Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hazard-specific pathways shape the flood protection gap in the USA

Koorosh Azizi

kazizi@tnstate.edu

Tennessee State University https://orcid.org/0000-0002-0786-7323

Hassan Dashtian

University of Texas at Austin

Margaret Garcia

Arizona State University https://orcid.org/0000-0002-2864-2377

Patrick Bixler

University of Texas at Austin

Dev Niyogi

Jackson School of Geosciences and Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX, USA

Paul Bates

University of Bristol https://orcid.org/0000-0001-9192-9963

Bridget Scanlon

University of Texas at Austin https://orcid.org/0000-0002-1234-4199

Article

Keywords:

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11080340/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Abstract

Financial flood protection gaps are often reported as a single measure, masking important differences in how protection falls short. We combine building-level estimates of rainfall, riverine and coastal flood losses with property characteristics, FEMA flood zones and National Flood Insurance Program (NFIP) records to trace how expected losses are distributed across the United States. We distinguish expected losses outside reconstructed NFIP payout ("public shortfall") from expected covered payout above the allocated building-coverage risk-premium base ("NFIP shortfall"). Public shortfall accounts for 81% of combined-hazard expected annual loss. Estimated NFIP participation is low for rainfall and riverine exposure (16 and 20%, respectively), compared with 59% for coastal exposure. Yet coastal NFIP shortfall reaches 32% of expected loss, versus about 11% inland. Across census tracts, similar financial gaps reflect different combinations of low participation, inadequate coverage and pressure within the insured pool. Effective policy must address these drivers differently across hazards and places.

Introduction

Flood-risk assessments combine hazard, exposure and vulnerability to estimate where damage may occur and its severity. These measures do not show who ultimately carries the resulting financial loss. That burden also depends on whether a property is insured, how much of the loss its coverage absorbs and whether the associated premiums support the expected payout. Flood risk is therefore a question of allocation as much as of exposure: who carries the expected burden when insurance participation, coverage and premiums vary across an uneven multi-hazard landscape?

This question is increasingly consequential because flood losses are large and persistent. Globally, flooding has produced more than $2 trillion in economic losses since 2000, including more than $42 billion in 2025 alone[1]. In the United States, inland flooding in the billion-dollar disaster record caused about $203 billion in losses from 1980 to 2024[2], and historical precipitation change explains roughly one-third of the increase in national flood damages[3]. Continued development in exposed areas and changes in heavy rainfall and coastal extremes are likely to sustain this pressure, increasing the importance of the institutions through which losses are insured, priced or left uncovered[3–8].

Standard homeowner insurance policies in the United States generally exclude flood, making separate flood insurance central to who bears the loss[9]. The National Flood Insurance Program (NFIP) remains the main public source of coverage, although private insurance has expanded[9,10]. Participation is shaped strongly by FEMA's Special Flood Hazard Area (SFHA), which triggers mandatory purchase for federally backed mortgages[11,12]. Yet the maps defining the SFHA vary in age and coverage and do not consistently represent pluvial flooding from rainfall, fluvial risk outside mapped zones, present-day conditions where the underlying studies are dated, or future climate conditions[9,12]. Participation is concentrated in coastal communities, while many buildings exposed to pluvial and fluvial flooding remain outside the program[10,13,14]. Premiums must also reflect risk without becoming unaffordable, and

evidence from Risk Rating 2.0 indicates that increases without affordability support can reduce enrollment[11,15–18].

Previous studies have examined flood exposure, insurance participation, coverage and pricing separately. National exposure analyses identify substantial risk outside regulatory floodplains[6,8,19], claims-based work links many NFIP claims beyond the SFHA to pluvial flooding[20], and insurance-demand studies link participation to risk perception, information, trust and willingness to pay[21–26]. Recent national work has quantified underinsurance among single-family residences[10], while other studies examine recurrent losses, pricing and pressure on NFIP finances[15–17,27]. However, these studies do not trace pluvial, fluvial and coastal losses through the same national insurance structure to distinguish burden outside coverage from pressure within the insured portfolio.

We address this gap through a building-scale reconstruction of flood losses and NFIP coverage across the contiguous United States (CONUS). We combine pluvial, fluvial and coastal flood estimates with building geometry, census-tract-informed replacement values, FEMA flood zones[28] and spatially redacted NFIP policy records[29] active as of October 2025. Because policy files do not identify insured parcels, observed records are allocated to likely buildings. “Reconstructed” refers to this allocation; reported policy coverage and written costs are observed, not simulated. This reconstruction represents the pathway observed in NFIP records and excludes privately underwritten flood insurance.

We define two burden channels from expected annual loss, uncovered loss, and covered payout. Public shortfall is modeled loss remaining outside expected NFIP payout through nonparticipation, deductibles or coverage limits. It falls first on property owners and communities and may later generate demand for post-disaster assistance; it does not measure government expenditure. NFIP shortfall is the positive difference between expected covered payout and the risk premium allocated to building coverage. It indicates pressure within the reconstructed portfolio, not the program’s full financial balance. Premium-funded payout accounts for the remainder; each expected loss dollar is assigned to one of three destinations (Fig. 1).

The reconstruction covers 12.9 million exposed buildings with modeled flood inundation, including 7.5 million with positive combined-hazard (pluvial, fluvial, and coastal) expected annual loss, and 74,240 census tracts used in the regime analysis. The 30-m hazard framework has been described and evaluated in prior U.S. applications[7,30], while spatial aggregation to census tracts reduces the influence of local prediction error. We examine how pluvial, fluvial and coastal flooding shape the national flood protection gap through different combinations of public and NFIP shortfalls, and how these combinations vary across places. The results show that similar expected losses can produce different insurance problems depending on the hazard and where the burden falls.

Results

National geography of multi-hazard flood loss

The combined hazard layer yields $25.7 billion yr^{-1} in expected annual loss (EAL) across the processed building inventory (Fig. 2). Pluvial flooding contributes $17.1 billion yr^{-1}, followed by fluvial ($5.3 billion) and coastal flooding ($3.9 billion). Pluvial and fluvial losses extend widely inland, particularly in Appalachia, while coastal loss is concentrated in fewer tracts. California accounts for 10% of combined EAL, the largest state share. New York and Pennsylvania rank among the highest-loss states, while Kentucky and West Virginia accumulate substantial burden almost entirely through inland hazards (Extended Data Fig. 2; Supplementary Table S3). Similar state totals conceal different hazard structures: Florida is strongly coastal-dominant, Washington and Oregon fluvial-dominant, and California and Texas have large pluvial components.

The combined layer uses the maximum hazard-specific depth at each return period before loss estimation. This prevents double counting where pluvial, fluvial and coastal layers overlap but does not represent compound or sequential flooding. Hazard-specific totals are therefore not additive. Even so, inland flood losses remain extensive, whereas insurance participation is concentrated along the coast (Extended Data Table 1).

Hazard-specific pathways through NFIP participation and coverage

Reconstructed NFIP coverage differs markedly by hazard (Fig. 3a). Coverage is present for 59% of coastal-exposed buildings, compared with 20% of fluvial-exposed and 16% of pluvial-exposed buildings. Across the combined layer, 1.87 million of 7.51 million exposed buildings carry reconstructed NFIP coverage. Coastal participation is higher across the expected-loss distribution, while inland participation remains low across most loss levels. Public shortfall equals 50% of coastal EAL, 85% of fluvial EAL and 87% of pluvial EAL (Table 1). For inland flooding, low participation is the main reason expected loss remains outside the NFIP.

Among insured exposed buildings, 23% of combined-hazard EAL remains uncovered (Fig. 3b). The uncovered share is elevated at low EAL, where a fixed per-event deductible absorbs a larger fraction of small losses. It declines through much of the middle range and rises sharply in the upper tail as coverage limits bind. Hazard averages range from 18% for pluvial flooding to 33% for fluvial and 22% for coastal flooding (Extended Data Table 1). Nationally, $1.29 billion yr^{-1} of expected loss on insured buildings lies above reconstructed coverage, compared with $0.10 billion absorbed by the deductible (Supplementary Table S18). Among insured buildings, most uncovered loss lies above coverage limits, but the larger national gap comes from exposed buildings without coverage.

The location of public shortfall relative to FEMA flood zones shows a similar inland–coastal contrast (Fig. 3c). Of pluvial public shortfall, 78% lies outside SFHA zones, compared with 59% for fluvial and 36% for coastal flooding. Across the combined layer, 71% of public shortfall lies outside the mapped high-risk core. Most of that amount is in mapped non-SFHA zones: 63.1 percentage points of the combined share,

compared with 7.8 percentage points in areas with missing or undigitized designations (Supplementary Table S22). Higher coastal participation is therefore associated with closer alignment between coastal loss and the mapped regulatory core, not complete or adequate coverage.

Table 1
National flood exposure, NFIP participation and shortfall by hazard.

Hazard	Exposed buildings	NFIP participation	EAL ($B yr^{-1})	Public shortfall ($B yr^{-1}; % EAL)	NFIP shortfall ($B yr^{-1}; % EAL)
Pluvial	4,890,253	15.7%	17.11	14.84 (86.8%)	1.87 (10.9%)
Fluvial	1,472,944	20.0%	5.32	4.53 (85.3%)	0.60 (11.3%)
Coastal	1,624,810	59.0%	3.88	1.94 (50.0%)	1.24 (32.1%)
Combined	7,506,151	24.9%	25.67	20.87 (81.3%)	3.59 (14.0%)

Burden reallocation across the expected-loss spectrum

The balance between public shortfall and NFIP shortfall changes with EAL (Fig. 4). Public shortfall dominates at low and middle EAL, although its share initially declines as a larger proportion of expected loss enters covered payout. Over this range, the NFIP-shortfall share increases and the gap between the two channels narrows. However, the NFIP-shortfall share remains below the public-shortfall share across the distribution, with premium-funded payout accounting for the remainder of EAL. The low participation rates in Fig. 3a suggest that nonparticipation is the main contributor to public shortfall across much of the low-to-middle EAL range (Fig. 3a; Supplementary Table S4).

The national accounting shows the scale of the two channels. Of $25.7 billion yr^{-1} in combined-hazard EAL, $20.9 billion, or 81%, remains as public shortfall and only $4.8 billion, or 19%, enters expected covered payout. Within that payout, $3.6 billion is classified as NFIP shortfall and ~$1.2 billion as premium-funded payout. These three destinations are exhaustive at the building scale. NFIP shortfall is proportionally greatest for coastal flooding, reaching 32% of coastal EAL compared with about 11% for pluvial and fluvial flooding (Extended Data Table 1). Because NFIP shortfall sums positive building-level gaps, it is a diagnostic of where pressure arises, not a net program-wide deficit; premiums above expected payout can offset gaps elsewhere.

Across the distribution, the binned NFIP-shortfall share rises to approximately 19% near $10,000 yr^{-1} and then declines as the uncovered-loss share turns upward. The per-event deductible contributes to uncovered loss at low EAL, but coverage limits govern the upper-tail reversal. When an event loss exceeds available coverage after the deductible, the uncovered portion contributes to public shortfall rather than expected payout. The unsmoothed bin means and confidence intervals support this broad reversal, but few buildings occupy the extreme upper tail and the precise threshold is less certain

(Supplementary Fig. S13 and Tables S23, S25). A single coverage-deficit measure would not show this shift between burden channels.

Spatial regimes and hazard composition of flood burden

At census-tract scale, public shortfall spans much of the country, while NFIP shortfall is concentrated more strongly along the Gulf and Atlantic coasts (Fig. 5a,b). Three combined-hazard measures summarize this variation: public loss absorption, the share of EAL remaining outside expected payout; premium-deficit pressure, NFIP shortfall relative to allocated written policy cost; and total shortfall burden, the sum of public and NFIP shortfalls. A four-component Gaussian mixture model fitted without hazard type or geography identifies empirical regimes differing mainly in burden magnitude and how much loss enters reconstructed coverage (Fig. 5c). The components overlap, so these descriptive regimes do not constitute fixed actuarial classes (posterior uncertainty and model-selection diagnostics appear in Supplementary Fig. S8 and Tables S7, S31).

The minor burden, uninsured regime contains 17,882 tracts with near-complete public loss absorption, negligible NFIP shortfall and a median total burden of $23,950. The minor burden, some coverage regime contains 5,747 tracts with a small reconstructed insured pool and a median total burden of $5,160. These regimes occupy much of the lower-loss interior. A high uncovered fraction does not imply a large dollar burden: the uninsured regime has median public loss absorption of one, but a public-shortfall concentration ratio of only 0.39 (Extended Data Table 3).

The major burden, mostly uncovered regime contains 21,197 tracts, or 28.6% of classified tracts, and has the largest median total burden, $214,770, with median public loss absorption of 0.927. It carries $12.89 billion in public shortfall, approximately 62% of public shortfall among classified tracts and 2.16 times the amount implied by its tract share. The major burden, substantially insured regime contains 29,414 tracts, or 39.6% of classified tracts, with a median total burden of $54,310, median public loss absorption of 0.457 and median premium-deficit pressure of 1.266. It carries $2.91 billion in NFIP shortfall, approximately 81% of NFIP shortfall among classified tracts and 2.04 times the amount implied by its tract share, alongside $5.89 billion in public shortfall (Fig. 5d; Extended Data Table 3). The mostly uncovered regime is common across inland and Appalachian regions, while the substantially insured regime is concentrated more strongly along the Gulf and Atlantic coasts and in other high-participation areas.

The first major regime concentrates economically significant loss outside expected payout. The second represents shared burden, with substantial NFIP shortfall alongside residual public shortfall. Across these regimes, greater entry of modeled loss into coverage is associated with a mixed burden, without eliminating uncovered loss.

Hazard composition differs across regimes (Fig. 6), although hazard type was excluded from their definition. Shares are normalized over the separately modeled hazard-specific shortfalls within each

regime and channel. In the mostly uncovered regime, pluvial flooding accounts for 76% of public shortfall, fluvial flooding for 22% and coastal flooding for 2%. Inland hazards still dominate public shortfall in the substantially insured regime, but coastal flooding contributes 27%. Its NFIP-shortfall composition is more balanced: pluvial flooding contributes 45%, coastal flooding 41% and fluvial flooding 14%. The mostly uncovered regime's smaller NFIP-shortfall total is 72% pluvial and 25% fluvial. Within the substantially insured regime, inland hazards therefore dominate public shortfall, while pluvial and coastal flooding contribute similar shares of NFIP shortfall. Coastal concentration does not imply that coastal flooding alone produces pressure within the insured portfolio.

Discussion

The reconstruction distinguishes two related deficits in financial flood protection under the NFIP. Public shortfall is expected loss outside reconstructed NFIP payout, while NFIP shortfall is expected covered payout above the allocated building-coverage risk-premium base. The first falls initially on property owners and communities; the second indicates premium-deficit pressure within the reconstructed portfolio, with implications for borrowing, future premiums and public finances[9,13,14]. This distinction separates places where loss largely bypasses insurance from places where more loss enters coverage but pressure remains within the insured portfolio. The hazard also matters because pluvial, fluvial and coastal losses are filtered differently through mapping, participation and coverage.

Our national estimates are similar in scale to earlier work, although comparisons differ in scope. Combined-hazard EAL of $25.7 billion is structure-only and excludes contents, vehicles and business interruption; the approximately $32 billion estimate from a related national hazard framework covers a broader property base[8]. A recent national analysis found roughly 70% of expected residential flood loss was uninsured[10], compared with an 81% public-shortfall share across all modeled building types here. The reconstructed NFIP portfolio's expected payout is $4.8 billion yr^{-1}, compared with the Congressional Budget Office (CBO) estimate of about $3.7 billion in expected annual NFIP claims[14,31]. Accounting scopes and dollar years differ, but estimates are of the same order. After netting premiums above expected payout against deficits elsewhere, our aggregate reconstructed gap is $1.8 billion yr^{-1}, close in scale to CBO's $1.4 billion program estimate under a different accounting framework. CBO found shortfall concentrated in coastal counties, consistent with our maps. However, pluvial flooding produces the largest absolute hazard-specific NFIP shortfall nationally, and pluvial and coastal flooding contribute similar shares within the major burden, substantially insured regime. Coastal concentration should not be interpreted as a coastal-hazard effect alone. These comparisons support plausibility, not direct validation.

The inland results extend claims-based evidence. A recent study[20] attributed 87% of analyzed NFIP claims outside the SFHA to pluvial flooding. In our reconstruction, pluvial flooding produces the largest hazard-specific EAL, but only 16% of exposed buildings have reconstructed NFIP coverage, and 78% of pluvial public shortfall lies outside SFHA zones. Mapping and mandatory-purchase rules provide

stronger signals in mapped high-risk areas, while participation elsewhere also depends on affordability, perceived risk, information and trust[21,25,26,32]. These mechanisms are consistent with the observed contrast, although the reconstruction does not identify them directly.

Coverage is concentrated at limits in the reconstructed portfolio (Supplementary Table S23), and loss above available coverage dominates the residual upper-tail burden (Supplementary Table S25). Because expected payout is capped at reconstructed coverage, additional modeled loss above that cap remains in public shortfall, consistent with the declining NFIP-shortfall share. Claims data could test this pattern in realized events by examining whether severe building claims more often reach policy limits among upper-EAL properties. The turning point around $10,000 yr^{-1} is descriptive, not an actuarial or regulatory boundary, and its location is less certain than the broader upper-tail decline[33–35].

The tract regimes show why higher participation alone does not necessarily eliminate shortfalls. In major burden, mostly uncovered tracts, economically significant loss remains outside coverage, pointing to needs for better hazard information, affordability support and broader participation. In major burden, substantially insured tracts, more loss enters the NFIP, but public shortfall persists alongside NFIP shortfall. Premium support, coverage limits, mitigation and risk pooling are additional considerations there. Risk selection may contribute if higher-risk properties more often purchase coverage where participation is voluntary and premiums remain below expected loss[26]. Participation, affordability and premium adequacy are distinct, but reforming one can affect the others. Premium increases without affordability support may reduce enrollment, while expanded enrollment without sufficient premium support may increase pressure within the program[6,15,17].

These differences limit the usefulness of a uniform approach to insurance reform. Where public shortfall is concentrated outside SFHA zones, continuous hazard information and affordability support are more relevant than premium reform alone. Similar expected losses can require different combinations of information, participation and financial capacity. This is especially important for pluvial flooding, whose diffuse geography is poorly represented by binary mapped zones, and coastal areas, where high participation can coexist with uncovered loss and NFIP premium-deficit pressure.

The results have several limitations. The analysis uses one national hazard model, and the division between pluvial and fluvial flooding is model-defined because some small-channel flooding is represented in the pluvial layer[7,30]. First-floor heights are assigned from Elevation Certificate strata, basement presence is sampled from regional prevalence and below-grade damage contributes materially to the national total. Removing the below-grade basement-damage pathway lowers combined EAL to $15.22 billion yr^{-1}, although pluvial flooding remains the largest hazard component (Supplementary Table S15). Events more frequent than the 5-year return period are not represented, and the maximum-depth combined layer avoids double counting but does not simulate compound flooding[36]. Removing the first 0.2 m of pluvial depth lowers combined EAL by 23%. The NFHL is not seamless, although only 7.8 percentage points of the 71% outside-SFHA share arise from missing or undigitized designations. Although absolute loss estimates vary, the predominance of public shortfall

and the inland–coastal contrast persist across the tested specifications (Supplementary Tables S14–S16).

Insurance is reconstructed from spatially redacted NFIP policies, not observed parcel-level contracts. Fewer than one policy in a thousand can be matched by footprint containment; 69% are placed deterministically within the reported tract and rated flood zone and 31% through stochastic allocation within the tract (Supplementary Table S11). The reconstruction is therefore more reliable at tract scale than for any individual building. Private and surplus-lines insurance are absent from the data and may reduce public shortfall where they are common. The analysis also excludes disaster assistance, household savings, credit and other recovery resources[37,38]. NFIP shortfall indicates where modeled payout is high relative to the reconstructed premium base; it is not an estimate of program solvency or portfolio-specific actuarial adequacy. The reconstruction has not been validated against event-level NFIP claims.

Expected flood loss alone does not identify a community's financial problem. Pluvial and fluvial losses are concentrated mainly in public shortfall, while high-participation areas can carry uncovered loss alongside NFIP pressure. Combining these conditions in a national protection-gap measure can obscure the appropriate response. Future work should compare reconstructed channels with event-level claims, examine how tract regimes align with socioeconomic vulnerability and recovery capacity, test alternative policy-allocation rules and incorporate private flood insurance as suitable data become available.

Methods

Overview and data sources

We developed a building-scale reconstruction of flood loss, National Flood Insurance Program (NFIP) coverage, and the resulting burden channels across the contiguous United States. The processed dataset contains 12,932,146 buildings with modeled inundation under at least one hazard and return period. Buildings in Alabama, Connecticut, and North Dakota were not carried through the upstream building-to-tract assignment and are absent from the processed dataset; national totals therefore exclude these states. Exposed buildings are defined as those with positive expected annual loss (EAL) for the hazard layer being analyzed. Under the combined hazard layer, 7,506,151 buildings meet this definition. Unless stated otherwise, insurance-participation shares use exposed buildings as the denominator.

The workflow integrates hazard-specific flood depths, building geometry, tract-level census information, Hazus replacement-cost data[39], FEMA flood-zone information, surveyed first-floor elevations, and observed NFIP policy records. These inputs are used to estimate direct event loss, EAL, expected covered payout, public shortfall, and NFIP shortfall for individual buildings, followed by aggregation to census tracts (Extended Data Fig. 1).

Flood depths were taken from the Fathom U.S. flood-model framework, which simulates pluvial, fluvial, and coastal inundation at 30 m grid resolution across the conterminous United States. The framework has been described and evaluated in prior peer-reviewed U.S. applications[7,30]. Building-level predictions retain local uncertainty, but aggregation to census tracts reduces the influence of local over- and under-prediction on the spatial comparisons used in this study.

Building exposure was represented using footprint polygons and LiDAR-derived heights[40]. Tract-level priors and housing-age information were taken from the American Community Survey (ACS), including Tables S2504 and B25034[41]. Prototype-specific replacement-cost information and above-floor depth-damage relationships were based on Hazus technical guidance, while below-grade damage relationships were taken from U.S. Army Corps of Engineers guidance. Flood-zone information was taken from FEMA's National Flood Hazard Layer (NFHL)[28]. Designations beginning with A or V were treated as Special Flood Hazard Area (SFHA) zones. All other and missing designations, including locations without digitized NFHL coverage, were treated as outside the mapped SFHA in the main analysis. The NFHL is assembled from community-level studies with gaps, overlaps, and non-digitized areas; these features were not manually infilled, and SFHA-related results inherit the resulting limitations. Observed insurance information was taken from spatially redacted OpenFEMA NFIP policy records[29]. Private flood-insurance policies are not observed in these files and are treated as a limitation, not as evidence of zero private coverage. The National Structure Inventory and related U.S. Army Corps of Engineers documentation were used where needed for structure classification and inventory harmonization.

Building-scale hazard depths

Let i index buildings, h index hazard type, and r index return period. Hazard type was defined as:

$$h \in \{PD, FD, CD\},$$

where PD, FD and CD denote pluvial, fluvial and coastal flooding, respectively. The hazard fields provide depths at the 5-, 10-, 100- and 500-year return periods. All modeled depths, building heights and first-floor heights are expressed in feet.

$$R = \{5,10,100,500\}\ yr.$$

For building i, let B_i denote its footprint polygon and let $D_{h,r}(x)$ denote modeled flood depth at location x for hazard h and return period r. Building-scale flood depth was defined as the maximum modeled depth within the footprint:

$$d_{i,h,r} = \underset{x \in B_i}{max} D_{h,r}(x).$$

Buildings with no modeled inundation were assigned $d_{i,h,r} = 0$. The maximum-within-footprint extraction preserves localized high-depth cells that may be missed by centroid-based or footprint-

average approaches.

For integrated multi-hazard loss estimation, combined depth at each return period was defined as:

$$d_{i,Comb,r} = max\left(d_{i,PD,r}, d_{i,FD,r}, d_{i,CD,r}\right).$$

This construction counts each building once where the hazard layers overlap. The combined layer is therefore a non-additive measure of integrated multi-hazard exposure, not a simulation of compound or sequential flooding and not the sum of hazard-specific depths.

Building geometry and gross floor area

Each building was assigned a planimetric footprint area and a LiDAR-derived maximum height[40,42]. Story count and gross floor area were inferred from these quantities. Building geometry supports occupancy-prototype assignment and provides the floor-area basis for replacement-value estimation and policy-allocation limits. The functional forms and parameter bounds are reported in the Supplementary Methods.

Occupancy prototype assignment by Bayesian fusion

Each building was assigned to a Hazus occupancy prototype, such as RES1I, RES3BI, COM3I, or IND2I, through a Bayesian fusion of geometry-based likelihoods[43,44] and tract-level prior information. For each candidate prototype, a geometry likelihood was computed from reference ranges for footprint area, story count, and gross floor area and combined with tract-level occupancy priors derived from ACS Table S2504. The maximum a posteriori prototype was assigned to each building. This procedure allows tract context to regularize the classification while retaining building-scale geometric information. Likelihood specifications, prior construction, posterior normalization, and tie-breaking rules are reported in the Supplementary Methods.

Replacement value estimation

Tract- and prototype-specific unit replacement costs were taken from the Hazus inventory[39]. Missing tract-prototype values were imputed using state-level medians for the same prototype. Unit costs were adjusted in four bounded steps: a local housing-value adjustment based on the ratio of tract- to state-level values for the relevant age cohort; a vertical adjustment for buildings taller than the canonical story count of their prototype; a bounded age factor; and tile-level winsorization[45,46] to limit the influence of sparse prototype assignments and extreme geometry.

The upstream valuation procedure produced an adjusted structure value R_i^{adj} and a bounded age factor f_i^{age}. Because the loss model is applied to full, undepreciated structure replacement cost, the central pipeline used:

$$R_i = \frac{R_i^{adj}}{f_i^{age}},$$

with f_i^{age} bounded between 0.05 and 1. The resulting value excludes land and contents. Relative spatial patterns in selected Austin, Texas tracts were compared with the publicly available Zillow Home Value Index as a plausibility check[47]. This comparison was not used to calibrate or rescale building values.

Reconstruction of building-level NFIP coverage

The OpenFEMA NFIP policy files do not identify the insured building footprint directly. Building-level coverage and written policy cost were therefore reconstructed by assigning observed policy records to physical buildings through a tiered spatial allocator.

Policy allocation was restricted to buildings in the processed inundation inventory. Each candidate assignment had to satisfy a coverage-cap constraint tied to estimated replacement value and a prototype-specific policy-count constraint scaled by gross floor area. Residential-status consistency was also enforced. Observed policy coverage amounts already contain the NFIP building-coverage limits, including the $250,000 residential and $500,000 non-residential limits. The allocation constraints govern where an observed policy can be placed; they do not create synthetic coverage.

Policies were assigned in six tiers, ordered from strongest to weakest spatial evidence: direct footprint matching, buffered footprint matching, nearest-eligible-building matching, deterministic within-tract matching constrained by FEMA flood zone, stochastic risk-distance allocation based on proximity and 100-year flood depth, and a final within-tract high-risk fallback. Distance thresholds, sub-passes, stochastic weights, policy-count constraints, tier-specific assignment totals, and unmatched-policy diagnostics are reported in the Supplementary Methods.

After allocation, reconstructed building coverage and annual written policy cost were:

$$C_i = \sum_{p \in P(i)} c_p, \; P_i^{act} = \sum_{p \in P(i)} \pi_p,$$

where $P(i)$ is the set of policy records assigned to building i, c_p is the observed building-coverage amount, and π_p is the observed annual policy cost. These quantities carry observed policy records to building scale. They are reconstructed building-level assignments, not directly observed parcel-level contracts or optimal coverage allocations.

First-floor elevation from surveyed records

First-floor height was derived from two Elevation Certificate fields in the OpenFEMA policy records: lowest-floor elevation and lowest adjacent grade. Their difference gives the height of the reported lowest floor above local grade. Records were retained when both elevations were present and physically plausible, and records whose lowest floor represented a basement, enclosure, or subgrade crawlspace were excluded from the first-floor-height sample. Repeated policy-term records with identical tract,

elevation pair, occupancy class, and flood-zone family were collapsed, leaving 2,378,132 distinct elevation records.

First-floor-height strata were constructed by county, occupancy class (residential or non-residential), and A/V/X flood-zone family. A county-level stratum was used when it contained at least 100 observations. Assignment otherwise fell back to state by occupancy and zone family, state by occupancy, and finally national occupancy by zone family. The central specification uses the stratum median. The first and third quartiles define the low and high sensitivity scenarios. Under this hierarchy, 37% of study buildings received a county-level estimate, 38% a state-by-zone-family estimate, 25% a state-level estimate, and fewer than 0.1% a national estimate. Because Elevation Certificates are disproportionately available for insured and post-FIRM properties in mapped flood zones, the resulting strata may not fully represent buildings without certificates; the quartile scenarios are used to assess this source of uncertainty.

Hazard-specific direct loss estimation

Direct event loss was estimated separately for the pluvial, fluvial, coastal, and combined hazard layers. Let e_i denote the assigned first-floor height and H_i the building height. Effective depth above the first floor was:

$$a_{i,h,r} = max\left(d_{i,h,r} - e_i, 0\right).$$

Where a valid building height was available, this quantity was capped at the available height above the first floor:

$$a_{i,h,r} = min\left[a_{i,h,r}, max\left(H_i - e_i, 0\right)\right].$$

Each occupancy prototype was mapped to a Hazus-family piecewise-linear above-floor depth-damage curve. Let $f\left(c_i\right)$ denote the curve family assigned to prototype c_i, $D_{f(c_i)}\left(\cdot\right)$ the corresponding curve, and η_{c_i} a prototype-specific multiplier. The above-floor damage ratio was:

$$\delta_{i,h,r}^{above} = \eta_{c_i} D_{f(c_i)}\left(a_{i,h,r}\right).$$

Below-grade damage was represented using the U.S. Army Corps of Engineers EGM 04 – 01 structure-damage functions. Because basement presence is not observed for each building, a local basement probability q_i was estimated from the share of basement-type foundation records in the building's county-by-occupancy stratum, with a minimum of 100 records. Missing county estimates fell back to state-by-occupancy strata with at least 50 records, followed by Census-division basement prevalence for residential buildings and a national non-residential prevalence of 16.7%. Basement presence was then assigned through a Bernoulli draw with a fixed random seed of 12345:

$$B_i \sim Bernoulli\left(q_i\right).$$

The central assignment identifies 3,110,951 buildings, or 24.1% of the processed inventory, as having a basement. Results under an alternative seed are included in the sensitivity analysis.

Let $z_{i,h,r} = d_{i,h,r} - e_i$ denote water level relative to the first floor, and let $D_s^{WB}(z)$ and $D_s^{NB}(z)$ denote the EGM 04 – 01 with-basement and without-basement structure-damage functions for story class s. The below-grade increment was:

$$\delta_{i,h,r}^{bas} = B_i \, (d_{i,h,r} > 0) \, max \left[D_s^{WB}(z_{i,h,r}) - D_s^{NB}(z_{i,h,r}), 0 \right].$$

Buildings with first-floor heights of at least 4 ft were also assigned an enclosure, foundation, and service-equipment component beneath the elevated floor. This component represented 10% of structure value and increased in proportion to the fraction of the grade-to-floor space filled by water:

$$\delta_{i,h,r}^{enc} = 0.10 \, (e_i \geq 4 \; ft) \, (d_{i,h,r} > 0) \, \frac{min(d_{i,h,r}, e_i)}{max\left(e_i, 10^{-3}\right)}.$$

The total damage ratio was bounded between zero and one:

$$\delta_{i,h,r} = min \left[max \left(\delta_{i,h,r}^{above} + \delta_{i,h,r}^{bas} + \delta_{i,h,r}^{enc}, 0 \right), 1 \right],$$

and direct event loss was:

$$L_{i,h,r} = \delta_{i,h,r} R_i.$$

No global rescaling was applied to the damage functions, and no parameter was calibrated to reproduce an external national loss total.

Expected annual loss

Expected annual loss was calculated by integrating the exceedance-loss curve. Let T_j, for $j = 1, \ldots, J$, denote the return periods on the integration grid, ordered from the most frequent to the rarest event, and let

$$p_j = \frac{1}{T_j}$$

denote annual exceedance probability. Before interpolation, the modeled depth sequence for each building and hazard was constrained to be nondecreasing with return period. Depths were then interpolated between the four available return-period layers in log-return-period space onto a 33-point grid spanning 5 to 500 years. Event loss $L_{i,h,j}$ was evaluated for building i, hazard h, and grid point j.

Building-level EAL was calculated by trapezoidal integration over annual exceedance probability:

$$EAL_{i,h} = \sum_{j=1}^{J-1} \frac{L_{i,h,j} + L_{i,h,j+1}}{2} \left(p_j - p_{j+1} \right) + p_J L_{i,h,J}.$$

The final term carries the 500-year event loss across events rarer than the 500-year return period. Events more frequent than the 5-year return period are not represented in the hazard layers. Alternative integration conventions, including lower- and upper-step approximations over the original return-period layers and a specification restricted to return periods of 10 years and longer, are reported in the Supplementary Information. These alternatives affect the absolute loss estimates but have limited effects on burden-channel shares and tract-level spatial patterns; rank correlations of the tract-level burden metrics against the central specification remain at or above 0.86 under every alternative tested (Supplementary Table S16).

Expected covered payout, public shortfall, and NFIP shortfall

For each building, reconstructed building coverage C_i and annual written policy cost P_i^{act} were combined with hazard-specific event losses to calculate expected covered payout, public shortfall, and NFIP shortfall.

A representative per-event deductible of $D = \$1{,}500$ was applied before the building-coverage limit. Covered payout at each return-period grid point was:

$$Q_{i,h,j} = min\left[max\left(L_{i,h,j} - D, 0\right), C_i\right].$$

For buildings without reconstructed NFIP coverage, $C_i = 0$, and covered payout is zero. Expected covered payout was annualized using the same integration grid and trapezoidal rule as EAL:

$$Pay_{i,h} = \sum_{j=1}^{J-1} \frac{Q_{i,h,j} + Q_{i,h,j+1}}{2}\left(p_j - p_{j+1}\right) + p_J Q_{i,h,J}.$$

Building-level public shortfall, also referred to as uncovered loss, was defined as:

$$S_{i,h}^{pub} = EAL_{i,h} - Pay_{i,h}.$$

This quantity includes loss arising through nonparticipation, the portion of an insured loss absorbed by the deductible, and loss exceeding reconstructed building coverage.

Building-coverage risk-premium base

The modeled payouts represent building damage only, while written policy cost contains premium components associated with both building and contents coverage and may also include fees, surcharges, and other policy components. The premium base was therefore adjusted to match the scope of the modeled payouts.

For each state s, we calculated ρ_s, the ratio of total insurance premium to total written policy cost, and b_s, the building share of combined building-and-contents coverage. For a building located in state $s(i)$, the allocated building-coverage risk-premium base was:

$$P_i^{risk} = P_i^{act} \rho_{s(i)} b_{s(i)}.$$

Nationally, the insurance-premium share is approximately 0.82 and the building-coverage share is approximately 0.81. Their product places the building-coverage risk-premium base at about 66% of total written policy cost, or approximately $\$ 3.0$ billion per year across the reconstructed portfolio.

Allocating premium in proportion to coverage assumes that building and contents premiums are proportional to their coverage amounts within each state. The resulting base is not a pure actuarial risk premium because the available policy fields do not separately identify every expense, underwriting, and increased-cost-of-compliance component. To the extent that such components remain in the base, the estimated NFIP shortfall is conservative.

Building-level NFIP shortfall was defined as the positive difference between expected covered payout and the allocated building-coverage risk-premium base:

$$S_{i,h}^{NFIP} = max\left(Pay_{i,h} - P_i^{risk}, 0\right).$$

NFIP shortfall is used as a measure of premium-deficit pressure within the reconstructed portfolio. It is not an estimate of program solvency because it does not include administrative costs, capital requirements, reserves, reinsurance, or other program-level balance-sheet components.

The portion of expected covered payout supported by the allocated premium base was defined as premium-funded payout:

$$F_{i,h} = Pay_{i,h} - S_{i,h}^{NFIP} = min\left(Pay_{i,h}, P_i^{risk}\right).$$

Building-level burden shares were calculated for buildings with positive EAL:

$$ULS_{i,h} = \frac{S_{i,h}^{pub}}{EAL_{i,h}},$$

$$NFS_{i,h} = \frac{S_{i,h}^{NFIP}}{EAL_{i,h}},$$

where $ULS_{i,h}$ is the uncovered-loss share and $NFS_{i,h}$ is the NFIP-shortfall share. The remaining share is premium-funded payout:

$$PFS_{i,h} = \frac{F_{i,h}}{EAL_{i,h}} = 1 - ULS_{i,h} - NFS_{i,h}.$$

The public-shortfall and NFIP-shortfall shares therefore do not sum to one because the remaining portion of EAL is premium-funded payout.

Aggregate net premium gap

NFIP shortfall sums positive building-level gaps and does not allow premium amounts above expected payout for one building to offset deficits for another. For comparison with program-level estimates, we also calculated the aggregate net premium gap under the combined hazard layer:

$$G_{Comb}^{net} = \sum {}_i Pay_{i,Comb} - \sum {}_i P_i^{risk}.$$

The premium total in this calculation includes all processed buildings with an allocated NFIP policy, including buildings with zero combined-hazard EAL. The aggregate net gap therefore differs from:

$$\sum {}_i S_{i,Comb}^{NFIP} = \sum {}_i max\left(Pay_{i,Comb} - P_i^{risk}, 0\right),$$

because the net calculation allows premium amounts above expected payout to offset deficits elsewhere in the reconstructed portfolio.

Accounting identity and incidence

Expected annual loss decomposes exactly into public shortfall, NFIP shortfall, and premium-funded payout:

$$EAL_{i,h} = S_{i,h}^{pub} + S_{i,h}^{NFIP} + F_{i,h}.$$

Equivalently,

$$EAL_{i,h} = S_{i,h}^{pub} + Pay_{i,h},$$

and

$$Pay_{i,h} = S_{i,h}^{NFIP} + F_{i,h}.$$

Each expected dollar of modeled loss is therefore assigned to one of three destinations. Public shortfall is loss that remains outside expected NFIP payout. It falls first on uninsured or under-covered property owners and their communities and may later generate demand for post-disaster assistance. The term does not denote government expenditure. NFIP shortfall is the portion of expected covered payout that exceeds the allocated building-coverage risk-premium base. Premium-funded payout is the portion of expected payout supported by that base.

A sensitivity grid varied the first-floor-height scenario, integration rule, return-period floor, below-grade and enclosure pathways, basement assignment, premium-base definition, and pluvial-drainage treatment (Supplementary Information, Part D). All main-text results use the central specification.

Aggregation and tract-level burden metrics

Building-level outputs were aggregated to census tracts. The building-level dataset spans 76,659 tracts, while the tract-level spatial layer contains 83,954 tracts. The regime analysis includes 74,240 tracts with positive combined-hazard EAL and the quantities required for classification.

For tract t and hazard h, tract-level expected annual loss was:

$$EAL_{t,h} = \sum_{i \in t} EAL_{i,h}.$$

Tract-level public shortfall was:

$$Pub_{t,h} = \sum_{i \in t} S_{i,h}^{pub}.$$

Tract-level NFIP shortfall was:

$$NFIP_{t,h} = \sum_{i \in t} S_{i,h}^{NFIP}.$$

Allocated written policy cost was aggregated as:

$$Prem_t^{act} = \sum_{i \in t} P_i^{act}.$$

The tract written-policy-cost total includes amounts allocated to processed buildings with zero combined-hazard EAL. These amounts do not contribute to building-level NFIP shortfall but remain part of the local premium volume used in the tract metric.

Three tract-level measures were constructed under the combined hazard layer. Public loss absorption was

$$PLA_t = \frac{Pub_{t,Comb}}{EAL_{t,Comb}},$$

and measures the share of tract-level EAL that remains outside expected NFIP payout.

For tracts with positive allocated written policy cost, premium-deficit pressure was:

$$PDP_t = \frac{NFIP_{t,Comb}}{Prem_t^{act}}.$$

The denominator is the tract's allocated written policy cost, so the ratio expresses NFIP shortfall relative to the premium volume recorded locally. The numerator is calculated against the building-coverage risk-premium base, whereas the denominator is total allocated written policy cost; PDP is therefore a relative indicator across tracts, not a like-for-like accounting ratio. It should not be interpreted as a break-even threshold, an actuarial loss ratio or a measure of total program liability. Tracts with both zero allocated written policy cost and zero NFIP shortfall were assigned a PDP value of zero. For clustering only, tracts with zero allocated written policy cost and positive NFIP shortfall were assigned the upper value used to

limit the influence of extreme ratios. Uncapped values were retained for descriptive summaries where the denominator was positive.

Total shortfall burden was:

$$TBA_t = Pub_{t,Comb} + NFIP_{t,Comb}.$$

Tract-level burden regimes

The tract classification used three variables:

$$PLA_t,\ \ log\left(1 + PDP_t\right),\ \ log\left(1 + TBA_t\right).$$

To limit the influence of extreme values during model fitting, PDP and total shortfall burden were capped at their respective 99th percentiles before logarithmic transformation. These caps were applied only to the classification variables; uncapped values were retained for regime summaries. The three variables were then centered on their medians and scaled by the difference between their 90th and 10th percentiles.

A four-component Gaussian mixture model with full covariance matrices was fitted to a random sample of 20,000 valid tracts and used to predict component membership for all classified tracts. The model used three initializations, a maximum of 150 iterations, covariance regularization of 10^{-6}, and a random seed of 42. A four-component solution was retained because it provided an interpretable separation of burden magnitude and insurance state. Component assignment was probabilistic. Tracts with a maximum posterior assignment probability below 0.70 were identified as uncertain, and assignment uncertainty is reported in the Supplementary Information.

The components were labelled from their empirical composition. The component with the largest share of tracts combining public loss absorption of at least 0.99 with zero NFIP shortfall was labelled minor burden, uninsured. Among the remaining components, the component carrying the largest total NFIP shortfall was labelled major burden, substantially insured. Of the two remaining components, the one carrying the larger total public shortfall was labelled major burden, mostly uncovered, and the remaining low-burden component was labelled minor burden, some coverage.

These labels describe the empirical characteristics of the fitted components and do not represent fixed actuarial thresholds. The components overlap in the classification space, and posterior assignment uncertainty is reported in Supplementary Fig. S8 and Supplementary Table S7. Hazard type and geographic location were not included in the classification, allowing the hazard composition and spatial distribution of each regime to be examined separately as empirical results.

Data availability

The flood-hazard fields used in this study are from the defended layers of the Fathom U.S. flood model and are subject to licensing restrictions; these data cannot be redistributed by the authors. National

Flood Insurance Program (NFIP) policy records are publicly available through OpenFEMA, and FEMA flood-zone data are available through the National Flood Hazard Layer. Other publicly available data sources are identified and cited in the Methods. Derived census-tract-level data supporting the analyses and figures are available in the project GitHub repository (https://github.com/Koazizi/Flood-Insurance). Additional supporting data, tables and diagnostics are provided in the Supplementary Information.

Code availability

The analysis code used to reconstruct building-level losses, allocate NFIP coverage and classify tract-level burden regimes is available at the project GitHub repository (https://github.com/Koazizi/Flood-Insurance). The repository includes environment specifications, processing scripts and supporting tables, but does not include licensed Fathom flood-hazard fields.

Declarations

Author Contributions Statement

K.A. conceptualized the study. K.A, H.D, and M.G. developed the methodology and software, curated and analyzed the data, produced the visualizations, and wrote the original draft. R.P.B, D.N., B.R.S contributed to the methodology. R.P.B, D.N., P.B, and B.R.S. reviewed and edited the manuscript. All authors discussed the results, contributed to manuscript revisions, and approved the final version.

Competing Interests

The authors declare no competing interests.

Funding Statement

This work was supported by the National Oceanic and Atmospheric Administration (NOAA) AdSci program (award no. NA21OAR4310285), the National Science Foundation (NSF) RISE program (award no. 2228205) as part of Project CRESSLE, PlanetTexas2050, a research grand challenge at The University of Texas at Austin. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NOAA or NSF.

References

1. Aon. (2026) Climate and Catastrophe Insight – Aon Report. https://www.aon.com/getmedia/3fd8e3d3-1e2f-4df8-89a8-d376b753b3a0/20260120-cci-2026.pdf (2026)
2. 2024 An active year of U.S. billion-dollar weather and climate disasters | NOAA Climate.gov. https://www.climate.gov/news-features/blogs/beyond-data/2024-active-year-us-billion-dollar-weather-and-climate-disasters

3. Davenport FV, Burke M, Diffenbaugh NS (2021) Contribution of historical precipitation change to US flood damages. *Proceedings of the National Academy of Sciences* 118, e2017524118
4. Mills E (2005) Insurance in a climate of change. Science 309:1040–1044
5. Rogers JS, Maneta MM, Sain SR, Madaus LE, Hacker JP (2025) The role of climate and population change in global flood exposure and vulnerability. *Nature Communications 2025 16:1* 16, 1287-
6. Wing OEJ, Pinter N, Bates PD, Kousky C (2020) New insights into US flood vulnerability revealed from flood insurance big data. *Nature Communications 2020 11:1* 11, 1444-
7. Bates PD et al (2020) Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates. *Water Resour. Res.* 57, eWR028673 (2021)
8. Wing OEJ et al (2022) Inequitable patterns of US flood risk in the Anthropocene. *Nature Climate Change 2022 12:2* 12, 156–162
9. Natural Disasters and the Homeowners Insurance Market | Congress.gov | Library of Congress. https://www.congress.gov/crs-product/IN12375
10. Amornsiripanitch N, Biswas S, Orellana-Li J, Zink D (2025) Measuring flood underinsurance in the USA. *Nature Climate Change 2025 15:9* 15, 971–977
11. Flood Insurance FEMA's New Rate-Setting Methodology Improves Actuarial Soundness but Highlights Need for Broader Program Reform | U.S. GAO. https://www.gao.gov/products/gao-23-105977
12. FEMA Flood Maps Better Planning and Analysis Needed to Address Current and Future Flood Hazards | U.S. GAO. https://www.gao.gov/products/gao-22-104079
13. Flood Insurance in Communities at Risk of Flooding | Congressional Budget Office. https://www.cbo.gov/publication/60042
14. The National Flood Insurance Program Financial Soundness and Affordability | Congressional Budget Office. https://www.cbo.gov/publication/53028
15. NFIP's Pricing Approach State Profiles | FEMA.gov. https://www.fema.gov/flood-insurance/risk-rating/profiles
16. Kousky C, Kunreuther H (2014) Addressing Affordability in the National Flood Insurance Program. 2345737614500018(01):1450001*https://doi.org/10.1142/S*
17. Mulder P, Kousky C (2023) Risk Rating without Information Provision. *AEA Papers and Proceedings* 113, 299–303
18. Ortega F, Petkov I (2024) To Improve Is to Change? The Effects of Risk Rating 2.0 on Flood Insurance Demand. *IZA Discussion Papers* https://ideas.repec.org/p/iza/izadps/dp17021.html
19. Smiley KT et al (2022) Social inequalities in climate change-attributed impacts of Hurricane Harvey. *Nature Communications 2022 13:1* 13, 3418-
20. Nelson-Mercer B, Kim T, Tran VN, Ivanov V (2025) Pluvial flood impacts and policyholder responses throughout the United States. *npj Natural Hazards 2025 2:1* 2, 8-

21. Bakkensen LA, Barrage L (2022) Going Underwater? Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics. Rev Financ Stud 35:3666–3709
22. Gourevitch JD et al (2023) Unpriced climate risk and the potential consequences of overvaluation in US housing markets. *Nature Climate Change 2023 13:3* 13, 250–257
23. Hino M, Burke M (2021) The effect of information about climate risk on property values. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2003374118
24. Hu Z (2022) Social interactions and households' flood insurance decisions. J financ econ 144:414–432
25. Netusil NR, Kousky C, Neupane S, Daniel W, Kunreuther H (2021) The Willingness to Pay for Flood Insurance. Land Econ 97:17–38
26. Katherine H, Wagner BR et al (2022) Adaptation and Adverse Selection in Markets for Natural Disaster Insurance. Am Econ J Econ Policy 14:380–421
27. Nayak A, Zhang M, Gentine P, Lall U (2025) Catastrophic hyperclustering and recurrent losses: diagnosing U.S. flood insurance insolvency triggers. *npj Natural Hazards 2025 2:1* 2, 83-
28. Flood Data Viewers and Geospatial Data | FEMA.gov. https://www.fema.gov/flood-maps/national-flood-hazard-layer
29. FIMA NFIP Redacted Policies - v2 | FEMA.gov. https://www.fema.gov/openfema-data-page/fima-nfip-redacted-policies-v2
30. Wing OEJ et al (2017) Validation of a 30 m resolution flood hazard model of the conterminous United States. Water Resour Res 53:7968–7986
31. Introduction to the National Flood Insurance Program (NFIP) | Congress.gov | Library of Congress. https://www.congress.gov/crs-product/R44593
32. Botzen WJW, Aerts JCJH, Van Den Bergh JC (2009) J. M. Dependence of flood risk perceptions on socioeconomic and objective risk factors. Water Resour Res 45
33. Haasnoot M, Di Fant V, Kwakkel J, Lawrence J (2024) Lessons from a decade of adaptive pathways studies for climate adaptation. Glob Environ Change 88:102907
34. Paarlberg AE, van der Bolt B, Vignola R, Le Clec'h S, Smolenaars WJ (2025) Exploring the use of adaptation tipping points: A systematic review of definitions, characteristics and applications. Environ Sci Policy 172:104211
35. Werners SE et al (2013) Thresholds, tipping and turning points for sustainability under climate change. Curr Opin Environ Sustain 5:334–340
36. Moftakhari HR, Salvadori G, AghaKouchak A, Sanders BF, Matthew RA (2017) Compounding effects of sea level rise and fluvial flooding. *Proc. Natl. Acad. Sci. U. S. A.* 114, 9785–9790
37. Cano Pecharroman L, Hahn CH (2024) Exposing disparities in flood adaptation for equitable future interventions in the USA. *Nature Communications 2024 15:1* 15, 8333-
38. Shi L, Moser S (2021) Transformative climate adaptation in the United States: Trends and prospects. Sci (1979). 372

39. Hazus Documentation | FEMA.gov https://www.fema.gov/flood-maps/tools-resources/flood-map-products/hazus/documentation
40. Che Y et al (2024) 3D-GloBFP: the first global three-dimensional building footprint dataset. Earth Syst Sci Data 16:5357–5374
41. Census Bureau Data and Maps https://www.census.gov/data.html
42. U.S. national categorical mapping of building heights by block group from Shuttle Radar Topography Mission data | U.S. Geological Survey. https://www.usgs.gov/data/us-national-categorical-mapping-building-heights-block-group-shuttle-radar-topography-mission
43. Stewart R et al (2016) A Bayesian machine learning model for estimating building occupancy from open source data. *Natural Hazards 2016 81:3* 81, 1929–1956
44. Urban M, Stewart R, Basford S, Palmer Z, Kaufman J (2023) Estimating building occupancy: a machine learning system for day, night, and episodic events. *Natural Hazards 2023 116:2* 116, 2417–2436
45. Rousseeuw PJ, Leroy AM (1987) Robust Regression and Outlier Detection. https://doi.org/10.1002/0471725382 doi:10.1002/0471725382
46. Wilcox RR (2021) Introduction to Robust Estimation and Hypothesis Testing. Introduction Robust Estimation Hypothesis Test 1–900. 10.1016/C2019-0-01225-3
47. Housing Data - Zillow Research https://www.zillow.com/research/data/

Figures

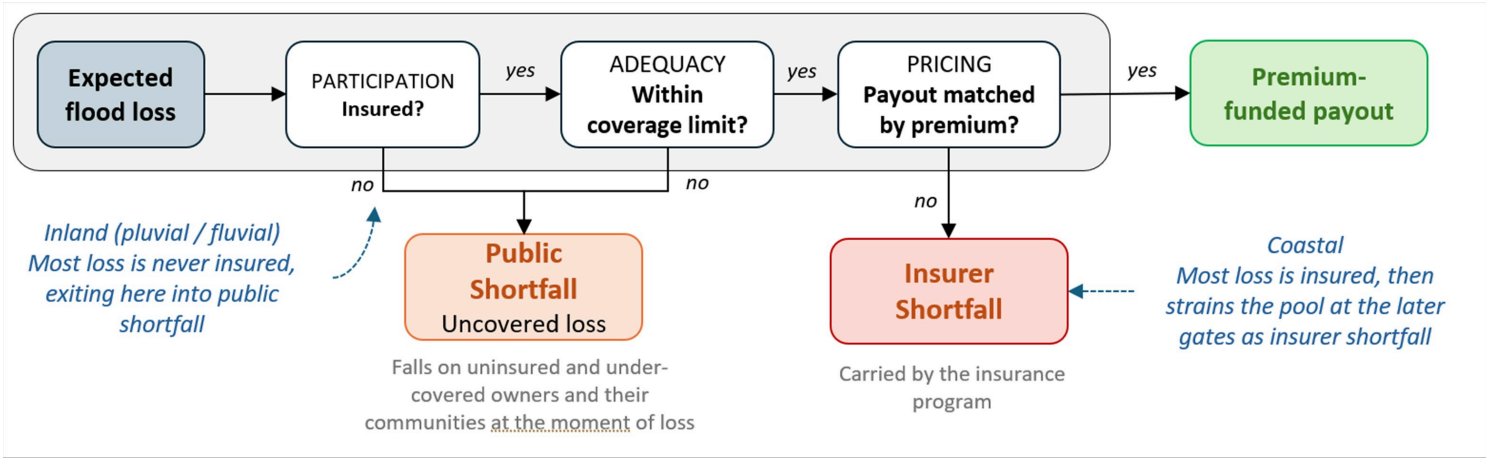


Figure 1

Sequential allocation of expected flood loss across burden channels. Each unit of expected annual loss passes through three conditions: insurance participation, coverage adequacy and premium pricing. Loss associated with nonparticipation, deductibles or coverage limits becomes public shortfall. Expected covered payout above the allocated building-coverage risk-premium base becomes NFIP shortfall, and

the remaining covered payout is premium-funded. Grey arrows show the available routes, and blue annotations indicate the pathways most common for each hazard.

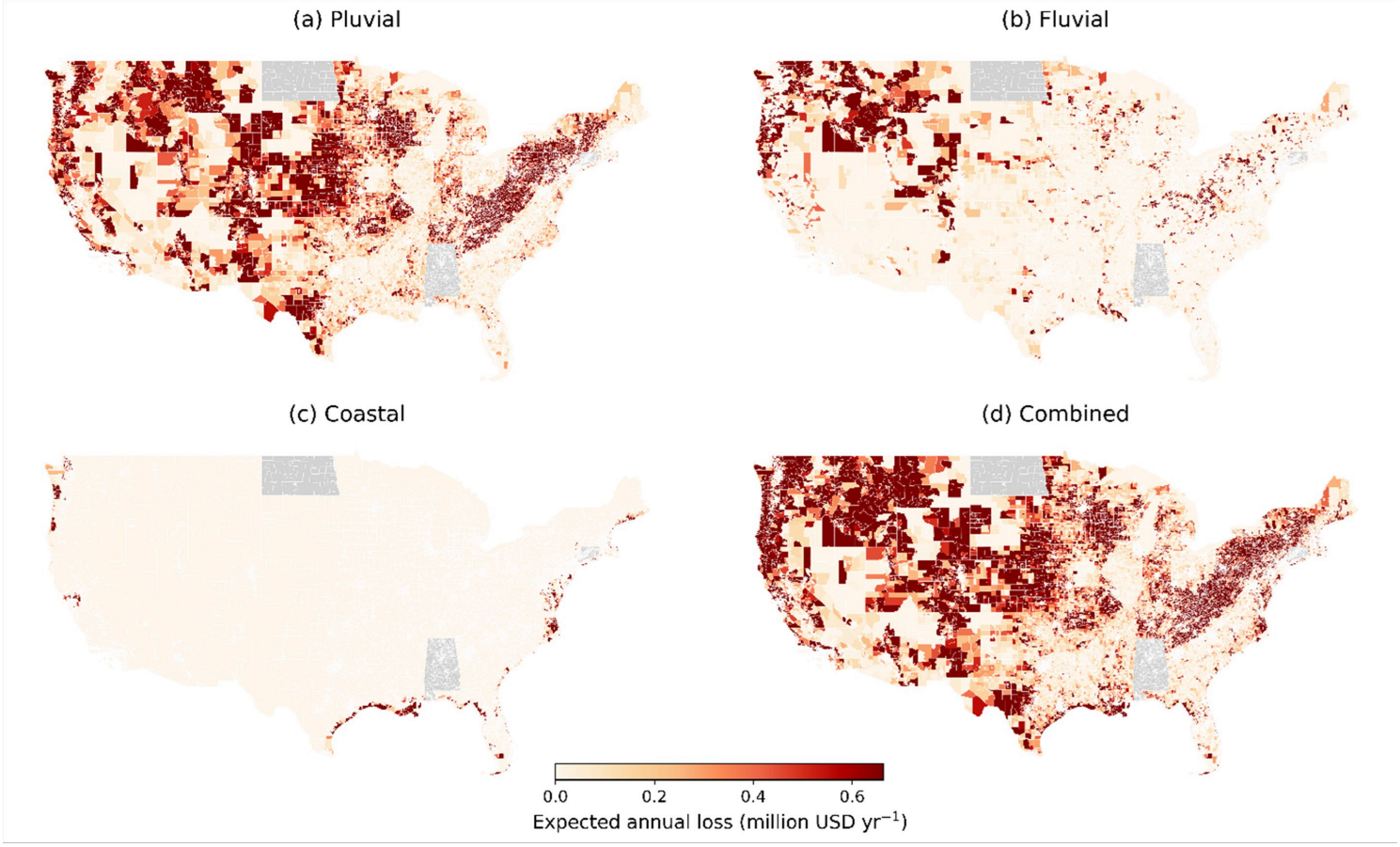


Figure 2

National geography of hazard-specific expected annual loss. Census tract-level expected annual loss is mapped separately for (a) pluvial, (b) fluvial and (c) coastal flooding, and (d) the combined hazard layer across the CONUS. Values are aggregated from building-level modeled losses annualized over the return-period set used in the analysis. The combined panel uses the maximum hazard-specific depth at each return period before loss estimation and represents an integrated multi-hazard loss layer, not the sum of the three hazard-specific layers. Darker shading indicates higher tract-level expected annual loss, and light grey indicates no-data tracts. Alabama, Connecticut and North Dakota are absent from the processed building inventory (Methods) and appear as no-data areas here and in subsequent maps. Census tracts are population-based units, so visual area is not proportional to national loss contribution; the state-level view in Extended Data Fig. 2 summarizes the corresponding concentration and composition.

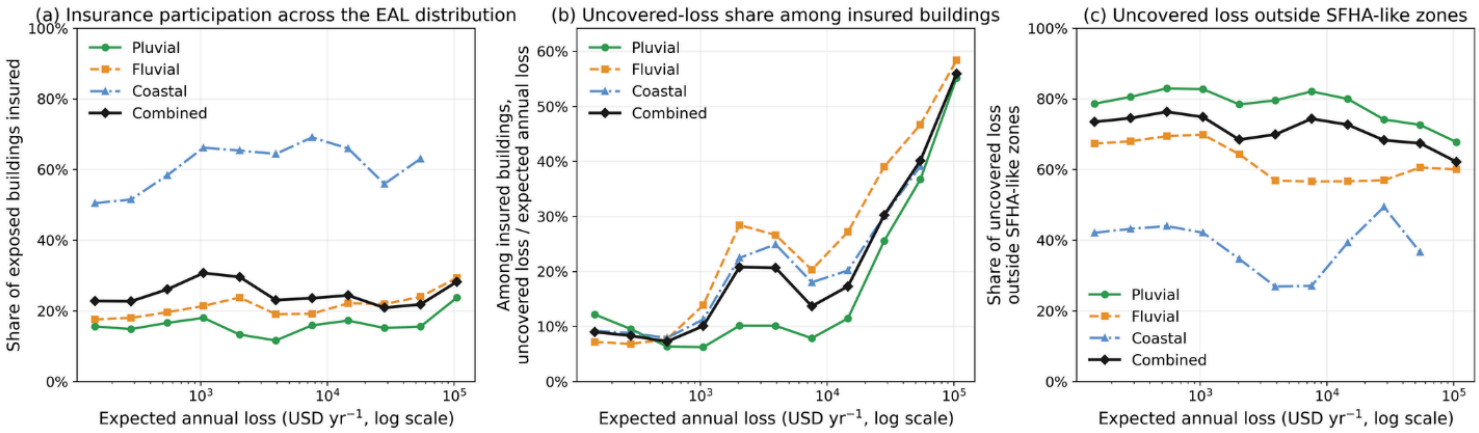


Figure 3

Hazard-specific pathways into and around coverage. The panels compare how flood losses enter reconstructed NFIP coverage or remain outside expected payout across pluvial, fluvial, coastal and combined hazard layers. Panel (a) shows the share of exposed buildings that are insured across the expected annual loss distribution. Panel (b) shows, among insured buildings, the uncovered-loss share, computed as summed uncovered loss divided by summed expected annual loss within each loss bin. Panel (c) shows the share of uncovered loss located outside FEMA Special Flood Hazard Area (SFHA) zones across the same expected-loss distribution. The panels distinguish whether hazard-specific differences arise mainly from insurance participation, residual loss among insured buildings, or the extent to which uncovered loss falls outside the mapped regulatory core.

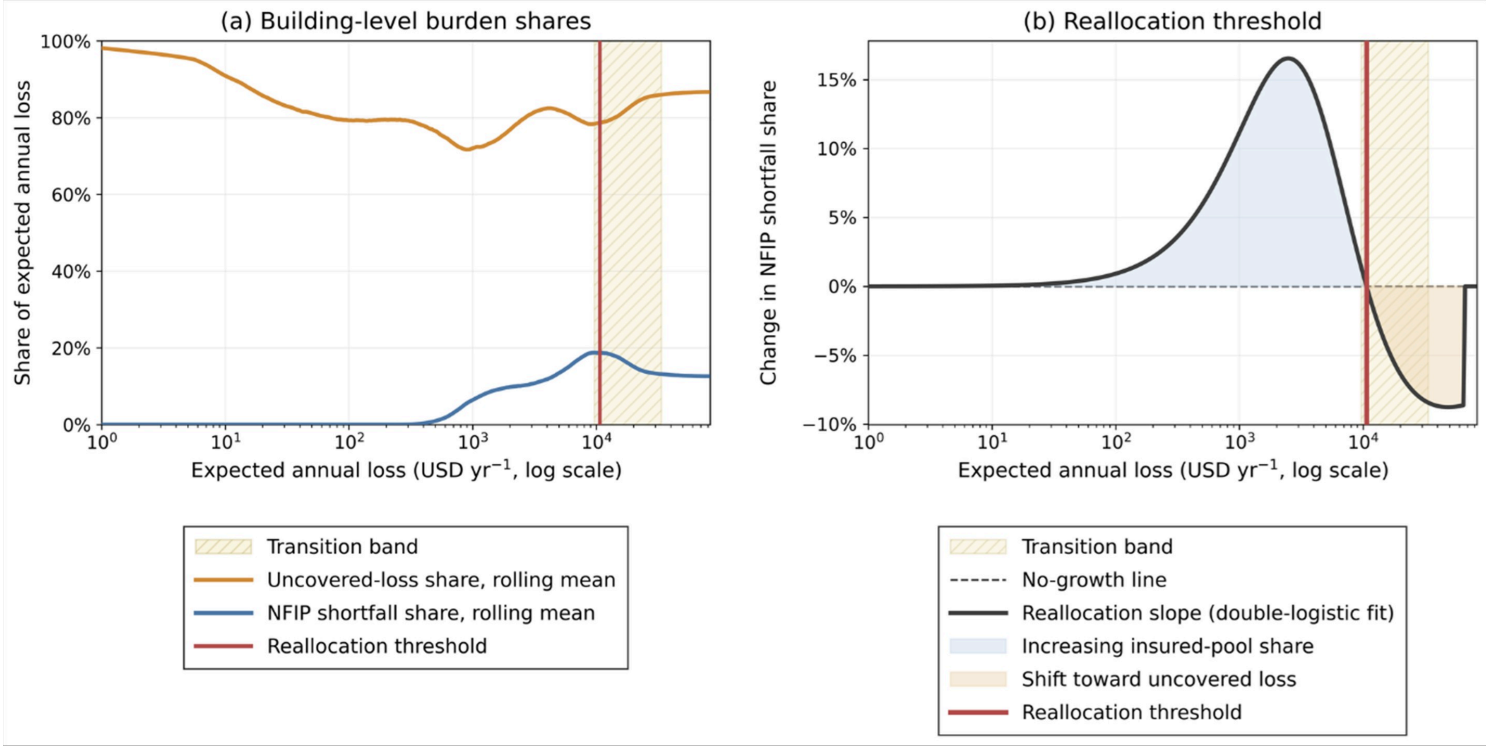


Figure 4

Burden reallocation across the expected-loss spectrum under the combined hazard layer. Panel (a) plots the building-level uncovered-loss and NFIP-shortfall shares against expected annual loss, with each share measured relative to building-level EAL. Curves are rolling means over windows of approximately 224,000 buildings. The two shares do not sum to one because the remaining share is premium-funded payout. Panel (b) plots the derivative of a two-transition double-logistic profile fitted to logarithmically binned means of the NFIP-shortfall share. Blue shading marks the range in which the NFIP-shortfall share rises, orange shading marks the range in which it declines, and the vertical line marks the approximate reallocation threshold near $10,000 yr⁻¹.

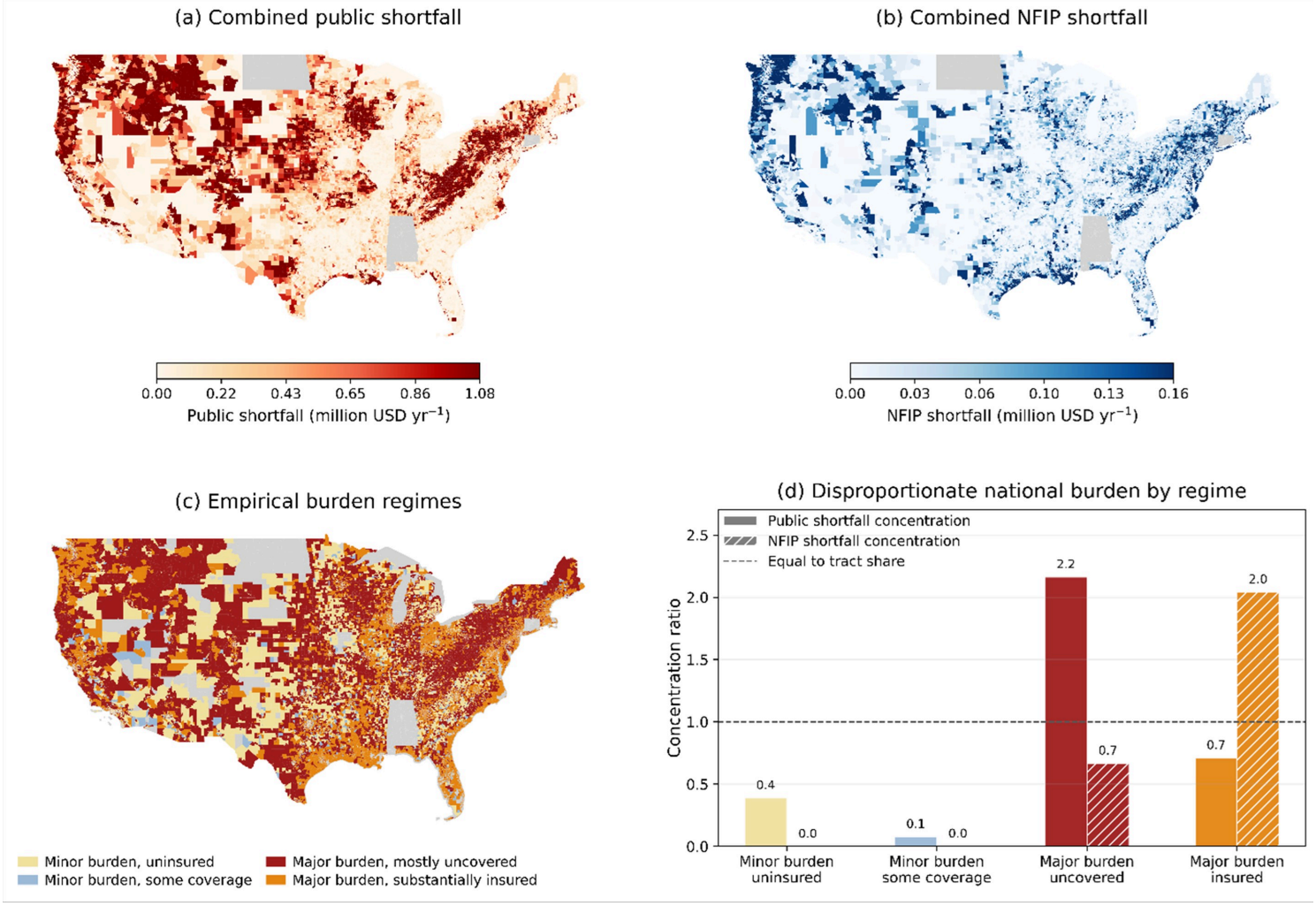


Figure 5

Tract-level burden regimes under the combined hazard layer. Panel (a) maps tract-level public shortfall and panel (b) maps tract-level NFIP shortfall, both in million USD yr⁻¹. Panel (c) shows the four empirical regimes derived from public loss absorption, log-transformed premium-deficit pressure and log-transformed total shortfall burden. Assignment is probabilistic and overlap among components is reported in the Supplementary Information. Panel (d) compares each regime's share of burden among classified tracts with its tract share; values above one indicate disproportionate concentration. Tracts

without positive combined-hazard EAL or the measures required for classification are shown in grey. Alabama, Connecticut and North Dakota are absent from the processed inventory.

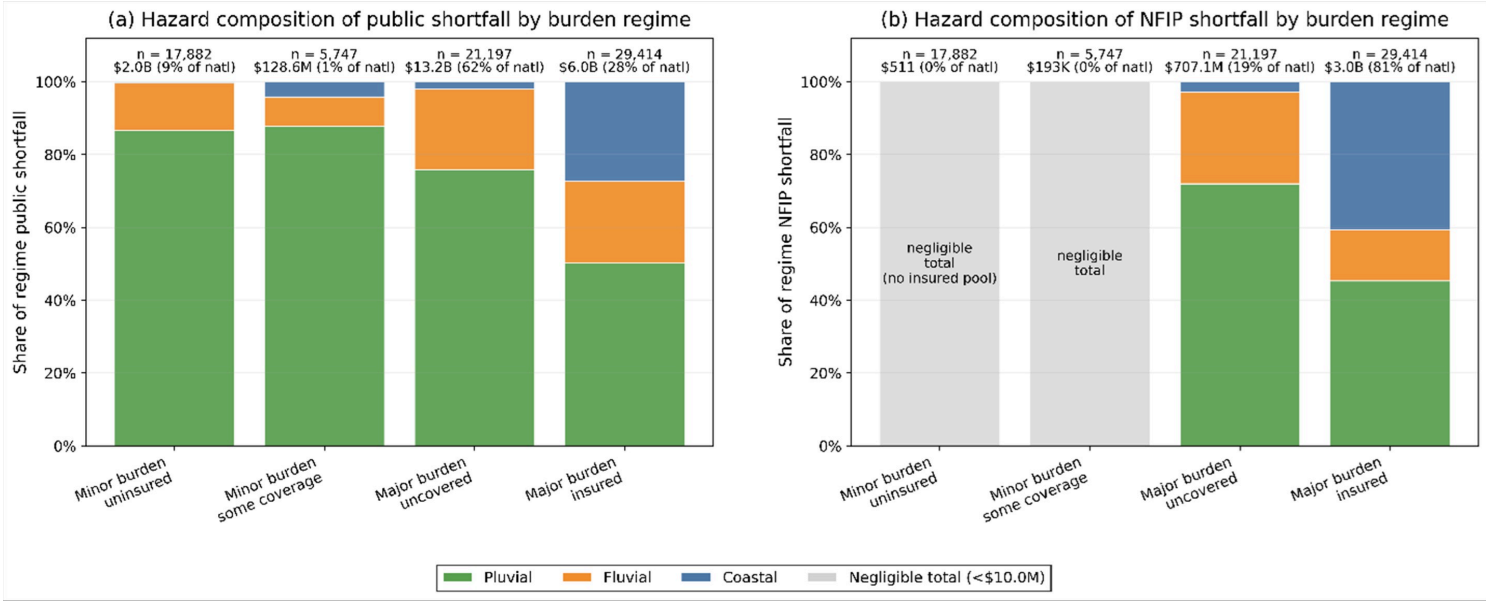


Figure 6

Hazard composition within burden regimes. Panel (a) shows the relative contributions of separately modeled pluvial, fluvial and coastal public shortfalls within each burden regime. Panel (b) shows the corresponding shares for NFIP shortfall. Grey bars indicate regimes in which the displayed channel is below $10 million and percentage composition is not interpreted. Percentages normalize the three hazard-specific estimates within each regime and channel; they are not a decomposition of the combined-layer totals in Extended Data Tables 1 and 3.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.docx
- ExtendedData.docx