0.01609	0.29620	0.35064	0.01249	0.29321	0.34722
Within R^2		0.03805	0.04218		0.03751	0.04055

Clustered (CBSAFP) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 9: Within-City Spatial Selection of Projects – Flood Risk And Levee Protection – Race

This table estimates the correlation (conditional on state fixed effects) between 1940 Census tract characteristics and subsequent flood protection by a levee. The HAND is the Height Above Nearest Drainage (Nobre et al. 2011) is inversely related to flood risk. A 10 meter increase in the HAND reduces the probability of being in a leveed area by 1.2 to 2.4 percentage points. Standard errors are clustered at the state level.

Dependent Variable:	in_levee					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	0.1596**			0.1498**		
	(0.0654)			(0.0625)		
HAND	-0.0012**	-0.0020***	-0.0021***	-0.0012**	-0.0020***	-0.0021***
	(0.0006)	(0.0006)	(0.0006)	(0.0005)	(0.0006)	(0.0006)
‘BlackPopulationFlow19301940‘	-0.8415	0.3917	0.2632			
	(0.7477)	(0.4698)	(0.3717)			
‘Out-migration19101940‘				0.0437	-0.0003	-0.0052
				(0.0927)	(0.0267)	(0.0199)
Fixed-effects						
‘State f.e.‘		Yes			Yes	
CBSAFP			Yes			Yes
Fit statistics						
Observations	12,560	12,560	12,560	12,220	12,220	12,220
R^2	0.01435	0.29508	0.34941	0.01296	0.29236	0.34718
Within R^2		0.03650	0.04034		0.03634	0.04049

Clustered (CBSAFP) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	in_levee					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	0.3211**			0.1646		
	(0.1258)			(0.2805)		
HAND	-0.0013**	-0.0020***	-0.0021***	-0.0013**	-0.0020***	-0.0021***
	(0.0005)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
log(‘AverageValue‘)	-0.0212**	-0.0101*	-0.0100**	-0.0222**	-0.0099*	-0.0095*
	(0.0099)	(0.0053)	(0.0049)	(0.0103)	(0.0054)	(0.0050)
log(distance_to_center)				0.0181	-0.0054	-0.0132
				(0.0359)	(0.0104)	(0.0083)
Fixed-effects						
‘State f.e.‘		Yes			Yes	
CBSAFP			Yes			Yes
Fit statistics						
Observations	11,947	11,947	11,947	11,942	11,942	11,942
R^2	0.02005	0.29737	0.35375	0.02246	0.29806	0.35408
Within R^2		0.03959	0.04284		0.03970	0.04380

Clustered (CBSAFP) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 10: Long-Run Evolution of Neighborhoods – Panel Tract F.E. Approach

This table presents a descriptive approach to understanding the evolution of neighborhoods protected by levees compared to the evolution of neighborhoods outside of leveed areas in the 1940–2010 sample. The reference decade is 1940, i.e. coefficients are relative to the 1940 housing value, fraction black, or the 1940 population. Each regression controls for tract fixed effects and double clusters standard errors at the year and tract levels. Tables 11 to 14 present results using boundary discontinuity designs.

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Leveed × 1940	Ref.	Ref.	Ref.	Ref.	Ref.
Leveed × 1950	0.0191	0.0148***	0.1009***	0.0964***	0.1322**
	(0.0365)	(0.0037)	(0.0301)	(0.0299)	(0.0522)
Leveed × 1960	0.1458***	0.0047	0.0988**	0.0908**	0.2659***
	(0.0471)	(0.0062)	(0.0427)	(0.0406)	(0.0714)
Leveed × 1970	0.1240***	-0.0129*	0.1073**	0.1413***	0.3110***
	(0.0364)	(0.0076)	(0.0449)	(0.0447)	(0.0709)
Leveed × 1980	0.0748**	-0.0151*	0.0066	0.0084	0.1261*
	(0.0373)	(0.0079)	(0.0450)	(0.0449)	(0.0722)
Leveed × 1990	0.1101***	-0.0201***	-0.0577	-0.0769*	0.0728
	(0.0368)	(0.0076)	(0.0449)	(0.0455)	(0.0718)
Leveed × 2000	0.0965***	-0.0232***	-0.0916**	-0.1172**	0.0337
	(0.0367)	(0.0074)	(0.0457)	(0.0461)	(0.0717)
Leveed × 2010	0.0892**	-0.0260***	-0.1178**	-0.1433***	-0.0044
	(0.0355)	(0.0073)	(0.0461)	(0.0463)	(0.0703)
Fixed-effects					
Tract	Yes	Yes	Yes	Yes	Yes
County × Year	Yes	Yes	Yes	Yes	Yes
Observations	188,076	188,476	188,476	188,476	188,476
R^2	0.85137	0.79928	0.64815	0.68241	0.83919

Clustered (GEOID10) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 11: Regression Discontinuity at the Boundary of Leveed Areas – Repeated Cross Section, 0.25 mile buffer, Gaussian RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1940	-0.0637	0.0154***	-0.1671	-0.2331*	-0.1519
	(0.0556)	(0.0045)	(0.1264)	(0.1235)	(0.1106)
in_levee × year_1950	-0.1148	-3×10^{-5}	-0.0771	-0.1282	-0.0931
	(0.1050)	(0.0072)	(0.0926)	(0.0831)	(0.1388)
in_levee × year_1960	0.0079	-0.0309	0.0410	-0.0595	0.0764
	(0.1125)	(0.0216)	(0.0292)	(0.0396)	(0.1328)
in_levee × year_1970	0.1994***	-0.0414	0.0692	-0.0045	0.3972***
	(0.0618)	(0.0255)	(0.0521)	(0.0537)	(0.1253)
in_levee × year_1980	0.1500**	-0.0340	0.0392	-0.0134	0.3393**
	(0.0621)	(0.0216)	(0.0331)	(0.0342)	(0.1463)
in_levee × year_1990	0.1906***	-0.0336**	0.0286	-0.0207	0.3377***
	(0.0498)	(0.0154)	(0.0239)	(0.0271)	(0.1086)
in_levee × year_2000	0.1519***	-0.0306***	0.0655***	0.0094	0.3455***
	(0.0460)	(0.0107)	(0.0223)	(0.0209)	(0.0879)
in_levee × year_2010	-0.0026	-0.0233**	0.0794***	0.0365	0.1676**
	(0.0325)	(0.0094)	(0.0253)	(0.0241)	(0.0722)
Fixed-effects					
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	133,718	133,969	133,969	133,969	133,969
R^2	0.76690	0.38967	0.37854	0.40259	0.68555
Within R^2	0.00174	0.00106	0.00068	0.00083	0.00296

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 12: Regression Discontinuity at the Boundary of Leveed Areas – Combined Regression Discontinuity and Tract f.e., County $\times$Year f.e., 0.25 mile buffer, Gaussian RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1950	-0.0179	-0.0131	0.0826	0.0923	0.1061
	(0.1156)	(0.0151)	(0.1373)	(0.1312)	(0.2194)
in_levee × year_1960	0.1005	-0.0560***	0.2336*	0.2013*	0.3014*
	(0.0972)	(0.0153)	(0.1253)	(0.1218)	(0.1767)
in_levee × year_1970	0.2772***	-0.0699***	0.2509*	0.2380*	0.6018***
	(0.0632)	(0.0184)	(0.1311)	(0.1285)	(0.1678)
in_levee × year_1980	0.2706***	-0.0672***	0.2044	0.2168*	0.5985***
	(0.0634)	(0.0138)	(0.1289)	(0.1260)	(0.1737)
in_levee × year_1990	0.3008***	-0.0669***	0.2027*	0.2153*	0.5916***
	(0.0757)	(0.0123)	(0.1228)	(0.1191)	(0.1805)
in_levee × year_2000	0.2579***	-0.0616***	0.2376*	0.2454**	0.5978***
	(0.0657)	(0.0130)	(0.1244)	(0.1202)	(0.1665)
in_levee × year_2010	0.1641**	-0.0546***	0.2416*	0.2590**	0.4880***
	(0.0747)	(0.0160)	(0.1236)	(0.1194)	(0.1702)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	133,718	133,969	133,969	133,969	133,969
R^2	0.88001	0.84054	0.64955	0.67273	0.84528
Within R^2	0.00221	0.00226	0.00116	0.00111	0.00330

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 13: Regression Discontinuity at the Boundary of Leveed Areas – Combined Regression Discontinuity and Tract f.e., County $\times$Year f.e., 0.5 mile buffer, Gaussian RD Kernel

Dependent Variables: Model:	log(Housing Value) (1)	Frac. Black (2)	log(Population) (3)	log(Housing Units) (4)	log(Agg. Housing Value) (5)
Variables					
in_levee $\times$ year_1950	0.0240	-0.0038	0.0573	0.0549	0.0951
	(0.0947)	(0.0123)	(0.1090)	(0.1012)	(0.1848)
in_levee $\times$ year_1960	0.0796	-0.0357***	0.1319	0.1164	0.1956
	(0.1096)	(0.0106)	(0.1031)	(0.0960)	(0.1814)
in_levee $\times$ year_1970	0.2401***	-0.0551***	0.1596	0.1489	0.4472***
	(0.0760)	(0.0127)	(0.1053)	(0.1009)	(0.1667)
in_levee $\times$ year_1980	0.2111***	-0.0545***	0.1232	0.1342	0.4205**
	(0.0760)	(0.0106)	(0.1020)	(0.0954)	(0.1647)
in_levee $\times$ year_1990	0.2505***	-0.0558***	0.1190	0.1282	0.4402***
	(0.0814)	(0.0096)	(0.0995)	(0.0924)	(0.1679)
in_levee $\times$ year_2000	0.2220***	-0.0525***	0.1399	0.1482	0.4479***
	(0.0774)	(0.0101)	(0.1005)	(0.0929)	(0.1640)
in_levee $\times$ year_2010	0.1711**	-0.0481***	0.1426	0.1515	0.3842**
	(0.0808)	(0.0112)	(0.0997)	(0.0923)	(0.1634)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	168,934	169,286	169,286	169,286	169,286
R^2	0.87298	0.83129	0.64622	0.67130	0.84126
Within R^2	0.00331	0.00443	0.00107	0.00107	0.00458

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 14: Regression Discontinuity at Leveed Area Boundaries with no Discontinuity in Height Above Nearest Drainage

This table focuses on the sample of tracts within ZIP codes that straddle the boundary of leveed areas, and for which there is no discontinuity in flood risk measured by the Height Above Nearest Drainage. In columns (1) and (2), the left-hand side is the average house value, defined as the tract-level aggregate housing value divided by the number of housing units. Columns (3) and (4) focus on the average contract rent. Statistics on the number of tracts and ZIPs per metro area are presented on Table A31.

Dependent Variables: Model:	log(Housing Value) (1)	Frac. Black (2)	log(Population) (3)	log(Housing Units) (4)	log(Agg. Housing Value) (5)
Variables					
in_levee × year_1950	-0.1029***	0.0106	0.0033	0.1873**	-0.1027
	(0.0088)	(0.0058)	(0.0221)	(0.0551)	(0.0756)
in_levee × year_1960	0.0237	0.0031	-0.1142***	0.0780**	-0.0448
	(0.0330)	(0.0067)	(0.0312)	(0.0310)	(0.0605)
in_levee × year_1970	0.0411	-0.0089	-0.1474**	0.0193	-0.0711*
	(0.0312)	(0.0119)	(0.0455)	(0.0239)	(0.0362)
in_levee × year_1980	0.0102	-0.0192	-0.2026***	-0.0120*	-0.1142***
	(0.0409)	(0.0121)	(0.0488)	(0.0051)	(0.0177)
in_levee × year_1990	0.0809*	-0.0230*	-0.2358***	-0.0432***	-0.0805***
	(0.0400)	(0.0103)	(0.0511)	(0.0057)	(0.0131)
in_levee × year_2000	0.1052**	-0.0169*	-0.2231***	-0.0318**	-0.0487**
	(0.0413)	(0.0089)	(0.0539)	(0.0111)	(0.0150)
in_levee × year_2010	0.1068**	-0.0231**	-0.2111***		
	(0.0376)	(0.0079)	(0.0518)		
in_levee × year_1940				0.1903**	0.0119
				(0.0569)	(0.0804)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	10,652	10,663	10,663	10,663	10,663
R^2	0.86182	0.82580	0.64940	0.68629	0.82325
Within R^2	0.00263	0.00203	0.00590	0.00542	0.00049

Clustered (GEOID10 & year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 15: Descriptive Statistics – Leveed (Treated) Census Tracts within Radiuses of Projects Appropriated by Congress

This table presents the number of observations within radiuses of 10 or 20 miles of funded projects. Printed texts of the Flood Control Acts of 1936, 1938, 1946, 1954, 1974, 1986 mention labelled locations. After appropriation by Congress, the US Army Corps determined the specific locations of embankments and floodwalls. The labels of the Federal appropriations were geocoded by Flood Control Act and matched to high resolution leveed area maps.

year	nobs_treated_10	nobs_treated_20	nobs_untreated_10	nobs_untreated_20
1940	383	748	10,928	10,563
1950	367	732	13,507	13,142
1960	355	714	14,488	14,129
1970	383	747	24,190	23,826
1980	374	738	30,189	29,825
1990	376	738	31,186	30,824
2000	379	741	31,256	30,894
2010	355	694	29,760	29,421

Table 16: Estimated Impacts of Leveed Areas – Within Radiuses of Locations Appropriated by Congress in 1936

These tables present the estimation of the impact of leveed areas, considering the set of neighborhoods in years after funding by one of the Flood Control Acts. A neighborhood (Census tract) is treated here if 1) it belongs to a leveed area and 2) it is observed in a year after the appropriation (1940 onwards for the 1936 and 1938 Acts, 1960 for the 1954 Act, etc) and 3) it is within a 10, 20, or 50-mile radius of a location indicated by the Act.

Dependent Variables:	log_average_value	frac_black	log(total_population)	log_housing_units	log_aggregate_value
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
after_treated × year_1950	0.0666***	0.0242*	0.1274**	0.1191***	0.1965***
	(0.0227)	(0.0124)	(0.0530)	(0.0331)	(0.0297)
after_treated × year_1960	0.1615***	-0.0032	0.0958	0.1046**	0.2807***
	(0.0579)	(0.0291)	(0.0638)	(0.0519)	(0.0865)
after_treated × year_1970	0.2729***	-0.0610**	0.0427	0.0706	0.3812***
	(0.0472)	(0.0234)	(0.0666)	(0.0564)	(0.0805)
after_treated × year_1980	0.2178***	-0.0948***	0.0614	0.0502	0.3143***
	(0.0504)	(0.0214)	(0.0707)	(0.0595)	(0.0853)
after_treated × year_1990	0.3368***	-0.1077***	0.0916	0.0329	0.3992***
	(0.0453)	(0.0225)	(0.0719)	(0.0624)	(0.0842)
after_treated × year_2000	0.3220***	-0.1086***	0.1107	0.0368	0.4230***
	(0.0423)	(0.0240)	(0.0729)	(0.0628)	(0.0777)
after_treated × year_2010	0.2082***	-0.1044***	0.1224*	0.0381	0.2835***
	(0.0622)	(0.0227)	(0.0726)	(0.0624)	(0.1002)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	14,995	15,008	15,008	15,008	15,008
R^2	0.83323	0.75828	0.63639	0.68176	0.81137
Within R^2	0.00401	0.01008	0.00159	0.00133	0.00345

Clustered (GEOID10 & county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 17: Mobility Towards the Leveed Area – Where do Households Come From?

These two panels estimate the difference in household mobility inside and outside leveed areas, measured in the 1950 Census (upper panel) and measured in the 1960 Census (lower panel). The regressions are performed by comparing tracts within a 2.5 mile buffer of the boundary of the leveed area. Regressions using a 1 mile buffer provide similar signs and significance.

	1950 Census			
	Frac. Same House in 1949 (1)	Frac Diff. House Same County (2)	Frac. Diff. County or Abroad (3)	Frac. Owner Occupied (4)
In Leveed Area	-0.0088* (0.0044)	0.0108*** (0.0026)	0.0006 (0.0020)	-0.0505** (0.0226)
Metro Area f.e.	Yes	Yes	Yes	Yes
Observations	2,961	2,961	2,961	2,294
R^2	0.29838	0.42441	0.21496	0.10125

Clustered (Metro Area f.e.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

	1960 Census			
	Frac. Same House in 1955 (1)	Frac Diff. House Central City (2)	Frac. Diff. House in SMSA (3)	Frac. Owner Occupied (4)
In Leveed Area	-0.0070 (0.0059)	-0.0223** (0.0100)	0.0194** (0.0088)	-0.0184 (0.0320)
Metro Area f.e.	Yes	Yes	Yes	Yes
Observations	4,978	4,978	4,978	4,966
R^2	0.28668	0.24689	0.28397	0.12950

Clustered (Metro Area f.e.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 18: Estimation of Structural Parameters

This table estimates the impact of levees on intertemporal values. Such values are estimated by using the flow of movers to each location and the number of stayers in each location. The first approach controls for tract × demographic group fixed effects, and year fixed effects. The second approach controls for these fixed effects and for lagged population, racial composition. Standard errors are double-clustered at the tract and year levels.

Dependent Variables:	log(Intertemporal value V^g_{jt})	log(Intertemporal value V^g_{jt})
Model:	(1)	(2)
Variables		
1950 × Leveed	1.291***	0.9093**
	(0.2113)	(0.2564)
1960 × Leveed	0.2803	-0.1125
	(0.1800)	(0.2379)
1970 × Leveed	0.4318***	0.1772
	(0.0308)	(0.1899)
1980 × Leveed	0.0683*	-0.2672
	(0.0312)	(0.1919)
1990 × Leveed	-0.0263	-0.3884*
	(0.0272)	(0.1724)
2000 × Leveed	0.0448	-0.2293
	(0.0391)	(0.1720)
log(Lagged population L_{jt-1})		0.6327**
		(0.1584)
Black × log(Fraction Black x_{jt-1})		1.070***
		(0.1307)
White × log(Fraction Black x_{jt-1})		-0.2073**
		(0.0580)
Additional Controls	—	Future value Π^g_{jt+1}
		log(Price q_{jt})
Fixed-effects		
GEOID10xrace	Yes	Yes
year	Yes	Yes
Fit statistics		
Observations	21,524	18,438
R^2	0.31473	0.32273
Within R^2	0.00285	0.15462

Clustered (GEOID10 & year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 19: Estimated Selection Equation – Commitment vs Time Consistent Case

These tables present the result of the maximization of welfare by the policymaker when committing at $t = 0$ and when the policymaker is time consistent (no commitment), maximizing welfare at each t. The first-order conditions are (17) and (21) respectively. The coefficients are those of the selection equation (50).

(a) Estimated Betas

Coefficient β	Commitment	No Commitment
HAND	-1.653	-1.215
$\log(\text{Price}_{j0})$	-0.600	0.106
$\log(\text{Population}_{j0})$	0.076	0.083
Year 1	4.560	-0.884
Year 2	4.518	-1.018
Year 3	4.518	-0.950
Year 4	4.518	-0.910
Year 5	4.518	-0.882

(b) Predicted Probabilities of Flood Protection

	Average Probability of Flood Protection	
Year	With Commitment at $t = 0$	Without Commitment
1	0.464%	1.342%
2	0.447%	1.215%
3	0.447%	1.278%
4	0.447%	1.316%
5	0.447%	1.343%

A Appendix Figures and Tables

Figure A15: Propensity Score for 1936 Flood Control Act Funding – Politics, Precipitation, and Great Depression Unemployment

These four figures illustrate the process of building the propensity score to weigh counties in the estimation of the impact of Flood Control Act funding on subsequent embankment and floodwall construction. Panel 1 shows the distribution of the estimated propensity score in the control and in the treatment groups. Panel 2 displays the Congressional districts of the 74th Congress (relevant for the 1936 Flood Control Act) whose representatives are members of the Flood Control Committee. Panel 3 displays the measured county-level precipitation in the year (June 1935–June 1936) preceding the signing of the 1936 Flood Control Act by FDR. Hearings suggest that the March 1936 floods played a key role in bringing floods to the attention of the House. Panel 4 presents the share unemployed (as a share of the total population of the county) in 1930. Hearings suggest that public policies for hiring unemployed workers during the Great Depression may have been a driver of infrastructure investment.

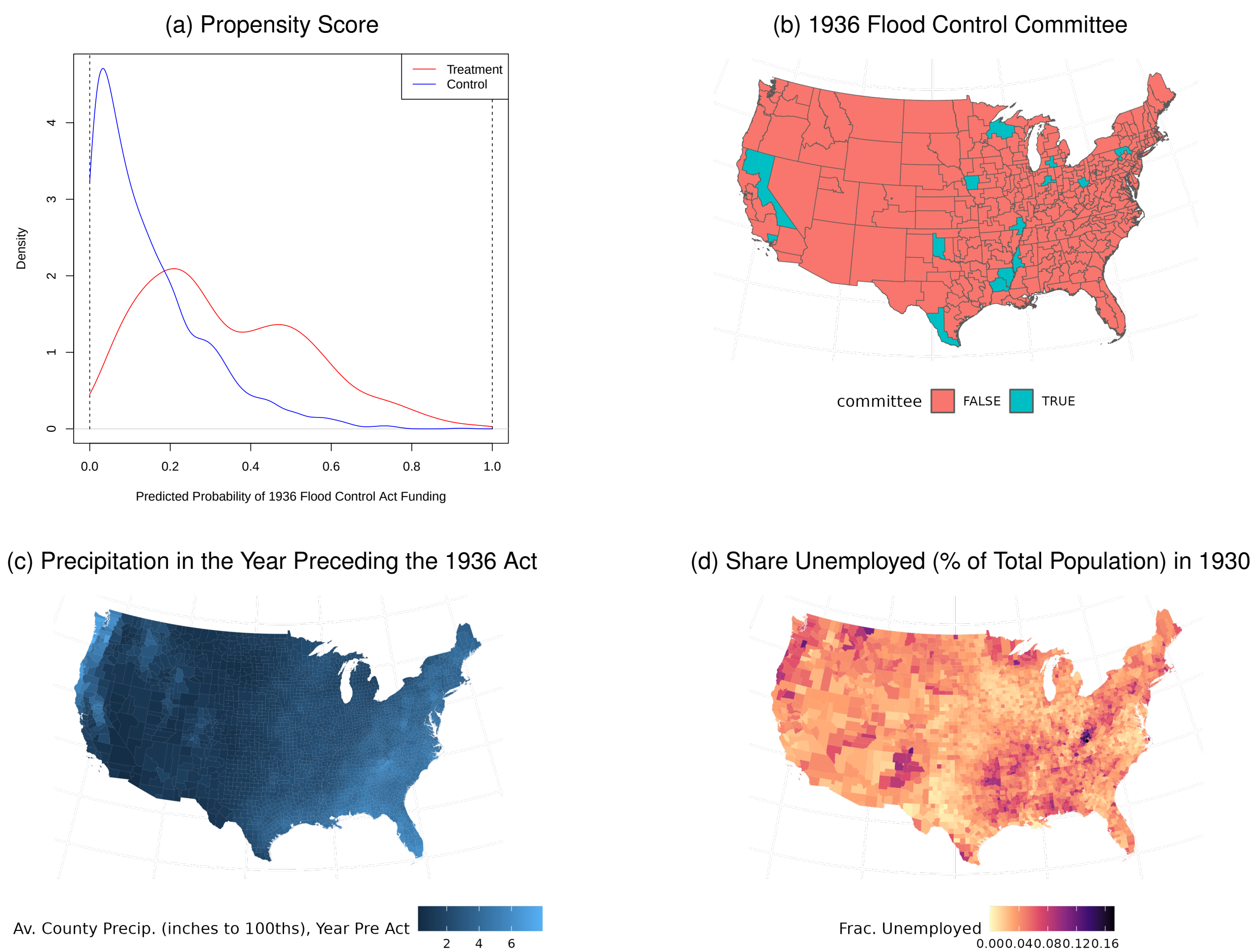


Figure A16: Building Census Tract Relationship Files

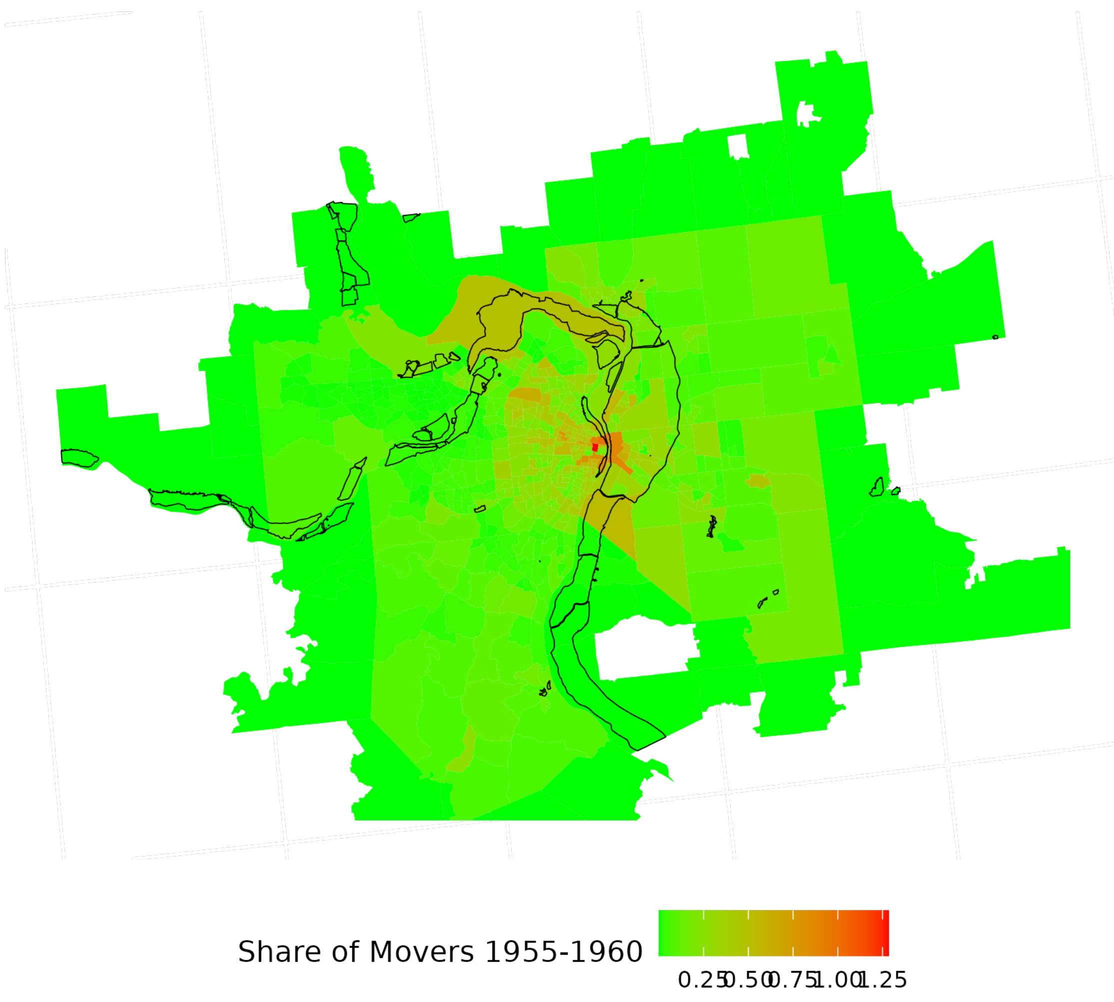
These two maps, focused on downtown Los Angeles, illustrate the process of building Census tract relationship files for constructing tracts with consistent boundaries over time. In the upper panel, the red boundary (—) is for the tracts of 1940. In the lower panel, the green boundary (—) is for the tracts of 1950. On both maps, the black boundaries are for the 2010 tracts. Areas calculated using the US National Atlas coordinate reference system 2163.

(a) Tract Relationship File, 1940–2010

(b) Tract Relationship File, 1950–2010

Figure A17: Identifying Values Using Mobility Flows – St Louis

This figures illustrates the use of tract-level data on movers to identify the structural model of location choice. Each color corresponds to the number of movers divided by the total number of movers. This is, $L^+_{jt}/\sum_{k=1}^{J} L^+_{kt}$*, or* $P(j,t+1|mobility) = \frac{U^{\theta}_{jt+1}}{\sum_{k=1,k\neq i}^{J} U^{\theta}_{kt+1}}$*. The map is for the 1960 Census, using the 'Residence in 1955' variable, table NBT18. The black line is for the boundary of the leveed area.*



Source: University of Missouri's National Historical Geographic Information System, 1960 Census. US Army Corps National Levee Database.

Table A20: The Impact of Congressional Appropriations - Flood Control Act and Subsequent Miles of Embankments and Floodwalls – 1954

*These eight regressions perform the event study separately by Act and by type of infrastructure: embankments in columns (1) and (3), floodwalls in columns (2) and (4). We present the acts with the largest number of projects funded for the sake of clarity. Clustered (county) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	asinh(length_infra)			
Model:	(1)	(2)	(3)	(4)
Variables				
within_buffer_law × decade_m6	-0.0317	-0.0787*	-0.0317	-0.0787*
	(0.0201)	(0.0411)	(0.0201)	(0.0411)
within_buffer_law × decade_m5	-0.0302	-0.0769*	-0.0302	-0.0769*
	(0.0201)	(0.0420)	(0.0201)	(0.0420)
within_buffer_law × decade_m4	-0.0302	-0.0332	-0.0302	-0.0332
	(0.0201)	(0.0583)	(0.0201)	(0.0583)
within_buffer_law × decade_m3	-0.0302	-0.0131	-0.0302	-0.0131
	(0.0201)	(0.0630)	(0.0201)	(0.0630)
within_buffer_law × decade_m2	-0.0150	0.0007	-0.0150	0.0007
	(0.0252)	(0.0720)	(0.0252)	(0.0720)
within_buffer_law × decade_0	0.1651**	0.3202**	0.1651**	0.3202**
	(0.0746)	(0.1361)	(0.0746)	(0.1361)
within_buffer_law × decade_1	0.1803***	0.4744***	0.1803***	0.4744***
	(0.0671)	(0.1502)	(0.0671)	(0.1502)
within_buffer_law × decade_2	0.3338***	0.7713***	0.3338***	0.7713***
	(0.0850)	(0.1743)	(0.0850)	(0.1743)
within_buffer_law × decade_3	0.1290*	0.1114	0.1290*	0.1114
	(0.0668)	(0.1180)	(0.0668)	(0.1180)
within_buffer_law × decade_4	0.1437**	0.2387**	0.1437**	0.2387**
	(0.0725)	(0.1202)	(0.0725)	(0.1202)
within_buffer_law × decade_5	0.1711**	0.2521**	0.1711**	0.2521**
	(0.0741)	(0.1113)	(0.0741)	(0.1113)
within_buffer_law × decade_6	0.0829	0.0597	0.0829	0.0597
	(0.0591)	(0.0873)	(0.0591)	(0.0873)
within_buffer_law × decade_7	0.0346	-0.0117	0.0346	-0.0117
	(0.0449)	(0.0593)	(0.0449)	(0.0593)
Fixed-effects				
county_fips	Yes	Yes	Yes	Yes
decade	Yes	Yes	Yes	Yes
Fit statistics				
Observations	36,195	36,195	36,195	36,195
R^2	0.13467	0.15423	0.13467	0.15423
Within R^2	0.00567	0.00597	0.00567	0.00597

Clustered (county_fips) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A21: Regression Discontinuity at the Boundary of Leveed Areas – Repeated Cross Section, 0.5 mile buffer, Gaussian RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1940	-0.1456***	0.0202***	-0.0782	-0.1260	-0.1993**
	(0.0425)	(0.0053)	(0.1044)	(0.1044)	(0.0894)
in_levee × year_1950	-0.1326**	0.0126**	-0.0289	-0.0721	-0.1328*
	(0.0550)	(0.0055)	(0.0630)	(0.0630)	(0.0688)
in_levee × year_1960	-0.0735	-0.0146	0.0270	-0.0382	-0.0382
	(0.1064)	(0.0109)	(0.0232)	(0.0278)	(0.1384)
in_levee × year_1970	0.0851*	-0.0312**	0.0558	0.0009	0.1967*
	(0.0476)	(0.0143)	(0.0444)	(0.0453)	(0.1149)
in_levee × year_1980	0.0386	-0.0273**	0.0306	-0.0067	0.1475
	(0.0443)	(0.0129)	(0.0284)	(0.0254)	(0.1061)
in_levee × year_1990	0.0830*	-0.0291***	0.0187	-0.0171	0.1639*
	(0.0447)	(0.0100)	(0.0180)	(0.0158)	(0.0915)
in_levee × year_2000	0.0498	-0.0274***	0.0450***	0.0075	0.1683*
	(0.0431)	(0.0081)	(0.0168)	(0.0126)	(0.0864)
in_levee × year_2010	-0.0292	-0.0215***	0.0567***	0.0234*	0.0695
	(0.0280)	(0.0078)	(0.0154)	(0.0137)	(0.0585)
Fixed-effects					
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	168,934	169,286	169,286	169,286	169,286
R^2	0.75148	0.37083	0.35879	0.38443	0.66759
Within R^2	0.00164	0.00201	0.00065	0.00066	0.00224

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A22: Regression Discontinuity at the Boundary of Leveed Areas – Repeated Cross Section, 1 mile buffer, Gaussian RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1940	-0.2079***	0.0219***	-0.0459	-0.0904	-0.2644***
	(0.0463)	(0.0053)	(0.1042)	(0.1068)	(0.0875)
in_levee × year_1950	-0.1720***	0.0198***	-0.0101	-0.0549	-0.1901***
	(0.0384)	(0.0054)	(0.0591)	(0.0630)	(0.0459)
in_levee × year_1960	-0.1329	-0.0042	0.0171	-0.0410	-0.1290
	(0.1075)	(0.0080)	(0.0241)	(0.0274)	(0.1479)
in_levee × year_1970	-0.0017	-0.0221*	0.0437	-0.0067	0.0535
	(0.0522)	(0.0123)	(0.0402)	(0.0423)	(0.1190)
in_levee × year_1980	-0.0500	-0.0232**	0.0190	-0.0101	0.0067
	(0.0461)	(0.0111)	(0.0327)	(0.0292)	(0.1020)
in_levee × year_1990	-0.0048	-0.0265***	3.16×10^{-6}	-0.0315**	0.0202
	(0.0511)	(0.0093)	(0.0199)	(0.0152)	(0.0983)
in_levee × year_2000	-0.0222	-0.0261***	0.0166	-0.0153	0.0359
	(0.0450)	(0.0080)	(0.0192)	(0.0126)	(0.0954)
in_levee × year_2010	-0.0471*	-0.0226***	0.0260	-0.0036	-0.0039
	(0.0276)	(0.0077)	(0.0169)	(0.0108)	(0.0548)
Fixed-effects					
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	178,181	178,553	178,553	178,553	178,553
R^2	0.74135	0.35327	0.33657	0.36543	0.65186
Within R^2	0.00281	0.00216	0.00028	0.00053	0.00150

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A23: Regression Discontinuity at the Boundary of Leveed Areas – Repeated Cross Section, 0.25 mile buffer, Rectangular RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1940	0.4052**	0.0117***	-0.7098**	-0.8656**	0.0841
	(0.1804)	(0.0035)	(0.3543)	(0.3387)	(0.2962)
in_levee × year_1950	-0.1187	-0.0453***	-0.3838	-0.4847*	-0.1230
	(0.4539)	(0.0152)	(0.2506)	(0.2497)	(0.5649)
in_levee × year_1960	0.4564**	-0.1469***	0.2073	-0.0719	0.7830***
	(0.1801)	(0.0504)	(0.1531)	(0.1999)	(0.0905)
in_levee × year_1970	0.7206***	-0.1678**	0.1756**	-0.0282	1.229***
	(0.1304)	(0.0719)	(0.0759)	(0.0854)	(0.1803)
in_levee × year_1980	0.7300***	-0.1245**	0.1710***	0.1257**	1.367***
	(0.1662)	(0.0503)	(0.0500)	(0.0588)	(0.2754)
in_levee × year_1990	0.6279***	-0.0599	0.1568**	0.1193*	1.142***
	(0.1284)	(0.0395)	(0.0609)	(0.0638)	(0.1949)
in_levee × year_2000	0.6319***	-0.0353	0.2209***	0.1761***	1.195***
	(0.1341)	(0.0282)	(0.0611)	(0.0548)	(0.1605)
in_levee × year_2010	0.1593**	-0.0274	0.1490**	0.1342	0.7065***
	(0.0623)	(0.0268)	(0.0662)	(0.0877)	(0.1879)
Fixed-effects					
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	12,884	12,897	12,897	12,897	12,897
R^2	0.77519	0.40171	0.38505	0.40858	0.69535
Within R^2	0.00619	0.00246	0.00251	0.00295	0.00803

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A24: Regression Discontinuity at the Boundary of Leveed Areas – Repeated Cross Section, 0.5 mile buffer, Rectangular RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1940	-0.0776	0.0151***	-0.1232	-0.1934*	-0.1135
	(0.0534)	(0.0051)	(0.1141)	(0.1133)	(0.1227)
in_levee × year_1950	-0.0436	0.0034	-0.0199	-0.0671	0.0623
	(0.1184)	(0.0072)	(0.0905)	(0.0822)	(0.1635)
in_levee × year_1960	0.0417	-0.0287*	0.0602**	-0.0402	0.1415
	(0.1036)	(0.0172)	(0.0273)	(0.0332)	(0.1278)
in_levee × year_1970	0.2756***	-0.0414**	0.1119*	0.0477	0.5504***
	(0.0573)	(0.0185)	(0.0590)	(0.0598)	(0.1028)
in_levee × year_1980	0.1946***	-0.0333*	0.0544	0.0069	0.4349***
	(0.0509)	(0.0186)	(0.0442)	(0.0436)	(0.1246)
in_levee × year_1990	0.2203***	-0.0349**	0.0449	0.0150	0.4361***
	(0.0461)	(0.0151)	(0.0298)	(0.0297)	(0.0988)
in_levee × year_2000	0.1861***	-0.0312***	0.0925***	0.0470*	0.4372***
	(0.0440)	(0.0119)	(0.0305)	(0.0242)	(0.0773)
in_levee × year_2010	0.0256	-0.0230**	0.1146***	0.0795**	0.2425***
	(0.0370)	(0.0111)	(0.0336)	(0.0323)	(0.0542)
Fixed-effects					
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	17,370	17,390	17,390	17,390	17,390
R^2	0.76144	0.37904	0.37318	0.39752	0.68050
Within R^2	0.00445	0.00187	0.00145	0.00109	0.00860

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A25: Regression Discontinuity at the Boundary of Leveed Areas – Repeated Cross Section, 0.5 mile buffer, Rectangular RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1940	-0.1595***	0.0214***	-0.0812	-0.1240	-0.2150**
	(0.0461)	(0.0059)	(0.0858)	(0.0847)	(0.0845)
in_levee × year_1950	-0.1322***	0.0158**	-0.0451	-0.0857*	-0.1547***
	(0.0452)	(0.0062)	(0.0510)	(0.0509)	(0.0556)
in_levee × year_1960	-0.0871	-0.0101	0.0127	-0.0509*	-0.0674
	(0.1053)	(0.0098)	(0.0254)	(0.0282)	(0.1369)
in_levee × year_1970	0.0790**	-0.0326**	0.0604	0.0012	0.1848*
	(0.0384)	(0.0128)	(0.0431)	(0.0461)	(0.0971)
in_levee × year_1980	0.0416	-0.0318***	0.0455	0.0044	0.1507*
	(0.0361)	(0.0117)	(0.0285)	(0.0273)	(0.0806)
in_levee × year_1990	0.0783*	-0.0323***	0.0303*	-0.0071	0.1622**
	(0.0426)	(0.0100)	(0.0166)	(0.0157)	(0.0808)
in_levee × year_2000	0.0475	-0.0309***	0.0525***	0.0146	0.1639**
	(0.0405)	(0.0092)	(0.0153)	(0.0157)	(0.0794)
in_levee × year_2010	-0.0142	-0.0240***	0.0645***	0.0268*	0.0859
	(0.0284)	(0.0089)	(0.0137)	(0.0143)	(0.0559)
Fixed-effects					
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	26,135	26,166	26,166	26,166	26,166
R^2	0.74723	0.36594	0.34970	0.37596	0.66249
Within R^2	0.00213	0.00320	0.00112	0.00094	0.00300

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A26: Regression Discontinuity at the Boundary of Leveed Areas – Combined Regression Discontinuity and Tract f.e., County $\times$Year f.e., 1 mile buffer, Gaussian RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee $\times$ year_1950	0.0297	0.0028	0.0437	0.0357	0.0765
	(0.1047)	(0.0115)	(0.1148)	(0.1031)	(0.2001)
in_levee $\times$ year_1960	0.0832	-0.0234**	0.0849	0.0722	0.1649
	(0.1252)	(0.0100)	(0.1084)	(0.0985)	(0.2069)
in_levee $\times$ year_1970	0.2172**	-0.0469***	0.1076	0.0964	0.3631*
	(0.0977)	(0.0114)	(0.1075)	(0.0995)	(0.1921)
in_levee $\times$ year_1980	0.1737*	-0.0491***	0.0701	0.0804	0.3177*
	(0.0959)	(0.0102)	(0.1057)	(0.0956)	(0.1884)
in_levee $\times$ year_1990	0.2186**	-0.0520***	0.0582	0.0638	0.3398*
	(0.0997)	(0.0094)	(0.1045)	(0.0942)	(0.1904)
in_levee $\times$ year_2000	0.2079**	-0.0509***	0.0697	0.0753	0.3597*
	(0.0973)	(0.0098)	(0.1055)	(0.0948)	(0.1893)
in_levee $\times$ year_2010	0.1949*	-0.0489***	0.0754	0.0794	0.3403*
	(0.1002)	(0.0107)	(0.1054)	(0.0945)	(0.1864)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	178,181	178,553	178,553	178,553	178,553
R^2	0.86506	0.82180	0.64086	0.66906	0.83461
Within R^2	0.00312	0.00549	0.00052	0.00044	0.00361

Clustered (county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A27: Regression Discontinuity at the Boundary of Leveed Areas – Combined Regression Discontinuity and Tract f.e., County $\times$Year f.e., 0.25 mile buffer, Rectangular RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee $\times$ year_1950	-0.2795	-0.0207	0.3066***	0.3114***	0.0388
	(0.3922)	(0.0216)	(0.0542)	(0.0652)	(0.3207)
in_levee $\times$ year_1960	0.1556	-0.1686***	0.9268***	0.7769***	0.8314**
	(0.2966)	(0.0477)	(0.1294)	(0.0953)	(0.3177)
in_levee $\times$ year_1970	0.3680	-0.1949**	0.8664***	0.8210***	1.193***
	(0.2249)	(0.0641)	(0.1439)	(0.1376)	(0.2337)
in_levee $\times$ year_1980	0.4703**	-0.1415**	0.8461***	0.9201***	1.414***
	(0.1853)	(0.0563)	(0.1348)	(0.1556)	(0.2891)
in_levee $\times$ year_1990	0.4106	-0.0791**	0.8297***	0.9308***	1.232***
	(0.2232)	(0.0300)	(0.1342)	(0.1619)	(0.2976)
in_levee $\times$ year_2000	0.3281	-0.0584***	0.8905***	0.9805***	1.182***
	(0.2112)	(0.0137)	(0.1474)	(0.1649)	(0.2317)
in_levee $\times$ year_2010	0.0265	-0.0506***	0.8606***	0.9670***	0.8827**
	(0.2766)	(0.0101)	(0.1498)	(0.1638)	(0.3052)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	12,884	12,897	12,897	12,897	12,897
R^2	0.88230	0.84577	0.65061	0.67325	0.84686
Within R^2	0.00207	0.00434	0.00414	0.00447	0.00398

Clustered (GEOID10 & year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A28: Regression Discontinuity at the Boundary of Leveed Areas – Combined Regression Discontinuity and Tract f.e., County $\times$Year f.e., 0.5 mile buffer, Rectangular RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee × year_1950	0.0596**	-0.0104	0.1014***	0.1238***	0.2240***
	(0.0251)	(0.0071)	(0.0280)	(0.0281)	(0.0324)
in_levee × year_1960	0.1496**	-0.0532***	0.2239***	0.1986***	0.3498***
	(0.0540)	(0.0101)	(0.0427)	(0.0321)	(0.0730)
in_levee × year_1970	0.3524***	-0.0681***	0.2644***	0.2673***	0.7194***
	(0.0462)	(0.0158)	(0.0569)	(0.0560)	(0.0895)
in_levee × year_1980	0.3265***	-0.0681***	0.1927**	0.2162**	0.6763***
	(0.0502)	(0.0154)	(0.0620)	(0.0664)	(0.1058)
in_levee × year_1990	0.3499***	-0.0698***	0.1914**	0.2233**	0.6806***
	(0.0553)	(0.0121)	(0.0617)	(0.0682)	(0.1050)
in_levee × year_2000	0.3103***	-0.0619***	0.2380**	0.2591***	0.6866***
	(0.0510)	(0.0098)	(0.0686)	(0.0717)	(0.1004)
in_levee × year_2010	0.2103***	-0.0546***	0.2429***	0.2749***	0.5572***
	(0.0407)	(0.0079)	(0.0646)	(0.0677)	(0.0822)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	17,370	17,390	17,390	17,390	17,390
R^2	0.87938	0.83535	0.65181	0.67533	0.84604
Within R^2	0.00434	0.00414	0.00198	0.00208	0.00681

Clustered (GEOID10 & year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A29: Regression Discontinuity at the Boundary of Leveed Areas – Combined Regression Discontinuity and Tract f.e., County $\times$Year f.e., 1 mile buffer, Rectangular RD Kernel

Dependent Variables:	log(Housing Value)	Frac. Black	log(Population)	log(Housing Units)	log(Agg. Housing Value)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
in_levee $\times$ year_1950	0.0364***	-0.0004	0.0491***	0.0458***	0.0921***
	(0.0049)	(0.0038)	(0.0116)	(0.0118)	(0.0128)
in_levee $\times$ year_1960	0.0773**	-0.0286***	0.1202***	0.1024***	0.1803***
	(0.0316)	(0.0041)	(0.0289)	(0.0224)	(0.0436)
in_levee $\times$ year_1970	0.2509***	-0.0536***	0.1649***	0.1481***	0.4581***
	(0.0282)	(0.0089)	(0.0370)	(0.0351)	(0.0583)
in_levee $\times$ year_1980	0.2260***	-0.0553***	0.1401***	0.1457***	0.4393***
	(0.0319)	(0.0092)	(0.0378)	(0.0385)	(0.0632)
in_levee $\times$ year_1990	0.2595***	-0.0554***	0.1331***	0.1391***	0.4574***
	(0.0327)	(0.0081)	(0.0378)	(0.0397)	(0.0629)
in_levee $\times$ year_2000	0.2359***	-0.0527***	0.1484***	0.1543***	0.4640***
	(0.0328)	(0.0072)	(0.0395)	(0.0407)	(0.0624)
in_levee $\times$ year_2010	0.1897***	-0.0479***	0.1546***	0.1584***	0.4087***
	(0.0284)	(0.0065)	(0.0389)	(0.0398)	(0.0560)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	26,135	26,166	26,166	26,166	26,166
R^2	0.86994	0.82537	0.64650	0.67247	0.83984
Within R^2	0.00440	0.00601	0.00163	0.00161	0.00645

Clustered (GEOID10 & year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A30: Estimated Impacts of Leveed Areas – Within Radiuses of Locations Appropriated by Congress in 1936

These tables present the estimation of the impact of leveed areas, considering the set of neighborhoods in years after funding by one of the Flood Control Acts. A neighborhood (Census tract) is treated here if 1) it belongs to a leveed area and 2) it is observed in a year after the appropriation (1940 onwards for the 1936 and 1938 Acts, 1960 for the 1954 Act, etc) and 3) it is within a 10, 20, or 50-mile radius of a location indicated by the Act.

Dependent Variables:	log_average_value	frac_black	log(total_population)	log_housing_units	log_aggregate_value
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
after_treated × year_1950	0.0834***	0.0074	0.1726***	0.1733***	0.2532***
	(0.0062)	(0.0062)	(0.0300)	(0.0134)	(0.0070)
after_treated × year_1960	0.2403***	-0.0119	0.2392***	0.2307***	0.4599***
	(0.0409)	(0.0112)	(0.0529)	(0.0409)	(0.0604)
after_treated × year_1970	0.2671***	-0.0366***	0.2241***	0.2217***	0.5296***
	(0.0279)	(0.0131)	(0.0577)	(0.0488)	(0.0633)
after_treated × year_1980	0.2082***	-0.0390***	0.2337***	0.2094***	0.4592***
	(0.0316)	(0.0130)	(0.0588)	(0.0519)	(0.0683)
after_treated × year_1990	0.2835***	-0.0429***	0.2442***	0.1903***	0.4968***
	(0.0313)	(0.0120)	(0.0581)	(0.0529)	(0.0672)
after_treated × year_2000	0.2374***	-0.0415***	0.2598***	0.1870***	0.4775***
	(0.0283)	(0.0114)	(0.0588)	(0.0531)	(0.0629)
after_treated × year_2010	0.1405***	-0.0405***	0.2599***	0.1762***	0.3424***
	(0.0393)	(0.0108)	(0.0584)	(0.0527)	(0.0696)
Fixed-effects					
GEOID10	Yes	Yes	Yes	Yes	Yes
county_fips_year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	25,576	25,600	25,600	25,600	25,600
R^2	0.86370	0.76950	0.67255	0.71338	0.82946
Within R^2	0.00381	0.00220	0.00485	0.00347	0.00607

Clustered (GEOID10 & county_fips_year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A31: Regression Discontinuity at Leveed Area Boundaries with no Discontinuity in Height Above Nearest Drainage – Metro Area-Level Statistics

For each 5-digit ZIP code j that covers both sides of a leveed area boundary, the difference in HAND within and outside the leveed area was estimated. ZIP codes with no statistically significant discontinuity at 95% were selected. This table presents the number of ZIP codes and tracts by MSA with no such statistically significant discontinuity.

Metropolitan Area	Number of 1940 Tracts	Number of 5-digit ZIP Codes with No HAND Discontinuity
Los Angeles-Long Beach-Anaheim, CA	659	71
Indianapolis-Carmel-Anderson, IN	76	11
Washington-Arlington-Alexandria, DC-VA-MD-WV	55	5
Dallas-Fort Worth-Arlington, TX	49	7
New York-Newark-Jersey City, NY-NJ-PA	48	6
St. Louis, MO-IL	45	7
Kansas City, MO-KS	42	7
Pittsburgh, PA	41	6
Dayton, OH	40	6
Minneapolis-St. Paul-Bloomington, MN-WI	36	4
San Francisco-Oakland-Hayward, CA	34	4
Cincinnati, OH-KY-IN	30	6
Providence-Warwick, RI-MA	27	4
Columbus, OH	26	3
Des Moines-West Des Moines, IA	26	3
Memphis, TN-MS-AR	26	3
Milwaukee-Waukesha-West Allis, WI	25	2
Louisville/Jefferson County, KY-IN	24	3
Richmond, VA	23	3
Portland-Vancouver-Hillsboro, OR-WA	20	3
Syracuse, NY	13	2
Hartford-West Hartford-East Hartford, CT	12	2
Toledo, OH	12	2
Augusta-Richmond County, GA-SC	9	1
Buffalo-Cheektowaga-Niagara Falls, NY	9	1
Chicago-Naperville-Elgin, IL-IN-WI	9	1
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	9	2
Flint, MI	7	1
Nashville-Davidson–Murfreesboro–Franklin, TN	6	1
Cleveland-Elyria, OH	4	1
New Haven-Milford, CT	4	1

Table A32: USGS Water Gages in the Metropolitan Areas of the 1940-2010 Sample

This table presents the number of active and inactive USGS gages with past records. Column (1) presents the number of gages per MSA (2010 boundaries), column (2) presents the number of such gages within leveed areas, according to the boundaries of the US Army Corps of Engineers' National Levee Database. Column (3) sums the number of peaks recorded for these gages, column (4) the number of peaks for those gages within leveed areas, and column (5) the number of gages whose streamflow (ft3/s) is in the 75th percentile of streamflows.

	# of Gages		# of Peak Streamflows		
Metro Area	Overall (1)	within Leveed (2)	Overall (3)	within Leveed (4)	In the 75th Percentile (5)
Louisville/Jefferson County, KY-IN	50	9	1848	309	6
Kansas City, MO-KS	60	8	1595	227	89
Los Angeles-Long Beach-Anaheim, CA	119	6	3733	148	9
Urban Honolulu, HI	114	3	4379	127	0
Riverside-San Bernardino-Ontario, CA	137	3	4938	103	18
Chicago-Naperville-Elgin, IL-IN-WI	162	1	6014	77	0
Providence-Warwick, RI-MA	64	1	1752	57	0
Oxnard-Thousand Oaks-Ventura, CA	34	1	1154	48	3
Seattle-Tacoma-Bellevue, WA	185	2	6303	29	0
Hartford-West Hartford-East Hartford, CT	71	1	2118	27	1
Memphis, TN-MS-AR	34	2	731	19	8
San Francisco-Oakland-Hayward, CA	81	2	2088	18	2
New Orleans-Metairie, LA	13	1	286	9	0
Akron, OH	26	0	649	0	0
Atlanta-Sandy Springs-Roswell, GA	151	0	4174	0	0
Atlantic City-Hammonton, NJ	5	0	172	0	0
Augusta-Richmond County, GA-SC	28	0	719	0	0
Austin-Round Rock, TX	64	0	1998	0	0
Bakersfield, CA	30	0	738	0	0
Baltimore-Columbia-Towson, MD	97	0	2931	0	0

Sources: Census Bureau Core Based Statistical Areas, National Levee Database, USGS Historical Streamflow Records.