- # **Referee #2**

Overall comments

Thank you for the opportunity to review this manuscript. The study aims to estimate the impact of two flood events on population movements in Japan, with the goal of assessing the effectiveness of flood mitigation strategies. Overall, the manuscript is well written and addresses an interesting and relevant research question. The empirical approach is generally sound, but I do have some concerns and suggestions for improvement. Below I provide several comments that could improve the paper. These points are presented in no particular order.

Response:

We sincerely thank the reviewer for carefully reviewing our manuscript and for the constructive comments. We are grateful for the positive evaluation of the writing, research question, and empirical approach. In response to the reviewer's comments, we revised the manuscript to clarify the data used to estimate the empirical relationships between flood magnitude and FIPMs, expand the explanation of the DiD method and its assumptions, improve the notation and definitions of variables in the equations, add supplementary figures showing the spatial coverage of the flood data used in the analyses, and clarify the temporal interpretation of the estimated FIPMs. Detailed responses to each specific comment are provided below.

Comment 1

My main concern relates to the timing of the analysis. The flood events occurred in 2019 and 2020, while population data are observed from 2005 to 2020 in five-year intervals. As a result, the study can only capture very short-term population movements in response to the floods. Since migration and population adjustments often occur gradually, it would be more informative to examine longer-term effects, for example by including population data for the period after 2020 (e.g., 2020–2025) if such data become available. As it stands, the analysis likely captures only immediate or short-run effects. This limitation also raises some concerns about the usefulness of conducting projections over several decades based on these short-term responses.

Response:

We thank the reviewer for this important comment. We agree that the temporal coverage of the available demographic data limits the interpretation of population adjustments after flood events. We therefore added this point to the limitations. In the added text, we clarify that the FIPMs quantified in this study should be interpreted as population changes and registered migration associated with flood exposure that are observable in the available demographic data, rather than as a complete representation of short-term evacuations or longer-term migration processes. We also clarify that the long-term projections in this study should not be interpreted as predictions of event-specific migration trajectories, but rather as scenario-based assessments of how empirical relationships between flood magnitude and FIPMs could affect future population distributions, land use, and fluvial flood damage estimates. Finally, we note that explicitly modeling these longer-term adjustment processes is outside the scope of the present framework and should be addressed in future research when post-2020 census and migration data become available.

Added text
Sixth, the temporal coverage of the demographic data limits the interpretation of population adjustments after flood events. The FIPMs quantified in this study should be interpreted as population changes and registered migration associated with flood exposure that are observable in the available demographic data. These datasets do not capture short-term evacuations or movements that occurred without changes in registered residence. Moreover, longer-term migration and population adjustment after flood events may be shaped by recovery processes, land markets, public investment, housing reconstruction, and changes in regional economies. Therefore, the long-term projections in this study should not be interpreted as predictions of event-specific migration trajectories, but rather as scenario-based assessments of how empirical relationships between flood magnitude and FIPMs could affect future population distributions, land use, and fluvial flood damage estimates. Explicitly modeling these longer-term adjustment processes is outside the scope of the present framework and should be addressed in future research when post-2020 census and migration data become available.

Comment 2

The Introduction is well written and clearly outlines the research problem. However, the empirical strategy, namely the use of a difference-in-differences (DiD) approach, is not sufficiently introduced or motivated. It would be helpful if the authors explained why they chose a DiD framework and whether alternative empirical approaches could have been used. Placing the chosen method more explicitly within the existing literature would improve the paper.

Response:

We thank the reviewer for this helpful comment. We agree that the motivation for using the DiD approach was not sufficiently explained in the original Introduction. We therefore added a paragraph before stating the aim of this study to place our empirical strategy more clearly within the existing literature. In this paragraph, we discuss previous empirical approaches, including propensity score matching with regression-based statistical modeling and panel gravity models, and clarify that DiD is well suited for this study because incorporating FIPMs into future population projections requires estimating population responses relative to a counterfactual scenario in which flooding did not occur.

Added text
Previous studies have used empirical approaches to analyze relationships between flood exposure or natural-hazard-related damage and population dynamics. For example, Shu et al. (2023) used propensity score matching followed by regression-based statistical modeling to estimate the relationship between flood exposure and population change, while Ton et al. (2025) applied panel gravity models to county-to-county migration flows to analyze the relationship between economic damage from natural hazards and internal migration. These approaches provide important insights into population responses to flood exposure and natural-hazard-related damage. Incorporating FIPMs into future population projections requires estimating population responses relative to a counterfactual scenario in which flooding did not occur. For this purpose, the difference-in-differences (DiD) method, a statistical causal inference technique developed in the fields of econometrics and sociology, is well suited when its identification assumptions are plausible. It quantifies FIPMs by estimating population changes and migration responses attributable to flooding through comparisons between affected and unaffected areas before and after flood events.

Added reference
Ton, M. J., de Moel, H., de Bruijn, J. A., Reimann, L., Botzen, W. J. W., and Aerts, J. C. J. H.: Economic damage from natural hazards and internal migration in the United States, Nat. Hazards, 121, 4985–5005, https://doi.org/10.1007/s11069-024-06987-2, 2025.

Comment 3

It would be useful to clarify earlier in the Introduction that the analysis focuses on only two flood events. When reading the Introduction, I initially expected a broader analysis of the effects of floods on population movements across Japan, whereas the study is in fact closer to a case study. In addition, a map showing the locations of the flood events would be helpful. Then we could also see whether the affected areas overlap because I was wondering this when reading the manuscript.

Response:

We thank the reviewer for this helpful comment. We agree that the data used to estimate the empirical relationships between flood magnitude and FIPMs should be clarified earlier in the Introduction, while noting that the subsequent future population projections and fluvial flood damage cost estimates are conducted nationwide. In the original manuscript, the Introduction and methodological framework sections did not sufficiently explain that the grid-cell-level estimation of FIPMs was based on two observed flood events, whereas the municipality-level estimation used data from flood-affected municipalities during 2015–2020. To address this, we revised the Introduction and methodological framework sections to clarify the data used to estimate the empirical relationships between flood magnitude and FIPMs. We also clarified that these empirical relationships are applied to future population projections using flood magnitude indicators derived from future fluvial flood inundation analyses, rather than by directly extrapolating the observed events themselves. In addition, following the reviewer's suggestion, we added supplementary maps showing (i) the areas covered by the flood inundation maps used in the 500 m grid-cell-level analysis and (ii) the treated municipalities used in the municipality-level DiD analysis. These are now presented as Figs. S1 and S2.

Original text	Revised text
Against this background, this study aims to evaluate the impacts of FIPMs on future flood damage costs in Japan. To achieve this, we propose a methodological framework for projecting future population and estimating flood damage costs that accounts for FIPMs. To quantify population movements between and within municipalities in response to varying flood magnitudes, this study employs the difference-in-differences (DiD) method – a statistical causal inference technique developed in the fields of econometrics and sociology. (p. 3, lines 66–70 in the original manuscript)	Against this background, this study aims to evaluate the impacts of FIPMs on future fluvial flood damage costs in Japan. To achieve this, we propose a methodological framework for projecting future population and estimating fluvial flood damage costs that accounts for FIPMs. Specifically, we use the DiD method to quantify FIPMs between and within municipalities in response to varying flood magnitudes and incorporate the estimated relationships into future population and land-use projections. These empirical relationships are estimated using grid-cell-level data from two flood events in 2019 and 2020 and municipality-

	level data from flood-affected municipalities during 2015–2020. They are then applied to future population projections using flood magnitude indicators derived from future fluvial flood inundation analyses, rather than by directly extrapolating the observed events themselves.
FIPMs were quantified at both the grid-cell and municipality levels utilizing available demographic data to capture population movements occurring within and between municipalities in response to varying flood magnitudes. Detailed methodologies for quantifying FIPMs are presented in Sect. 3, while methodologies for the subsequent processes of the framework are described in Sect. 4. (p. 4, lines 77–80 in the original manuscript)	FIPMs were quantified at both the grid-cell and municipality levels using available demographic data to capture population movements occurring within and between municipalities in response to varying flood magnitudes. Specifically, FIPMs were quantified using two types of observed datasets that combine flood-related and demographic information. Within-municipality population changes were estimated using flood inundation maps for areas affected by Typhoon Hagibis (2019) and the heavy rainfall event of July 2020, together with 500 m grid-cell population data. Inter-municipality population movements were estimated using municipality-level flood damage records and annual migration data. These data were used to estimate empirical relationships between flood magnitude and FIPMs, rather than to directly extrapolate the observed events themselves. The areas covered by the flood inundation maps used in the 500 m grid-cell-level analysis and the treated municipalities used in the municipality-level DiD analysis are shown in Figs. S1 and S2, respectively. Detailed methodologies for quantifying FIPMs are presented in Sects. 2.2 and 2.3, while methodologies for the subsequent processes of the framework are described in Sects. 2.4–2.6.

Added figures

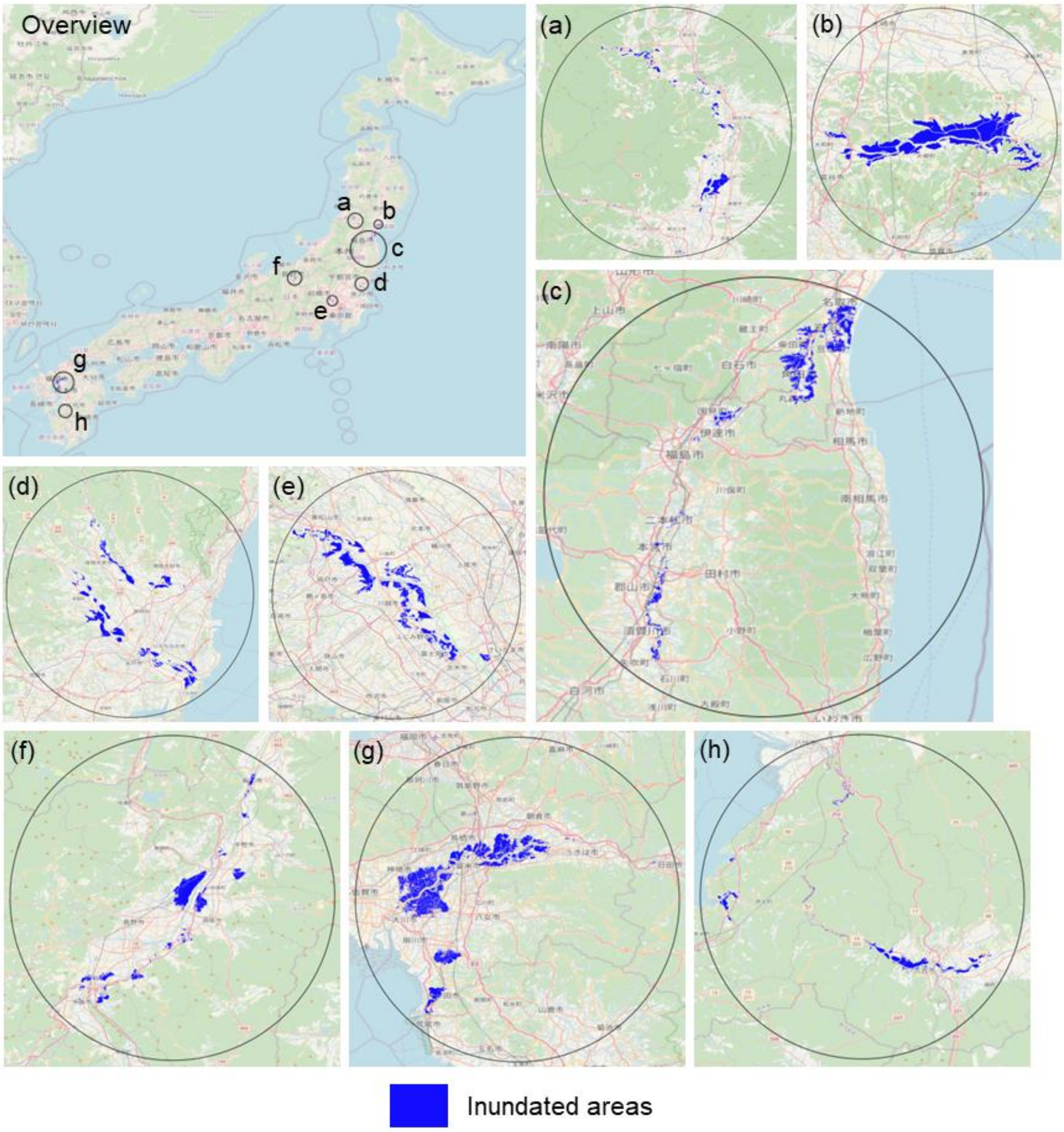

Figure S1: Inundated areas used in the 500 m grid-cell-level analysis. The blue areas indicate inundated areas identified from the flood inundation maps for Typhoon Hagibis (2019) and the heavy rainfall event of July 2020. The overview map shows the locations of the enlarged panels (a)–(h). Panels (b)–(f) correspond to Typhoon Hagibis (2019), whereas panels (a), (g), and (h) correspond to the heavy rainfall event of July 2020. Nearby mapped regions were grouped for display. Basemap: © OpenStreetMap contributors.

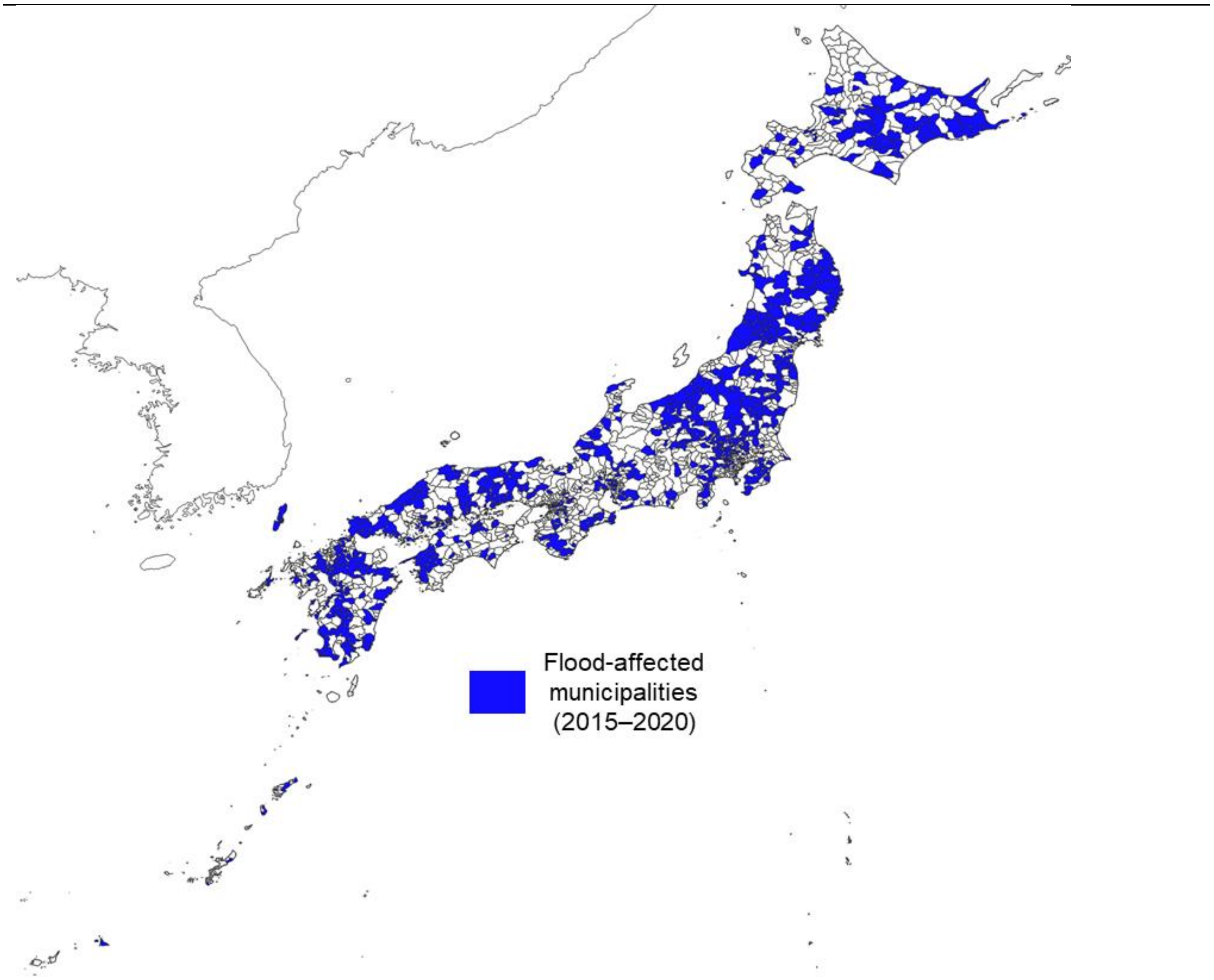

Figure S2: Municipalities affected by flood disasters during 2015–2020 and included as treated municipalities in the municipality-level DiD analysis.

Comment 4

A large part of the methodology is devoted to explaining the variables, coefficients and fixed effects, which makes the paper somewhat difficult to follow. I would encourage the authors to have another look at the equations and see where they can simplify the notation. For example, ¥delta_{bf}^{mun} R_{i', t', g'}^{bf} is not very readable. The equations might become clearer if the authors use simpler notation for coefficients, such as ¥beta_{0}, ¥beta_{1} and ¥beta_{2}, and call the variables BF, AF and CD, for example. Reducing the number of subscripts and superscripts would likely improve readability. Another suggestion would be to present the variables, coefficients, and fixed effects in a table or in a bullet-point list, so that the reader has a more clear overview.

Response:

We thank the reviewer for this helpful suggestion. We agree that the notation in the original equations was difficult to follow because many variables, coefficients, fixed effects, subscripts, and superscripts were defined directly in the text. We therefore revised the notation in the DiD regression equations throughout the main text and Supplement to improve readability and avoid overlapping symbols. Specifically, we removed superscripts where possible and simplified the notation for the treatment variables with reference to the reviewer's suggestion. At the same time, we retained the subscripts needed to identify the panel structure required for the DiD analysis, such as spatial unit, year, and year-specific dataset. To further improve readability, we moved the definitions of variables and parameters in Eqs. (1) and (2) to Tables 1 and 2, respectively.

Added tables

Table 1: Definitions of variables and parameters in Eq. (1).

Symbol	Definition
i	Grid-cell index, $i = \{1, 2, \dots, n\}$
n	Number of grid cells analyzed
t	Census year, $t = \{2005, 2010, 2015, 2020\}$
$P_{i,t}$	Total population (persons)
γ_i	Grid-cell fixed effects
τ_t	Year fixed effects
$\rho_i t$	Grid-cell-specific linear time trends
$\boldsymbol{Z}_{i,t-5}$	Covariate vector excluding time-invariant covariates
$\boldsymbol{\eta}$	Parameter vector for $\boldsymbol{Z}_{i,t-5}$
$H_{i,t}$	Maximum inundation depth (m)
ϕ	Parameter for $H_{i,t}$
$u_{i,t}$	Error term

Table 2: Definitions of variables and parameters in Eq. (2).

Symbol	Definition
j	Municipality index, $j = \{1, 2, \dots, n_g\}$
g	Year-specific dataset index, $g = \{1, 2, \dots, 6\}$
n_g	Number of municipalities analyzed within each year-specific dataset
y	Year, $y = \{2014, 2015, \dots, 2020\}$
$N_{j,y,g}$	Net migration rate (%)
$\mu_{j,g}$	Municipality fixed effects for each year-specific dataset
$\nu_{y,g}$	Year fixed effects for each year-specific dataset
$\kappa_{j,g}y$	Municipality-specific linear time trends within each year-specific dataset
$\boldsymbol{W}_{j,y-1,g}$	Covariate vector excluding time-invariant covariates
$\boldsymbol{\lambda}$	Parameter vector for $\boldsymbol{W}_{j,y-1,g}$
$BF_{j,y,g}$	Proportion of households affected below floor level by flooding (%)
$AF_{j,y,g}$	Proportion of households affected above floor level by flooding (%)
$CD_{j,y,g}$	Proportion of households that were completely destroyed by flooding (%)
β_1	Parameter for $BF_{j,y,g}$
β_2	Parameter for $AF_{j,y,g}$
β_3	Parameter for $CD_{j,y,g}$
$\upsilon_{j,y,g}$	Error term

Comment 5

In Equations (1) and (2), the authors include a time trend for each cross-sectional unit (grid cells in Equation 1 and municipalities in Equation 2). At the same time, the paper emphasizes the importance of satisfying the parallel trends assumption so that the estimated effect can be attributed to the flood rather than to other factors. The manuscript presents evidence suggesting that the parallel trends assumption holds. Given this, it is unclear why unit-specific time trends are still included in the regression. While the supplementary information cites one study that adopts a similar specification, this alone does not justify the modeling choice. I would like to see a clearer rationale for including these trends, or even better, a robustness check presenting results without the unit-specific time trends, especially since time fixed effects are already included.

Response:

We thank the reviewer for this important comment. We agree that the rationale for including unit-specific linear time trends was not sufficiently explained in the original Supplement. We revised the Supplement to clarify, using the grid-cell-level model as an example, that these trends were included to reduce the risk that heterogeneous pre-existing demographic trends are attributed to flooding. Specifically, we clarified that year fixed effects control for shocks common to all grid cells in each census year, whereas grid-cell-specific linear time trends control for gradual, pre-existing population trends that differ across grid cells, such as urbanization or depopulation. Although this explanation is provided for Eq. (1), the same rationale applies to the municipality-specific linear time trends in Eq. (2), which are included to account for gradual pre-existing trends that differ across municipalities. We also clarified that the indirect validation of the parallel trends assumption supports the absence of statistically significant pre-flood differences between the treatment and control groups. The inclusion of unit-specific linear trends further reduces the risk of attributing such pre-existing trends to flooding. In addition, following the reviewer's suggestion, we conducted sensitivity analyses without unit-specific linear time trends and added the results to Tables S12 and S13. We revised the main text to briefly report these additional analyses and to clarify that the estimates that were statistically significant in Tables S10 and S11 remained statistically significant and retained the same signs. Although we do not further discuss the changes in coefficient magnitudes in the manuscript to avoid over-interpreting the sensitivity analyses, the statistically significant estimates became larger in the negative direction at both the grid-cell and municipality levels when unit-specific linear time trends were omitted. This pattern is consistent with our concern that estimates without unit-specific linear time trends may be more susceptible to gradual pre-existing demographic trends, particularly in a country such as Japan where many areas face population decline. We therefore retain Eqs. (1) and (2), which include unit-specific linear time trends, as our main models.

Original text	Revised text
In the basic DiD model shown in Eq. (S1), the	In the basic DiD model shown in Eq. (S1), the

third term on the right-hand side of Eq. (1) was not included. However, if the time trends in the total population differ between the treatment and control groups, the basic DiD model fails to satisfy the parallel trends assumption. To address this issue, we followed Angrist and Pischke (2015) and added the third term to control for grid-cell-specific linear time trends and enhance the validity of the parallel trends assumption. (p. 3, lines 63–66 in the original Supplement)	third term on the right-hand side of Eq. (1) was not included. However, if the time trends in the total population differ between the treatment and control groups, estimates from the basic DiD model may be affected by heterogeneous pre-existing demographic trends. To reduce this risk, we followed Angrist and Pischke (2015) and included the third term to control for grid-cell-specific linear time trends. Year fixed effects control for shocks common to all grid cells in each census year, whereas grid-cell-specific linear time trends control for gradual, pre-existing population trends that differ across grid cells, such as urbanization or depopulation. The indirect validation of the parallel trends assumption in Sect. S6 supports the absence of statistically significant pre-flood differences between the treatment and control groups. Including grid-cell-specific linear trends further reduces the risk of attributing such pre-existing demographic trends to flooding.
Detailed results of parameter estimation for Eqs. (1) and (2) can be found in Tables S10 and S11. (p. 13, lines 326–327 in the original manuscript)	Detailed results of parameter estimation for Eqs. (1) and (2) can be found in Tables S10 and S11. As a sensitivity analysis, we also estimated Eqs. (1) and (2) without unit-specific linear time trends; the results are shown in Tables S12 and S13. The statistically significant estimates in Tables S10 and S11 remained statistically significant and retained the same signs in this sensitivity analysis.

Added tables

Table S12: Estimated parameters for Eq. (1) without grid-cell-specific linear time trends.

	Parameter (standard error)
Maximum inundation depth	−0.042*** (0.006)
Adjusted R^2	0.963
Within adjusted R^2	0.060
Sample size	19,120

Note: ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % levels, respectively. Robust standard errors clustered at the grid-cell level are reported in parentheses. Results are shown for the maximum inundation depth; other results are omitted.

Table S13: Estimated parameters for Eq. (2) without municipality-specific linear time trends within each year-specific dataset.

	Parameter (standard error)
Proportion of households affected below floor level by flooding	0.012 (0.034)
Proportion of households affected above floor level by flooding	−0.041 (0.037)
Proportion of households completely destroyed by flooding	−0.098*** (0.021)
Adjusted R^2	0.685
Within adjusted R^2	0.023
Sample size	5,512

Note: ***, **, and * indicate statistical significance at the 1 %, 5 %, and 10 % levels, respectively. Robust standard errors clustered at the municipality level for each year-specific dataset are reported in parentheses. Results are shown for the proportions of households affected by flooding; other results are omitted.

Comment 6

Equations (3), (4), and (5) are also somewhat difficult to follow. The notation b,c = {1,2,3,4,5} is not immediately clear. It might be simpler and more transparent to refer directly to the corresponding years instead of using index notation.

Response:

We thank the reviewer for this helpful suggestion. We agree that the notation was not sufficiently clear in the original manuscript. Regarding b and c, directly replacing these indices with calendar years would be difficult because the corresponding years depend on the projection year T. We therefore retained the index notation but revised the explanation to clarify its meaning. Specifically, we clarified that b and c represent annual-step indices within each five-year projection interval rather than fixed calendar years. We also added an example showing how these indices correspond to years within a projection interval. In addition, we summarized the variables and parameters in Eqs. (3) and (4) in Table 3 to improve readability. Because the variables in Eq. (5) are fewer and are explained immediately after the equation, we retained the original explanatory format for Eq. (5).

Added text
The indices b and c represent annual-step indices within each five-year projection interval rather than fixed calendar years. For example, when $T = 2020$, $b = 1$ represents the one-year population change from 2020 to 2021, whereas $b = 5$ represents the change from 2024 to 2025. The same interpretation applies to c.

Added table

Table 3: Definitions of variables and parameters in Eqs. (3) and (4).

Symbol	**Definition**
T	Projection year, $T = \{2015, 2020, \ldots, 2100\}$
b, c	Annual-step indices within each five-year projection interval, $b, c = \{1, 2, 3, 4, 5\}$
S	Sex
X	Lower bound of each five-year age group
$C_{T+b-1\to T+b,S,(X+b-1\text{ to }X+b+3)\to(X+b\text{ to }X+b+4)}$	Population changes due to FIPMs over one year from year $T+b-1$ to $T+b$, representing individuals of sex S who are aged $X+b-1$ to $X+b+3$ in year $T+b-1$ and transition to ages $X+b$ to $X+b+4$ in year $T+b$, within each municipality (persons)
$\hat{\beta}_1$	Change in the net migration rate at the municipality level per 1 % increase in the proportion of households affected below floor level by flooding (percentage points)
$\hat{\beta}_2$	Change in the net migration rate at the municipality level per 1 % increase in the proportion of households affected above floor level by flooding (percentage points)
$\hat{\beta}_3$	Change in the net migration rate at the municipality level per 1 % increase in the proportion of households that were completely destroyed by flooding (percentage points)
$\widehat{BF}_{T+b}$	Expected annual proportion of households affected below floor level by flooding in year $T+b$ (%)
$\widehat{AF}_{T+b}$	Expected annual proportion of households affected above floor level by flooding in year $T+b$ (%)
$\widehat{CD}_{T+b}$	Expected annual proportion of households that were completely destroyed by flooding in year $T+b$ (%)
$P^{\text{mun}}_{T+b-1,S,X+b-1\text{ to }X+b+3}$	Population of sex S, aged $X+b-1$ to $X+b+3$ in year $T+b-1$, within each municipality (persons)
$C_{T\to T+5,S,(X\text{ to }X+4)\to(X+5\text{ to }X+9)}$	Population changes due to FIPMs over the five-year period from year T to $T+5$, representing individuals of sex S who are aged X to $X+4$ in year T and transition to ages $X+5$ to $X+9$ in year $T+5$, within each municipality (persons)
$C_{T+c-1\to T+c,S,(X+c-1\text{ to }X+c+3)\to(X+c\text{ to }X+c+4)}$	Population changes due to FIPMs over one year from year $T+c-1$ to $T+c$, representing individuals of sex S who are aged $X+c-1$ to $X+c+3$ in year $T+c-1$ and transition to ages $X+c$ to $X+c+4$ in year $T+c$, within each municipality (persons)