Public Warning and Outreach to Reduce Pre-Evacuation Delay in Mountain Flash Flooding: Evidence from Liulimiao Town in Beijing

Zequn Wu
Tsinghua University
wuzq23@mails.tsinghua.edu.cn

Lida Huang
Tsinghua University
hld999@yeah.net

Tao Chen
Tsinghua University
chentao.b@tsinghua.edu.cn

Hui Zhang
Tsinghua University
zhhui@tsinghua.edu.cn

ABSTRACT

Flash floods compress decision time, making warning quality and outreach central to timely evacuation. Building on the revised Protective Action Decision Model (PADM), we distinguish two sources of delay: information frictions in warnings (clarity, credibility, actionability, and verification burden) and action frictions tied to preparedness, capability, and constraints. We then model evacuation initiation with a discrete-time event-history approach. We analyze a post-disaster survey of 197 households in Liulimiao Town, Beijing, following the late-July 2025 "25·7" flash-flood event. Departures are front-loaded, with 42.6% occurring within 15 minutes, yet 20.8% extend beyond 120 minutes. Latent class analysis identifies three friction profiles with distinct departure trajectories. Model-based standardization indicates that improving warning actionability and strengthening household mobilization readiness, defined by preparedness, route knowledge, and self-efficacy, especially in combination, are associated with earlier departures and substantially lower tail-delay risk.

Keywords

Flash flood, pre-evacuation delay, warning actionability, public information outreach, discrete-time hazard model.

INTRODUCTION

Climate warming is intensifying heavy precipitation and increasing the likelihood of rapid-onset flood hazards (Intergovernmental Panel on Climate Change, 2021). The late-July 2025 Haihe River Basin "25·7" event, which produced mountain flash flooding in Beijing's northern mountainous districts, illustrates how these shifts translate into high-impact, short-lead-time risk. Public reports indicate rainfall totals approaching annual norms within days alongside localized extreme hourly intensities, with cascading disruptions to roads, electricity, and communications (Reuters, 2025; China Daily, 2025). Such events can outpace routine expectations and local situational awareness. When transport and communication infrastructure fail simultaneously, warning messages and local cues can become unreliable at the last mile, while household mobility can be constrained precisely when decisions must be made.

People-centered, end-to-end early warning systems are commonly described as four co-equal elements: disaster risk knowledge; detection, monitoring, and forecasting; warning dissemination and communication; and preparedness and response capability (United Nations Office for Disaster Risk Reduction, n.d.; World Meteorological Organization, n.d.). We focus on the latter two, closest to household response: last-mile communication and interpretation, and the conversion of warnings into feasible protective action when floods escalate quickly and disruptions constrain mobility. In this framing, improving message actionability and reducing verification burden target communication-side frictions, while mobilization readiness and binding constraints shape whether timely departure is feasible after receipt. Evidence shows heterogeneous flash-flood warning response, especially by trust and preparedness (Morss et al., 2016); studies in China's mountainous basins highlight short lead times, complex terrain, and local uncertainty that weaken warning-to-action links under constraint (Ren et al., 2023). Syntheses similarly stress message specificity, credibility, and actionable guidance, while noting uneven evidence across settings and warning designs (Kuller et al., 2021), and reviews of pluvial/flash-flood early warning systems (EWS) point to persistent last-mile and people-centered design gaps (Acosta-Coll et al., 2018).

Despite extensive work on flood warnings and evacuation, two gaps remain for crisis communication and outreach. Reviews note rapid growth in evacuation modeling and optimization, yet behavioral components—especially time-to-depart under short lead times—remain comparatively under-developed (Li et al., 2025). Related studies show that preparedness, self-efficacy, and trust pathways differ across vulnerable groups (Li & Gilbert, 2024), evacuation strategies in mountainous China vary by terrain and settlement type (Huang et al., 2025), and peer-to-peer transmission across multiplex channels shapes who receives warnings and when under disruption (Koll et al., 2023). First, many studies emphasize binary outcomes or intentions, while warning-to-departure time is less systematically analyzed, even though pre-evacuation delay can dominate risk when lead times are short (Proulx & Fahy, 1997). Second, variation in comprehension, trust, and actionability is well documented, but their effects on the distribution of departure timing—especially the long tail most relevant for entrapment risk and operations—are less often quantified. We address these gaps by building on the revised Protective Action Decision Model (PADM; Lindell & Perry, 2012) and developing a PADM-informed friction framework. The framework links warning messages with environmental and social cues and emphasizes three perception domains—threat perceptions, protective action perceptions, and stakeholder perceptions—preceded by reception, attention, and comprehension processes. The outcome is warning-to-departure delay, defined as time from first awareness of a credible evacuation-relevant cue to leaving home to begin evacuation (National Weather Service, n.d.). For intervention design, we apply a bottleneck lens consistent with PADM and people-centered EWS guidance, separating information frictions (losses from reception to comprehension, trust, and actionability, including verification under conflicting messages) from action frictions (losses from intention to movement due to mobilization constraints, e.g., route knowledge, resources, caregiving, and vulnerability-related barriers). Under short lead times, improving warning actionability and strengthening mobilization readiness are complementary levers for reducing tail-delay risk.

Guided by this framework, we ask two research questions. First, which information frictions and action frictions are associated with earlier evacuation initiation after warning awareness, beyond threat perception and prior experience? Second, do these frictions cluster into distinct household bottleneck profiles, and do these profiles explain heterogeneity in the tail of the evacuation-initiation delay distribution under short lead times? These questions matter because average response can obscure a smaller but operationally critical subset of households that depart dangerously late. Empirically, we analyze post-event, door-to-door household surveys from a severely affected mountainous peri-urban area in northern Beijing, China, following the "25·7" extreme rainfall and flood episode. Using post-event household survey data, we estimate when households begin evacuation, identify distinct friction profiles, and compare how feasible warning and outreach improvements could shift the delay distribution. This allows us to show more directly how better warning actionability and stronger mobilization support could increase early departures and reduce dangerously long delays.

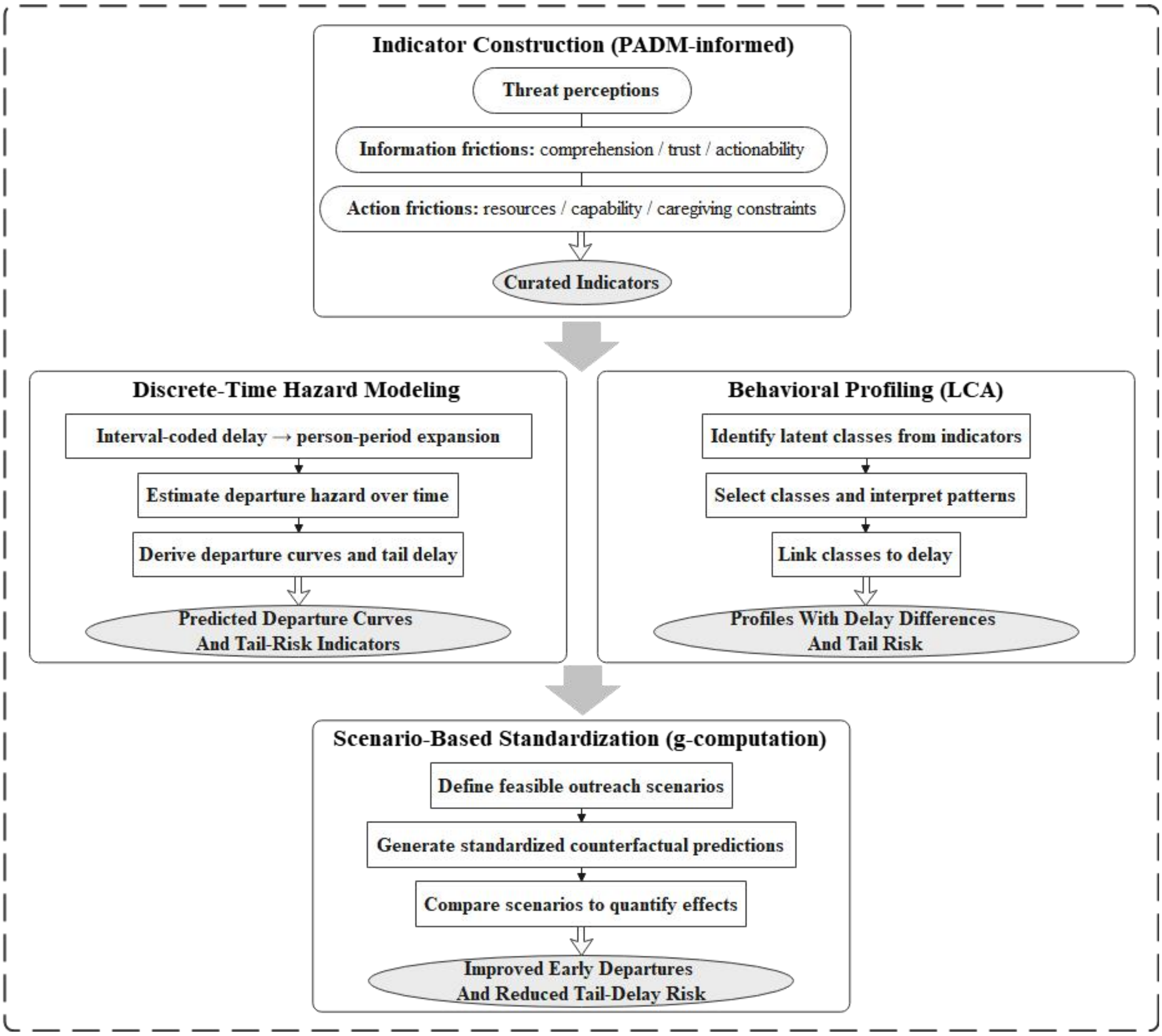

Figure 1. Analytical Framework

METHODS

Study Area and Data

This study focuses on Liulimiao Town in Huairou District, Beijing, one of the areas most severely affected by the late-July 2025 Haihe River Basin ("25·7") flood event, during which the town experienced mountain flash flooding. In Liulimiao, field debriefs and respondent accounts indicate that flooding emerged around 22:00 on July 25 and propagated along upstream–downstream villages; within roughly half an hour, communications became unavailable in some locations. Public reporting likewise documents severe "three-cut" disruptions—road access, electricity, and communications—across multiple villages in the town, conditions that plausibly hinder both last-mile warning delivery and evacuation mobility. Liulimiao was selected because the event there concentrated the conditions central to our research question: severe evacuation demand under a short-lead-time mountain flood, together with simultaneous disruption to roads, electricity, and communications that strained both last-mile warning delivery and household mobility. Its terrain, settlement pattern, infrastructure dependence, and population aging are also broadly comparable to other areas in North China and to other regions with similar characteristics, making it useful for examining mechanisms that may recur under comparable mountain-flood conditions. We conducted a post-disaster household survey in mid-November 2025, using the household as the unit of analysis and administering one questionnaire per household. To capture variation in disaster severity, we purposively selected five administrative villages spanning an impact gradient: Detiangou and Qingshiling (lightly affected) and Sunhugou, Xiwanzi, and Qifengcha (heavily affected). Recruitment combined station-based surveying and door-to-door interviews within the same villages. Station-based survey points were placed at accessible village locations, such as village committee activity rooms or other common gathering areas, to facilitate participation. Door-to-door visits were used to broaden coverage beyond these fixed points and were directed especially toward older residents, households with mobility limitations, and geographically more remote parts of the villages that were less likely to be reached through centralized survey points. Taken together, this mixed recruitment strategy was intended to cover the full village area rather than only central or convenience-accessible locations. Trained volunteers used standardized scripts and anchored probing (e.g., village broadcasts, power outages, officials' visits) to help respondents place experiences within predefined delay intervals. In total, 197 valid household questionnaires were collected. The study area and sampled villages are shown in Figure 2. Data collection was conducted with permission and operational support from local emergency management authorities and township administration. Participation was voluntary and based on informed oral consent; the survey collected no direct personal identifiers, and data were analyzed in de-identified form, stored on access-controlled devices, and used solely for research purposes.

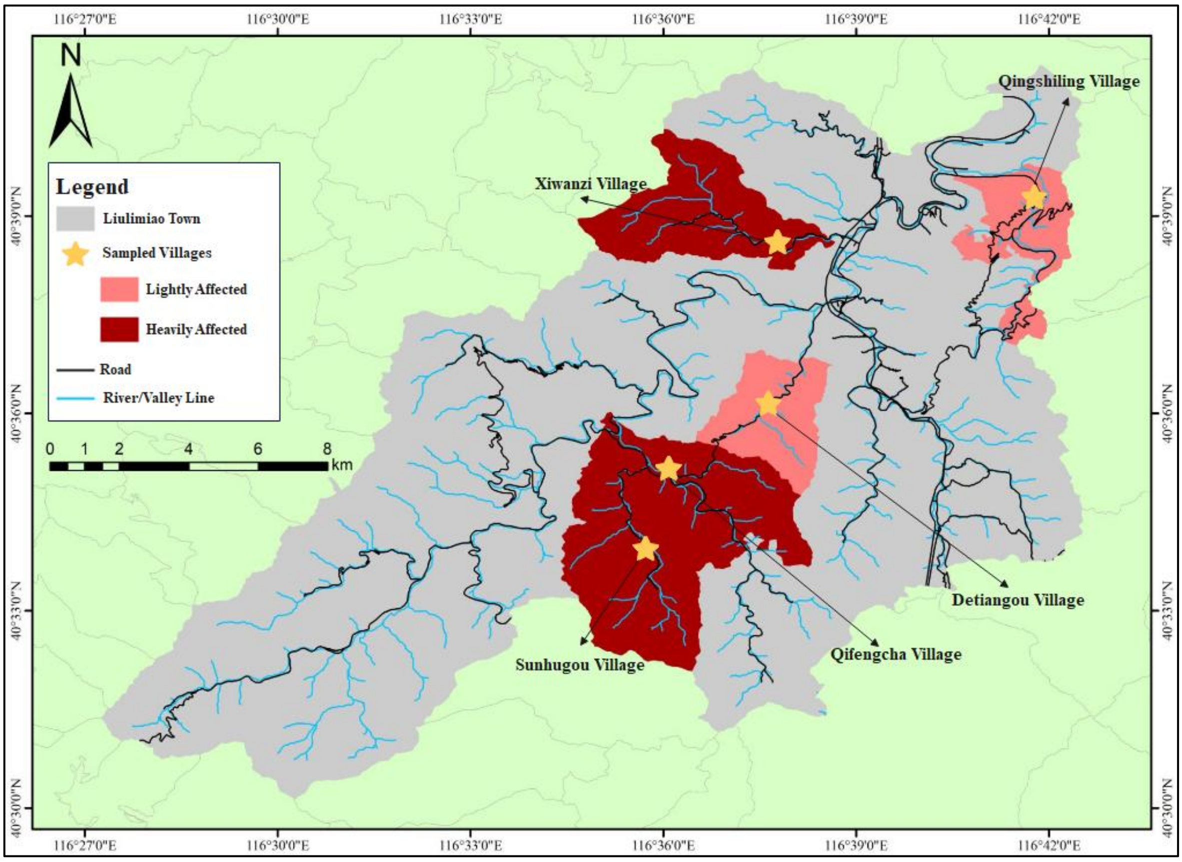

Figure 2. Study area

Measures and Operationalization

Measures are anchored in PADM and implemented as a friction framework aligned with the outcome: time lost from first awareness of a credible cue to evacuation initiation. Indicators capture two proximate delay mechanisms. Information friction reflects unclear or unactionable messages, weakened reliance on authoritative sources amid conflicting signals, and the verification burden of cross-channel checking (National Oceanic and Atmospheric Administration, 2016; de Moraes, 2023). Action-side readiness/constraints represent the feasibility of mobilization given resources and capabilities—preparedness, route/shelter knowledge, self-efficacy—and binding burdens such as caregiving and vulnerability-related barriers. We treat these measures as mechanism-proximal determinants rather than outcome proxies; threat perception enters as an upstream covariate that conditions urgency but may not yield rapid departure when downstream frictions bind. Mechanisms are operationalized as item composites or single-item indicators linked to the questionnaire (Table 1). Item wording and response options for the survey questions referenced in Table 1 are provided in the Supplementary Material. Items are reverse-coded so higher values indicate more information friction, while action-side measures are coded so higher values indicate greater readiness. Indicators are standardized prior to aggregation. Experimental evidence is consistent with this framing: coupling risk information with concrete protective-action guidance increases stated preparedness and intended protective actions in vulnerable communities (Wong-Parodi et al., 2018).

The outcome is warning-to-departure delay in six ordered intervals (0–15, 15–30, 30–60, 60–90, 90–120, >120 minutes), with finer early bins to capture front-loaded departures and wider late bins to limit burden and sparse cells. We convert the interval outcome to discrete-time event-history data by expanding each household into person–period records and coding a binary departure indicator by interval. The >120 category is treated as right-censoring at 120 minutes in the primary analysis and as a final open interval in sensitivity analyses; we also code ">120" as non-departure within the observation window as a robustness check. Interval-coded pre-evacuation time is common in evacuation simulation/estimation, supporting the practicality of discretized delay measures (Chen et al., 2019).

Indicators are curated using pre-specified rules, including collapsing sparse categories, screening near-zero variance, reducing redundancy among correlated items, and harmonizing directionality. Table 1 summarizes constructs and operationalization. Because measures are post-event self-reports, recall and common-method bias are possible; we assess robustness in sensitivity checks.

Table 1. PADM-informed mapping used to operationalize a friction-based framework

Study component	PADM-informed anchor	Survey block	Analytic variable family
Outcome: initiation delay	Behavioral response	Q35	Interval-coded departure delay → discrete-time event
Upstream context	Threat perceptions	Q15–Q17	Threat appraisal indicators
Mechanism A: information frictions	Predecision processing (comprehension/interpretation)	Q18–Q21	Message actionability/clarity friction; Verification burden
Mechanism B: action frictions	Situational facilitators/impediments	Q29–Q32	Action-friction / mobilization-readiness indicators

Note: PADM provides the conceptual structure for organizing mechanism domains. The empirical specification focuses on two timing-relevant mechanisms—information frictions and action frictions—to explain variation in departure timing. Multi-item constructs are averaged (reverse-coded as needed); continuous covariates are z-standardized using the full-sample mean and SD.

Analytical Strategy

Evacuation initiation is modeled as a discrete-time departure process (i.e., a discrete-time hazard model) rather than an ordinal outcome (Hasan et al., 2013). Household data are restructured into person–period format, with one record per interval until departure or the end of observation. The discrete-time hazard is defined as:

$$\mathrm{h}_{it}=P(y_{it}=1 \mid y_{i,t-1}=0) \tag{1}$$

where i indexes households and t indexes discrete time intervals; y_{it} indicates that household i first leaves home in interval t. We estimate a logistic model with interval fixed effects:

$$logit(\mathrm{h}_{it})=\alpha_t+\beta^{\top} X_i \tag{2}$$

where α_t denotes interval fixed effects capturing the time-varying baseline hazard, β is a vector of regression coefficients, and X_i is the vector of household-level covariates, including information frictions, action frictions, and upstream threat perceptions. Because intervals are unequally spaced, odds ratios from Eq. (2) reflect multiplicative changes in the interval-specific conditional odds of departure (given not yet departed), not constant per-minute effects. The term $\beta^{\top} X_i$ represents the linear predictor, i.e., the weighted sum of covariates and their associated coefficients. Standard errors are clustered at the household level to account for repeated person–period observations, consistent with standard practice in discrete-time event-history models estimated via logistic regression. Discrete-time hazard/survival formulations are also increasingly used in evacuation demand modeling and forecasting, including behaviorally integrated approaches that update predictions as new observations arrive (Guan & Chen, 2021).

The estimated interval hazards are used to construct cumulative departure curves and tail-delay risk measures. For each household (or for standardized group or scenario profiles), survival through interval t is computed as $S_i(t) = \prod_{j=1}^{t} (1 - h_{ij})$, and cumulative departure is defined as $F_i(t) = 1 - S_i(t)$. Delay risks are summarized using interpretable functionals of $F_i(t)$, including the tail probability P(delay>120) and cumulative departure by interval. The open-ended ">120 minutes" category is treated as right-censoring at 120 minutes in the primary analysis.

Behavioral heterogeneity is examined using latent class analysis (LCA) based on curated indicators from the information- and action-friction domains. Delay is excluded from the LCA measurement model and treated as a distal outcome; to account for classification uncertainty, we link delay to class membership using the bias-adjusted three-step Bolck–Croon–Hagenaars (BCH) procedure (Asparouhov & Muthén, 2014; Vermunt, 2010). Class enumeration is guided by Bayesian information criterion (BIC), entropy, and substantive interpretability.

Policy scenarios are examined using model-based standardization (g-computation) to generate counterfactual delay distributions by applying within-support shifts to mechanism indicators (e.g., -1 SD or the 75th percentile) while holding other covariates at observed values (Hernán & Robins, 2020). We report scenario-specific cumulative departure and tail-risk functions for the full sample and prespecified subgroups; 95% CIs are obtained via household bootstrap (B=4,000).

RESULTS

Descriptive Statistics

The analytic sample comprised 197 households. Demographic aging was pronounced: 78.2% of households reported at least one resident aged 60 or older (mean number of elderly members = 1.35). Vulnerability-related constraints were prevalent: 68.5% reported at least one vulnerable member, including 32.0% with mobility impairment and 45.7% requiring regular medication or medical care for chronic conditions. Private vehicle ownership was limited (29.9%), suggesting constraints on independent mobility during rapid-onset floods. Such vulnerability profiles have been linked to disproportionate post-disaster health and support needs among older adults, including evacuation- and information-related needs in urban disaster response (Ahmadi et al., 2018; Phraknoi et al., 2023).

The delay distribution shows a short-delay concentration alongside a substantial long tail: 42.6% of households initiated evacuation within 0–15 minutes, whereas 20.8% reported delays >120 minutes. Descriptive statistics are reported in Table 2 to make sample composition transparent along key vulnerability dimensions relevant to the PADM framing, including age structure, mobility constraints, health-related care needs, and mobility resources.

Table 2. Sample characteristics and delay distribution (N=197)

a. Key household characteristics

Characteristic	Value
Households (N)	197
Households with ≥1 elderly member (≥60)	154 (78.2%)
Mean number of elderly members	1.35
Households with any vulnerable member	135 (68.5%)
Mobility-impaired member present	63 (32.0%)
Chronic disease member present	90 (45.7%)
Private car ownership	59 (29.9%)

b. Warning-to-departure delay

Delay interval	Households	Percent
0–15 min	84	42.6%
15–30 min	36	18.3%
30–60 min	20	10.2%
60–90 min	10	5.1%
90–120 min	6	3.0%
>120 min	41	20.8%

Figure 3 plots the baseline cumulative departure function F(t) (right-censored at 120 minutes). Departures reached 60.9% within 30 minutes and 79.2% by 120 minutes, leaving a sizable tail of non-departure within the observation window.

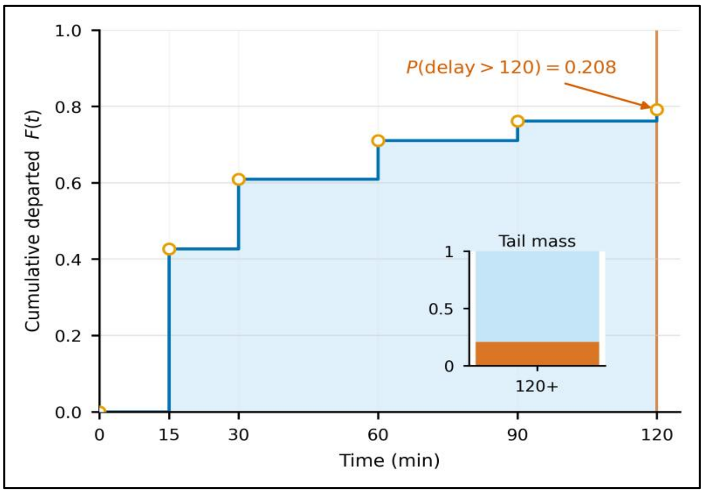

Figure 3. Baseline cumulative departure curve F(t)

Discrete-Time Hazard Models

We estimate nested discrete-time hazard models with interval fixed effects and household-clustered standard errors, treating ">120 minutes" as right-censoring at 120 minutes (Table 3). Coefficients are interpreted as changes in the interval-specific conditional odds of departure (given not yet departed). Model 1 includes threat perception; Model 2 adds information-side frictions; Model 3 adds action-side readiness; and Model 4 further controls for household mobility and health constraints.

Table 3. Discrete-Time Hazard Models of Evacuation Initiation

Variable	M1 Threat	M2 +Info	M3 +Action	M4 Full
Threat perception (z)	3.46*** [2.67, 4.47]	2.88*** [2.18, 3.79]	3.12*** [2.36, 4.13]	2.71*** [1.99, 3.69]
Info clarity/credibility/actionability friction (z, higher=worse)		0.47*** [0.36, 0.62]		0.50*** [0.37, 0.67]
Verification burden (z, higher=more)		0.89 [0.70, 1.14]		0.84 [0.65, 1.09]
Evacuation preparedness (z, higher=more prepared)			2.47*** [1.89, 3.23]	2.36*** [1.75, 3.18]
Self-efficacy (z)			0.86 [0.69, 1.07]	0.83 [0.64, 1.08]

Note: Odds ratios (ORs) are reported with 95% confidence intervals (CIs) in brackets. Standard errors are clustered at the household level. Model 4 (M4) additionally adjusts for mobility/resources and health-related constraints. * $p < .05$, ** $p < .01$, *** $p < .001$. OR = odds ratio; CI = confidence interval; SD = standard deviation.

Across specifications, threat perception remains a robust accelerator of departure timing. In the fully adjusted model (M4), a one-standard deviation increase in threat perception is associated with higher conditional odds of departure (OR=2.71, 95% CI [1.99, 3.69], p<0.001), consistent with faster movement once urgency is perceived.

Two proximate bottlenecks emerge. Higher clarity, credibility, and actionability friction is associated with lower conditional odds of departure (OR=0.50, 95% CI [0.37, 0.67], p<0.001), net of threat perception and constraints, indicating delayed initiation when warning information is harder to interpret and act on. Mobilization readiness is also strongly associated with earlier departure: greater preparedness is linked to higher conditional odds of departure (OR=2.36, 95% CI [1.75, 3.18], p<0.001).

Other covariates (verification burden, channel type, self-efficacy, car ownership, and health or mobility constraints) do not exhibit stable independent associations in the full model. Results are substantively unchanged in sensitivity analyses that treat ">120 minutes" as a final open interval or alternatively as non-departure within the observation window, and key coefficient directions remain the same.

Latent Class Analysis

To identify heterogeneity in friction mechanisms, we conducted a latent class analysis (LCA) using a parsimonious set of indicators spanning information frictions and mobilization readiness. Information-friction indicators capture message clarity and actionability, source credibility orientation under conflicting messages, and verification burden from cross-channel checking; mobilization-readiness indicators capture preparedness and evacuation planning, route and shelter knowledge, and evacuation self-efficacy. The delay outcome is excluded from the LCA measurement model and treated strictly as a distal outcome.

Models were estimated by maximum likelihood via the expectation–maximization (EM) with multiple random starts to reduce the risk of local maxima. We compared two- to five-class solutions and selected the final model based on BIC, entropy, class size, and substantive interpretability, while avoiding degenerate classes with implausibly small or unstable profiles (Table 4). Although the two-class model minimized BIC, it produced an overly coarse partition. The three-class solution improved fit and classification quality while maintaining balanced class proportions, whereas adding additional classes yielded only modest gains with increased complexity and reduced stability; we therefore retained the three-class solution as the most robust and policy-relevant specification.

Table 4. Latent Class Model Fit Statistics

Classes	LogLik	Params	AIC	BIC	Entropy	Min class share
2	-1973.626	57	4061.251	4248.394	0.815	0.459
3	-1913.476	86	3998.953	4281.308	0.844	0.202
4	-1869.525	115	3969.049	4346.618	0.897	0.206
5	-1839.110	144	3966.220	4439.002	0.923	0.113

Note: Models were estimated using maximum-likelihood EM (ML-EM) with multiple random starts. Entropy summarizes classification certainty, and class selection balanced statistical fit and interpretability. Delay was modeled as a distal outcome linked to class membership via the three-step BCH procedure.

The three latent classes correspond to distinct friction profiles. Because class numbering is arbitrary, we use substantive labels consistently across analyses. Class 1 represents a vulnerable, low-readiness profile, characterized by elevated vulnerability and limited mobilization readiness (lower preparedness, route/shelter knowledge, and self-efficacy). Class 2 represents a high-readiness, resourceful profile, characterized by lower information friction and higher mobilization readiness, consistent with greater capacity to act promptly. Class 3 represents a vulnerable-but-constrained profile with the highest prevalence of binding constraints, including vulnerable household members and mobility or health burdens, consistent with limited feasible mobility despite warning awareness.

To relate class membership to warning-to-departure delay while accounting for classification uncertainty, we use the bias-adjusted three-step BCH procedure and compare class-specific cumulative departure functions and tail-delay risks using BCH weights (Figure 4). We treat the resulting profiles as empirically grounded typologies in this study setting, with stability and transferability to be assessed through replication across additional events and communities.

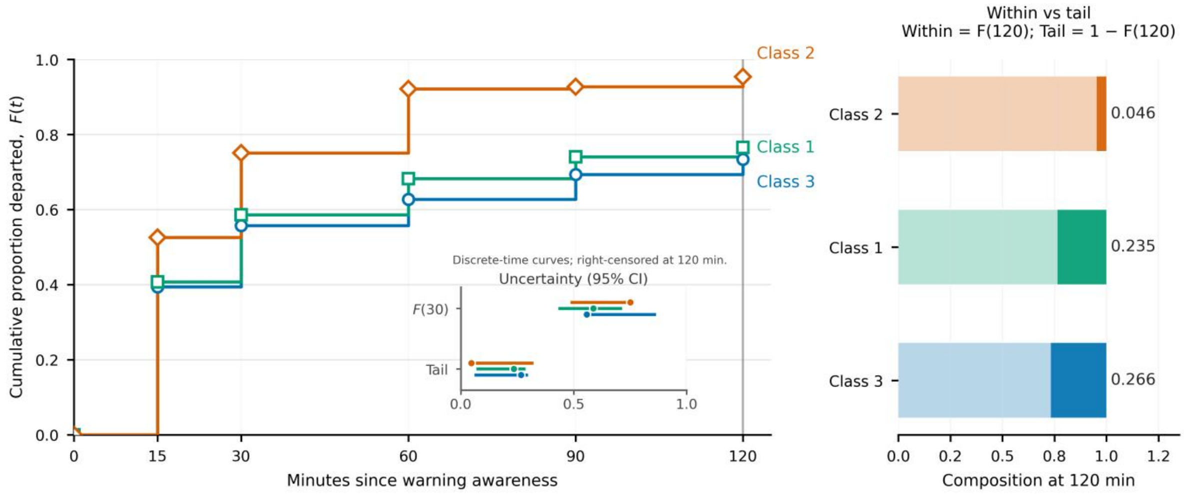

Figure 4. BCH-adjusted empirical class-specific cumulative departure functions and tail risk (Tail = 1 − F(120); right-censored at 120 min)

Scenario-Based Standardization

Building on the fitted model, we applied model-based standardization (g-computation) to quantify how the predicted delay distribution would change under feasible outreach scenarios. These projections are conditional on the fitted model and observed covariates and are used to compare leverage across mechanisms and household profiles. Relative to the model-implied baseline (F(30)=0.600; P(delay>120)=0.193; empirical baselines in Figure 4 and model-implied baselines in Figure 5), the information-focused scenario (−1 SD in the clarity/credibility/actionability friction composite and −1 SD in verification burden) increases F(30) to 0.740 and reduces tail risk to 0.060. The action-focused scenario (setting preparedness, route/shelter knowledge, and self-efficacy to their observed 75th percentiles) yields F(30)=0.807 and tail risk 0.042. The combined scenario produces the largest shift (F(30)=0.936) and near-eliminates tail risk (0.002). For each scenario, we derive household-level interval hazards, map them to cumulative departure and tail-risk functions, and average across households, including village indicators to absorb shared exposure and disruption.

To assess heterogeneity, we incorporated latent behavioral profiles using posterior class membership probabilities (rather than modal assignment), retaining classification uncertainty. We report class-standardized g-computation estimates with 95% CIs from a household bootstrap (B=4,000). Figure 5 shows that the vulnerable-but-constrained profile (Class 3) has the highest baseline tail risk (0.219) and the largest reduction under action support (0.038; Δ=−0.181). For the vulnerable/low-readiness profile (Class 1; baseline 0.162), information and action scenarios produce similar tail risks (0.045 and 0.044), while the combined scenario approaches zero (0.002). The high-readiness/resourceful profile (Class 2; baseline 0.167) shows smaller absolute gains but remains responsive, especially to action support (0.059); the combined scenario again approaches zero (≈0.002). Overall, integrated interventions compress the long tail of delayed departures, with profile-specific leverage points.

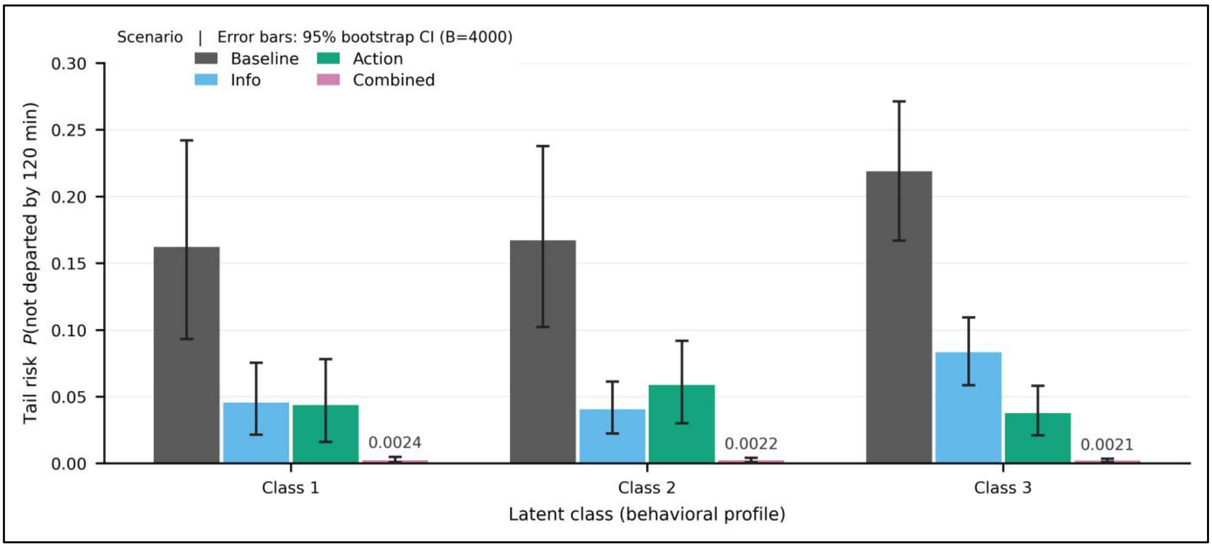

Figure 5. Model-implied baseline and scenario tail risks from standardization (g-computation) based on the fitted discrete-time hazard model. Scenarios use within-support shifts (e.g., −1 SD or 75th percentile) and are averaged over the observed covariate distribution

DISCUSSION

This study shows that evacuation initiation in rapid-onset floods should be understood not only as whether households evacuate, but also as when they begin to move. In Liulimiao, many households departed quickly, but a smaller group delayed much longer. This matters because average response times can hide the households facing the greatest entrapment risk.

Theoretically, the findings suggest that threat perception alone is not enough to explain timely evacuation. Within PADM, perceived threat helps activate protective decision-making, but the transition from awareness to movement depends heavily on downstream frictions. In our results, poorer warning clarity, credibility, and actionability were associated with later evacuation, while stronger mobilization readiness was associated with earlier departure. This means that even when households recognize danger, they may still delay if warnings are difficult to interpret or trust, or if they are not practically ready to move quickly. The study therefore extends PADM in a way that is useful for short-lead-time flooding by distinguishing between information frictions and action frictions as two related but separable sources of delay.

Methodologically, the study shows the value of treating evacuation delay as a distribution rather than relying only on a binary outcome or an average response time. Discrete-time event-history models make it possible to estimate when households are likely to depart, while cumulative departure and tail-risk functions translate model outputs into metrics that are easier to interpret for operational purposes. The latent class analysis further shows that delay is shaped by combinations of warning quality, readiness, and constraints rather than by isolated variables alone.

Practically, the results point to two complementary levers for reducing dangerously late evacuation: more actionable communication and stronger mobilization readiness. Scenario projections suggest that joint improvements compress the delay distribution more than either change alone. In rapid-onset flood settings, better communication alone may not be enough if households are not ready to act, and readiness alone may not be enough if warning information is confusing or unconvincing. For practice, this implies that warning design and local outreach should be treated as linked rather than separate tasks.

More broadly, the results suggest that evacuation performance in rapid-onset hazards should not be judged only by average improvement. A more useful benchmark may be whether interventions compress the long tail of dangerously late departures. In this sense, reducing tail-delay risk may be more operationally meaningful than improving central tendency alone, especially in settings where road access, communications, and household mobility are stressed at the same time.

LIMITATIONS AND FUTURE WORK

Several limitations qualify interpretation. First, key measures are post-event self-reports recorded in intervals, which may introduce recall and common-method bias as well as potential simultaneity. Second, scenario contrasts should be read as model-implied projections conditional on observed covariates and may still be sensitive to residual confounding by hazard severity and disruption. Third, generalizability is limited by the single-town, single-event design and purposive village selection. The study therefore provides mechanism-focused rather than statistically representative evidence: the warning and mobilization frictions identified here may also arise in other mountainous flood settings with short lead times and infrastructure disruption, although the extent and distribution of delay will likely vary with local demographic structure, mobility resources, communication resilience, and local evacuation policy arrangements.

Future work will focus on three extensions: (i) robustness and measurement validation, including sensitivity checks for delay coding and censoring choices; (ii) profile stability and within-context validation, including alternative indicator sets and class solutions; and (iii) improved intervention mapping and external testing by adding more direct measures of warning exposure and mobilization support, alongside replication across additional events and communities. A longer-term extension is to develop an agent-based model as a complementary tool to explore coordination spillovers and infrastructure-disruption dynamics under rapid-onset disruption.

CONCLUSION

This study examined evacuation initiation following warning awareness in a rapid-onset flood context using household survey data from Liulimiao Town. Modeling delay as a discrete-time departure process revealed front-loaded departures alongside substantial tail-delay exposure.

By aligning PADM with a timing-based outcome, the study identifies information frictions and action frictions as intervention-facing mechanisms associated with earlier departure beyond threat activation. Latent class analysis captures heterogeneous friction profiles with distinct delay trajectories, and standardization suggests that improving warning actionability and mobilization readiness, especially together, is associated with substantial reductions in tail-delay risk. Overall, the findings underscore the value of treating evacuation delay as a distributional problem, with performance benchmarks that prioritize compression of dangerously late departures. Future work will further assess robustness and transportability across additional events and communities.

ACKNOWLEDGMENTS

This work was supported by the Innovative Research Group Project of the National Natural Science Foundation of China (No. 72521001, Public Safety Emergency Management); the Chinese Academy of Engineering institute–local cooperation project (Local, No. 2025-AHYJY-06); the National Key R&D Program of China (No. 2024YFC3017000); and the National Natural Science Foundation of China (Grant No. 72334003).

REFERENCES

Acosta-Coll, M., Ballester-Merelo, F., Martinez-Peiró, M., & De la Hoz-Franco, E. (2018). Real-time early warning system design for pluvial flash floods—A review. *Sensors*, *18*(7), 2255. https://doi.org/10.3390/s18072255

Ahmadi, S., Khankeh, H. R., Mohammad Khan Kermanshahi, S., & Farrokhi, M. (2018). Health needs of older adults after natural disasters: A systematic literature review. *Trauma Monthly*, *23*(4), e60622. https://doi.org/10.5812/traumamon.60622

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329–341. https://doi.org/10.1080/10705511.2014.915181

Chen, J., Yu, J., Wen, J., Zhang, C., Yin, Z., Wu, J., & Yao, S. (2019). Pre-evacuation time estimation based emergency evacuation simulation in urban residential communities. *International Journal of Environmental Research and Public Health*, *16*(23), 4599. https://doi.org/10.3390/ijerph16234599

China Daily. (2025, July 29). 30 dead amid heavy rainfall in Beijing, rescue efforts ongoing. *China Daily*. https://www.chinadaily.com.cn/a/202507/29/WS68880594a310c26fd717c279.html

de Moraes, O. L. L. (2023). An impact-based forecast system developed for hydrometeorological hazards. *International Journal of Disaster Risk Reduction*, *93*, 103803. https://doi.org/10.1016/j.ijdrr.2023.103803

Guan, X., & Chen, C. (2021). A behaviorally integrated individual-level state-transition model that can predict rapid changes in evacuation demand days earlier. *Transportation Research Part E: Logistics and Transportation Review*, *152*, 102381. https://doi.org/10.1016/j.tre.2021.102381

Hasan, S., Mesa-Arango, R., & Ukkusuri, S. V. (2013). A random-parameter hazard-based model to understand household evacuation timing behavior. *Transportation Research Part C: Emerging Technologies*, *27*, 108–116. https://doi.org/10.1016/j.trc.2011.06.005

Hernán, M. A., & Robins, J. M. (2020). *Causal inference: What if*. Chapman & Hall/CRC. Retrieved February 25, 2026, from https://miguelhernan.org/whatifbook

Huang, X., Fang, Y., He, X., Yang, L. E., Xu, Y., Liu, J., Zhu, A., & Guo, Y. (2025). Diversity and differences of flood evacuation strategies among rural households in the Hengduan Mountains, Southwest China. *Geomatics, Natural Hazards and Risk*, *17*(1). https://doi.org/10.1080/19475705.2025.2608248

Intergovernmental Panel on Climate Change. (2021). *Climate change 2021: The physical science basis. Summary for policymakers*. https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM.pdf

Koll, C., Ahrens, C., Bearden, M., Coulliette, C., Jones, K., McNease, K., & Thompson, N. (2023). Emergency warning dissemination in a multiplex social network. *Journal of Artificial Societies and Social Simulation*, *26*(1), 7. https://doi.org/10.18564/jasss.4946

Kuller, M., Schoenholzer, K., & Lienert, J. (2021). Creating effective flood warnings: A framework from a critical review. *Journal of Hydrology*, *602*, 126708. https://doi.org/10.1016/j.jhydrol.2021.126708

Li, C., Jiang, C., Lam, M. Y., & Ahmadian, R. (2025). An overview of flood evacuation planning: Models, methods, and future directions. *Journal of Hydrology*, *656*, 133026. https://doi.org/10.1016/j.jhydrol.2025.133026

Li, R., & Gilbert, C. (2024). Impact of direct experience on disaster preparedness and evacuation: A Protective Action Decision Model analysis in low-income New York City communities. *International Journal of Disaster Risk Reduction*, *114*, 105009. https://doi.org/10.1016/j.ijdrr.2024.105009

Lindell, M. K., & Perry, R. W. (2012). The protective action decision model: Theoretical modifications and additional evidence. *Risk Analysis*, *32*(4), 616–632. https://doi.org/10.1111/j.1539-6924.2011.01647.x

Morss, R. E., Mulder, K. J., Lazo, J. K., & Demuth, J. L. (2016). How do people perceive, understand, and anticipate responding to flash flood risks and warnings? Results from a public survey in Boulder, Colorado, USA. *Journal of Hydrology*, *541*, 649–664. https://doi.org/10.1016/j.jhydrol.2015.11.047

National Oceanic and Atmospheric Administration, Office for Coastal Management. (2016). *Seven best practices for risk communication (quick reference)*. https://coast.noaa.gov/data/digitalcoast/pdf/risk-communication-best-practices.pdf

National Weather Service. (n.d.). *Flood warning vs. watch*. Retrieved February 25, 2026, from https://www.weather.gov/safety/flood-watch-warning

Phraknoi, N., Sutanto, J., Hu, Y., Goh, Y. S., & Lee, C. E. C. (2023). Older people's needs in urban disaster response: A systematic literature review. *International Journal of Disaster Risk Reduction*, *96*, 103809. https://doi.org/10.1016/j.ijdrr.2023.103809

Proulx, G., & Fahy, R. F. (1997). The time delay to start evacuation: Review of five case studies. In Y. Hasemi (Ed.), *Proceedings of the Fifth International Symposium on Fire Safety Science* (pp. 783–794). International Association for Fire Safety Science. https://doi.org/10.3801/IAFSS.FSS.5-783

Ren, Z., Sang, Y., Yang, M., Wang, Y., & Shang, L. (2023). Progress of research on the methods for the early warning of mountain flash flood disasters. *Progress in Geography*, *42*(1), 185–196. https://doi.org/10.18306/dlkxjz.2023.01.015

Reuters. (2025, July 28). Thirty dead in Beijing following heavy rain, Xinhua reports. *Reuters*. https://www.reuters.com/sustainability/climate-energy/thirty-dead-beijing-following-heavy-rain-xinhua-reports-2025-07-28/

United Nations Office for Disaster Risk Reduction. (n.d.). *Early warning system*. Retrieved February 25, 2026, from https://www.undrr.org/terminology/early-warning-system

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*(4), 450–469. https://doi.org/10.1093/pan/mpq025

Wong-Parodi, G., Fischhoff, B., & Strauss, B. (2018). Effect of risk and protective decision aids on flood preparation in vulnerable communities. *Weather, Climate, and Society*, *10*(3), 401–417. https://doi.org/10.1175/WCAS-D-17-0069.1

World Meteorological Organization. (n.d.). *Early warning system*. Retrieved February 25, 2026, from https://wmo.int/topics/early-warning-system

SUPPLEMENTARY MATERIAL

This Supplementary Material reports the English translations of the survey items used to operationalize the constructs in Table 1 and construct the analytic variables, including the delay outcome. Item wording and response options are reported verbatim to support transparency and replicability; unless otherwise noted, items use a five-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree). Due to space constraints, only analysis-critical items are included.

Q35 (Outcome) — Warning-to-departure delay (single choice)

"From the time you received a flood warning (or an equivalent urgent notice) to the time you began to evacuate, approximately how long was the delay?"
Options: 0–15 minutes; 15–30 minutes; 30–60 minutes; 60–90 minutes; 90–120 minutes; >120 minutes.

Q15–Q17 (Threat perceptions / upstream context)

Q15 (Likert): "I think extreme rainfall is likely to cause severe urban flooding."

Q16 (single choice): "Regarding flood risk, which view do you agree with most?"
A. "Modern flood-control systems are reliable; flood risk is controllable."
B. "Flood risk exists, but warnings and preparedness can effectively cope with it."
C. "Flood risk is hard to predict and highly damaging; it cannot be fully avoided."

Q17 (ranking; 1 = most concerned, 5 = least concerned): "If severe flooding occurs, which impacts are you most concerned about?"
Rank: (a) inconvenience (transport disruption / daily activities affected);
(b) property loss;
(c) life safety (injury/death and health risks);
(d) environmental pollution;
(e) disruption of social services (healthcare, water, electricity, etc.).

Q18–Q21 (Information frictions / warning comprehension, actionability, verification, conflict)

Q18 (Likert): "When I receive clear sheltering/evacuation guidance from authoritative sources or the community, I will act according to the guidance immediately."

Q19 (Likert): "Overall, heavy-rainfall/flood warning information released via authoritative sources is reliable, clearly expressed, and provides actionable guidance."

Q20 (Likert): "Rather than acting immediately upon receiving an official warning or evacuation notice, I cross-check key information through at least two different channels (e.g., whether it affects my area, the risk level, and specific recommended actions)."

Q21 (ranking; 1 = highest priority, 5 = lowest): "When different channels provide inconsistent heavy rainfall/flooding, what is your usual priority order for reference?"
Rank: (a) authoritative issuing channels (e.g., emergency management SMS, meteorological warnings);
(b) community/neighborhood notifications;
(c) traditional media (TV/radio);
(d) internet platforms (news/social media/apps/websites);
(e) family/friends.

Q29–Q32 (Action-side readiness / mobilization capability and preparedness)

Q29 (Likert): "I know which safe shelters I should go to if evacuation is needed, and how to get there."

Q30 (Likert): "When facing flood hazards, I think my response and escape capability is sufficient."

Q31 (Likert): "I believe that evacuating in a timely manner can effectively protect life safety and help avoid injuries and deaths."

Q32 (single choice): "Have you/your household developed an emergency flood-evacuation plan or made related preparations?"

A. “Yes—detailed household evacuation plan (clear routes, contact methods, and supplies, etc.).”
B. “Some preparation, but no written/specific plan (have thought about what to do, but no formal plan).”
C. “No plan or preparation.”

Key controls / constraints used descriptively or as covariates

Q13 (single choice): “Does your household own a private car?” Options: Yes (skip to Q14); No.

Q14 (multiple choice): “Does your household include any of the following vulnerable members? (Select all that apply.)”
Options: pregnant/postpartum; infants/toddlers; persons with disabilities or mobility impairment; chronic disease requiring regular care/medication; home medical dependence (e.g., dialysis, oxygen); none; other.