Gone with the Flood: Natural Disasters, Selective Migration, and Media Sentiment

Yi Fan[*], Qiuxia Gao[†], Yinghao Elliot Sitoh[‡], and Wayne Xinwei Wan[§]

January 15, 2024

Abstract

Exploiting variations in flood timing from the United States as a quasi-natural experiment, this study shows that floods cause population replacement in affected regions with 1.9% inflow and 2.7% outflow migration, respectively. They trigger younger, highly-educated, and employed residents out of, and attract older, less-educated, and unemployed ones into affected zones. Furthermore, the selective migration patterns are amplified by media information provision. The flood-induced selective migration has significant impacts on local economic development, causing a 5.3% decrease in housing prices but a 7.4% increase in housing rent, suggesting a structural change in the housing markets of flood-prone regions. A back-of-envelope calculation shows that flood-induced selective migration conditional on education and age profiles leads to net annual losses of $9.3 million and $4.7 million, respectively, in the flooded counties. Our results shed light on how information provision interacts with migration incentives in wake of natural disasters.

Keywords: Flood, migration, media sentiment, population replacement, residential location choice
JEL Codes: Q54, R23, D83

[*]Department of Real Estate, National University of Singapore; Email: yi.fan@nus.edu.sg
[†]Department of Real Estate, National University of Singapore; Email: e0610864@u.nus.edu
[‡]Department of Real Estate, National University of Singapore; Email: e.sitoh@nus.edu.sg
[§]Department of Banking and Finance, Monash University; Email: wayne.wan@monash.edu

With global climate change, concerns over flood risk have been escalated in the past decade. Recent studies suggest that flood risk is expected to grow by 26.4% by 2050, causing an estimated average annual loss of US$32.1 billion in the US (Wing et al., 2022). Coastal areas in the US could see flooding occurrences no less frequently than annually within the turn of this century (Marsooli et al., 2019), with the likelihood of an event capable of producing catastrophic flood doubled (Huang and Swain, 2022). Higher flood risks induced by climate change are especially troubling because flood-prone areas have historically been hotspots for settlements and urban development (Rentschler et al., 2023; Devitt et al., 2023): in the United States alone, nearly 42% of the population are living in coastal areas (Fleming et al., 2018). As sea levels continue rising, this puts an increasing number of neighbourhoods at risk of floods and inundation. By the end of this century, 2.5 million properties will be at risk of chronic flooding (Dahl et al., 2018).

Human settlement patterns alter in response to the escalating flood risks (Rigaud et al., 2018), which have a profound influence on local economic development beyond direct damage caused by floods. In fact, more than 650 million people worldwide have been displaced by floods in the past three decades (Kocornik-Mina et al., 2020). In the US, lower net migration inflows have been observed in states with higher flood risks in the past two decades. Figure 1a plots the net migration statistics at the state level between 2006 and 2019, while Figure 1b visualizes geographic flood risk within the United States. States with higher flood risk in the western and northeastern regions of the US show net outward migration, while lower-risk states see a general net inflow.

[Insert Figure 1 Here]

Although the impact of floods on human settlements becomes increasingly salient, existing research on migration trends has focused on net migration only. Mixed findings have been documented, with strong net outflow migration observed in response to some natural disasters (Hornbeck and Naidu, 2014; Chen et al., 2017) but little net migration in other cases (Bohra-Mishra et al., 2014). One plausible reason for the puzzling minimal net migration is that the post-disaster recovery plans introduced by local governments may, paradoxically, attract inward migration to these regions. Also, expected flood risks can lower local housing prices (Bernstein et al., 2019), attracting inward migrants who are more sensitive to housing affordability. Furthermore, some sociodemographic features, such as partisanship (Bernstein et al., 2022), are associated with heterogeneous beliefs in flood risks, resulting in selective residential sorting into flood-prone areas. Therefore, while the threats to human life and economic activities are expected to drive some outward migration from the flooded areas, some inward migration will be concurrently attracted to these regions. However, there is little knowledge about the discernible differences in the socioeconomic profiles of the migration inflows and outflows.

To the best of our knowledge, this study is the first to look below the surface, investigating the selective patterns of inflow and outflow migration post-floods separately. Such a partition is important because each of these statistics impacts the affected regions differently. Migrant outflows bleed out human capital and local investment from affected areas, undermining economic development. On the other hand, inflows bring about their own set of problems; demand-pull inflation and sorting effects influence housing prices in receiving counties (Daepp et al., 2023). Assuming the socioeconomic profiles of both migrant groups are similar, immigrants that replace emigrants should mitigate their effects and vice versa. However, if this assumption is breached, the extent of each effect can vary. For instance, if the emigrating residents are more educated than the immigrating ones, the net loss in human capital will adversely affect local economic growth. Also, having older immigrants than emigrants will substantially impact the local fiscal planning for healthcare. Thus, it is crucial that we open this black box of migration and uncover any heterogeneous trends within the different demographics in order to understand both the short- and long-term consequences of flood shocks.

Moreover, the disclosure of climate risk by the media helps to increase awareness among the affected populations and reduce environmental inequality. A rich body of economics literature has documented how the provision of information influences households' decision-making. It is important to understand whether the selective migration patterns will be amplified with improved information transparency on climate risk. Richler (2019) demonstrates that information provision can prompt more risk-aware decisions, successfully nudging consumers towards purchasing flood insurance. Other studies document evidence of information provision causing a price discount for houses in flood-prone zones (Lee, 2022; Niu et al., 2023; Hino and Burke, 2021; Gourevitch et al., 2023). These studies suggest that the supply of climate risk information helps residents better assess their flood risk, prompting movement out of dangerous counties. At the same time, however, such information nudges can also work in reverse—while news articles reporting floods increase the awareness of risks residents may face, these articles also frequently report governmental aid programs introduced as a response to floods (Jia et al., 2022), which would prompt inflow migration into affected counties through a moral hazard effect (Kocornik-Mina et al., 2020; Deng et al., 2023).

Thus, this study investigates two questions. First, we analyze whether floods trigger selective migration—both outflow and inflow migration in different socioeconomic groups. Second, we investigate the role of media sentiment in flood-induced selective migration. In order to answer these questions, we obtain an account of all flood events in the US between 2006 and 2019 from the National Center for Environmental Information and investigate their impact on both immigration and emigration patterns. Our migration data is extracted from the American Community Survey in the same period. Our news dataset is collected from the online database Factiva. Combining the three datasets, we employ a staggered Difference-in-Differences (DID) strategy to get to our answers.

Our results reveal that flood shocks cause both inflow and outflow migration to increase by 1.9% and 2.7%, respectively. More importantly, selective migration patterns are observed across distinct socioeconomic groups: Flood events lead to inflows of lower-potential individuals (older, less educated, unemployed) into affected counties, while simultaneously prompting outflows of higher-potential individuals (younger, more educated, employed). We also find evidence that media sentiment of flood risks influences selective migration patterns. Positive news sentiment (e.g., government's post-disaster relief programs) is associated with a decrease in outflows of higher-potential individuals and an increase in inflows of lower-potential individuals into flooded counties. These selective migration patterns induced by floods have salient economic consequences. Taking the local housing market as an example, they result in a 5.3% decrease in housing prices and a 7.4% increase in housing rent post-flood, suggesting a structural change in the housing markets of flood-prone regions. Using aggregate income as a proxy for the economic output, our conservative back-of-envelope calculations show that flood-induced selective migration by education and age leads to net annual losses of $9.3 million and $4.7 million, respectively, in the flooded counties.

Our study makes the following contributions: Firstly, we build on the literature documenting disaster-triggered migration (Bohra-Mishra et al., 2014; Hornbeck, 2012). We study how flood events affect intra-national inward and outward migration patterns in the US. To the best of our knowledge, this is the first work studying the effect of flood events on inflow and outflow migration separately. Prior studies have focused on net migration outflows, due to these regions becoming less attractive residential choices post-disaster. However, we theorize that disaster relief programs may stimulate economic activity (Heger and Neumayer, 2019) within these areas that incentivize immigration into these regions. We also further study the heterogeneous patterns of flood-induced migration, relating to literature reporting the inequality in the impact of floods across different socioeconomic groups. Wing et al. (2022) document evidence that flood damage in the US is borne disproportionately by the poor. Lindersson et al. (2023) report an association between income inequality and flood mortality. Our heterogeneous findings complement these studies, as we provide insight into the reasons behind this disparity. Finally, we also add to the active literature studying the impact of information nudges on residential choice (Hino and Burke, 2021; Lee, 2022). Our findings demonstrate how information sentiment can trigger heterogeneous migration patterns across different socioeconomic groups in wake of natural disasters.

Results

Event Study

We estimate the impact of flood events on triggering migration using an event-study strategy from Equation (1); that is, we investigate the year-by-year changes in migration in flooded counties ("treatment group") relative to the corresponding changes in adjacent non-flooded counties ("control group") over the same period. Figure 2 plots event study results on the differences between the annual migration flows in the treatment and control groups in the $[-3, +3]$ year window around any flood event, estimated from two methods—staggered TWFE and CSDID (details in **Methods**). With either method, we consistently observe positive growth in both outward and inward migration in the flooded counties after a flood shock, relative to the migration numbers in surrounding non-flooded counties. The by-year increases in outward migration are also slightly larger than the increases in inward migration, but the differences are not statistically significant, implying a small increase in net outflow migration numbers post-floods.

[Insert Figure 2 Here]

Notably, the validity of the DID strategy hinges on the assumption of parallel pre-trends, which states that in the absence of treatment (a flood event), both treatment (flooded counties) and control groups (adjacent non-flooded counties) would have seen migration trends growing at identical rates. Figure 2 also reveals that, with either estimation method, the differences between the treatment and control groups for both outflow and inflow migration are not statistically different from zero in the pre-treatment period, while the difference turns positive post-treatment. Thus, the assumption of parallel pre-trends is likely held in our analysis.

Average Treatment Effect

We further quantify the average treatment effects using Equation (2) in **Methods**, and the results are reported in Table 1. The coefficient of $Treat \times Post$ is interpreted as the average post-flood change in migration in flooded counties, relative to their surrounding non-flooded counties in the same period. Dependent variables in Columns (1) to (3) are the logarithmic forms of migration outflows, inflows, and net outflows, respectively, so the coefficients of $Treat \times Post$ represent their percentage changes. We find that three years after flood events, the outflow and inflow migration in flooded counties increase by 2.7% (Column (1)) and 1.9% (Column (2)), respectively, relative to adjacent non-flooded counties. The two estimates are statistically significant at the 1% level. Consistent with previous studies, we also find that floods result in significant net migration out of flooded counties, as shown in Column (3). The patterns are consistent when the actual numbers of outflow/inflow/net migration, rather than their logarithmic forms, are used as dependent variables under investigation (Columns (4)–(6)). In summary, these results reveal that flood events cause the number of migrants both out of and into flooded counties to increase.

[Insert Table 1 Here]

Selective Migration

Since floods cause increases in both immigration and emigration, we proceed to investigate the demographics within these immigrant and emigrant populations. If selective migration patterns (i.e., distinct sociodemographic statuses) exist in the inflow and outflow population, assessing the effects of flood shocks on a local population through net migration flows alone may conceal underlying economic consequences. Notably, it is important to examine the different behaviours between higher-potential individuals (the younger, employed, and more educated) and lower-potential individuals (the older, unemployed and less educated) for two reasons. First, we posit that these groups of people are likely to respond differently to natural disaster shocks. Second, the distributions of these demographics have substantially different long-term consequences on the various aspects of the local economy in these counties, such as labor supply, housing, education, and health.

Specifically, to investigate the selective migration patterns, we separate the outflow and inflow of migrants across three dimensions: (1) education, (2) employment and (3) age. We compare high education (at least college degree holders) versus low education (up to high school diploma), employed versus non-employed (in the original location of residence in 1 year before the flood events), young (< 40 years old) versus old (≥ 40 years old).

Figure 3 presents the different migration responses to the flood events between the higher-potential and lower-potential individuals, estimated from Equation (2) (the complete regression results are reported in Appendix Table A2). For each subgroup of migrant population, the point estimates represent the corresponding percent changes in migrant numbers in the flooded counties over the [-3, +3] year window around the year of the flood, relative to adjacent non-flooded counties over the same period. Figure 3a illustrates the selective patterns in emigration: Higher-potential individuals in flood-prone counties—more educated, employed, or younger residents—are more likely to move out. After a flood event, the number of outflow migrants in these groups increases by 4.4%, 3.0% and 3.2%, respectively. However, we do not find a statistically significant increase in emigration among the unemployed and older groups. There is an increase in emigration among the less educated, but the size of the impact (1.7%) is much smaller than and statistically different (at 90% level) from the impact on the more educated group.

[Insert Figure 3 Here]

In contrast, the patterns are entirely different in the immigration flows to affected counties (Figure 3b). After a flood shock, the number of immigrants who are less educated, unemployed, or older increases by 2.7%, 3.7%, and 4.1%, respectively. However, we do not find statistically significant increases in immigration among the converse socioeconomic groups. Thus, these findings imply that, unlike the patterns in outflow migrants, lower-potential individuals are more likely to move into flood-prone counties in spite of recent floods.

These empirical results demonstrate distinct patterns of selective migration decisions being made across different demographics. Upon a flood event, we see a clear trend of higher-potential individuals moving out of affected counties, while these events attract lower-potential individuals into affected counties. We hypothesize that such patterns emerge as flood shocks cause current residents of affected counties to re-evaluate the risk of their current location, prompting them to emigrate. As higher-potential individuals are more able to relocate (Florida, 2019), we see higher levels of emigrants within these demographics. Conversely, lower-potential people are attracted to flooded regions as they see new opportunities that may arise from these events; as floods shock the local economy, the devaluation of property in the area makes these areas more affordable (Borenstein, 2023).

Strengthening Effect of Media Sentiment

Finally, we investigate whether media sentiment influences our observed post-disaster selective migration patterns. We posit that these patterns partially result from different flood risk perceptions between the higher-potential and lower-potential individuals. Since a rich expanse of literature has documented how information provision alters the climate risk evaluation of agents (Lee, 2022; Niu et al., 2023; Hino and Burke, 2021; Richler, 2019), we hypothesize that different socioeconomic groups will respond to identical sources of information differently, leading to a strengthening effect on selective migration. This is of great importance in today's landscape, where information on flood risk is becoming increasingly accessible. As the most mainstream source of information for the everyday person, the influence that news has on general society is massive. Thus, we look to study how the transmission of flood risk information through news media influences migration decisions across the different demographics.

The media sentiment is measured using a machine learning method that extracts the news sentiment from a comprehensive database of newspaper articles published in the US over the study period from 2006 to 2019 (details in **Methods**). This period under examination encompasses the period newspapers played the predominant source for information distribution. Even in the present day, the pre-eminent choice for the majority of U.S. residents in obtaining news continues to be news websites and publishers.[1]This enduring preference reinforces our justification for utilizing newspapers as a reliable medium for information transmission, as opposed to alternative sources like social media.

In order to establish a causal inference of news sentiment on migration, we interacted this

[1]See: https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/

variable with our initial study to conduct a triple-difference analysis. To rule out potential confounders that may bias our analysis, several additional controls are included in this regression, including flood-related subsidies, housing price, firm entries and exists at the county level, because these variables are also likely to impact migration patterns. Specifically, subsidies entice potential immigrants/existing residents as it provides monetary incentive to stay within these counties. Also, the housing market may fluctuate post-disaster, and cheaper housing prices motivate immigration into the county. Furthermore, firm entries result in job creation, thus promoting movement into these regions, and vice versa for firm exits. Last, the flood event fixed effects are included to capture the severity of and damage caused by individual floods, which can correlate with the news sentiment.[2]

Table 2 reports the estimated impacts of news sentiment on inflow and outflow migration after floods. The coefficient of interest here is $Treat \times Post \times Score$, which represents the impact of news sentiment on the percent changes on migration numbers in the flooded counties, relative to adjacent non-flooded counties. We see that "good news" (e.g., post-disaster recovery plan) triggers a decrease in movement out of flooded counties for the high education and employed individuals, where the number of emigrants in these demographics are lowered respectively by 2.8% and 1.6% with a standard deviation increase in news sentiment. All three estimates are statistically significant at the 1% level. These patterns are mirrored for the young, which also see a 1.8% decrease in emigration. Conversely, we see an inverse trend for inflow migration, where positive news sentiment attracts low education and unemployed individuals into affected counties post-flood. We see a one-standard-deviation increase in news sentiment is associated with a 1.6%, 2.4%, and 1.6% increase in immigration for low-educated, unemployed and old immigrants, respectively, with statistical significance at conventional levels.[3]

[2]We concede that the inclusion of the additional controls mentioned above may create potential "bad control" issues if the channel through which flood shocks and media sentiment affect migration flows through these variables. However, should this be the case, the bias that arises from their inclusion will attenuate our estimates. Our results will still provide meaningful insight, as they provide a lower bound (in magnitude) of the true effects of media awareness.

[3]One example of "good" news comes from an article published by USA Today in September 2017, which reported that *"The House passed a $7.9 billion aid package Wednesday for victims of Hurricane Harvey, and the Senate is expected to attach a debt-limit extension to that bill...McConnell said that he would be supportive of the plan and intended to offer it as an amendment to the flood relief bill that passed the House on Wednesday"*. An example of

[Insert Table 2 Here]

Discussion

This paper provides the first empirical evidence detailing the effects of floods on selective migration patterns across different sociodemographic groups within the United States. Higher-potential individuals are more likely to leave flooded areas, while lower-potential individuals are unexpectedly attracted to move into the flooded areas, plausibly due to better housing affordability and post-disaster recovery subsidies from governments. These selective migration patterns thus give rise to a replacement effect in the local population, as incoming immigrants are systematically different from recent emigrants. This trend is potentially a cause for concern in affected counties, as it has both short and long term implications on the local economy within them. In the short term, local markets may see demand shocks as recent emigrants/immigrants may have different priorities and thus make different decisions. In the long-term, the development of the local economy may be hindered as more-abled workers migrate out. We therefore detail the profound implications of these effects on local economic developments.

Shorter-term effect We first use the local housing market as an indicator of the local economic consequences of selective migration in the short term (Baerlocher et al., 2023). The housing market reflects the immediate decisions made by migrants within treatment and adjacent control counties after a flood event, as moving into a new region requires them to find new housing arrangements. Using housing price and monthly rent data from Zillow, we adopt the same DID empirical design to investigate the impacts of flood events on housing price and rent. Since the housing rent data from Zillow is available from 2015 onward, we use the monthly observations of housing price and rent from 2015 to 2019 for the analysis,[4] and the results are reported in Table 3.

"bad" news is an article published by The New York Times in September 2006, which reported: *"The early estimates suggested insured property damage around $5 billion or less from Hurricane Rita, not including the effects of flooding and the impact on offshore oil rigs, which are excluded in most of the calculations. The storm struck less heavily populated areas, with less force than Hurricane Katrina, mostly bypassing Galveston, Tex., and Houston, where damage up to $30 billion had been feared"*.

[4]The housing price data is available in our entire study period from 2006 to 2019, and we confirm the results are consistent using the full sample of housing price.

A flood event results in a 5.3% decrease in housing prices in flooded counties, relative to adjacent non-flooded counties. However, the housing rent in the flooded counties increases by 7.4% post-flood. These figures are significant at the 95% confidence interval, and they translate to an average $8,030 drop in total housing prices and a $53.50 spike in monthly rent. These patterns observed in the housing market are plausibly due to selective migration patterns that arise from flood shocks. As higher-potential residents move out and lower-potential residents move in, the resulting population sees a shift in priorities: lower-potential immigrants with less wealth seek to rent housing, while the higher-potential emigrants seek to sell their recently deserted property. This causes an increase in the supply of property sold, while an increase in the demand of rental properties. This coordination failure within the new emigrants and immigrants causes housing prices to fall and rental prices to increase. These drastic fluctuations in the housing market demonstrate the striking short-term economic consequences of selective migration in affected counties.

[Insert Table 3 Here]

Longer-term effect These structural changes in the housing markets have further implications in the long term as well. As flood-prone counties cater to increasing renters, it gives rise to a moral hazard whereby homeowners are unwilling to renovate their homes in order to maintain low renting fees. This is especially concerning as many residential buildings in the US are not built to handle climate change, and would leave the rental population more vulnerable to the elements (Sisson, 2023). This may exacerbate the already present disparity in flood risk and flood damage borne across the rich and poor (Wing et al., 2022; Lindersson et al., 2023).

Apart from housing markets, flood-induced selective migration also has salient impacts on the local labor markets. As flood events trigger movements of higher-educated individuals out of and lower-educated ones into flood-prone counties, the resulting brain drain also has adverse effects on local economic growth. We use the aggregate changes in individual income as a proxy for GDP and conduct a back-of-envelope estimation of the impacts (Appendix Tables A5). Our regression results in Figure 3 reveal that flood events result in an increase in high-educated outflow migrants by 4.4% and low-educated outflow migrants by 1.70%. After adjustment for the sampling rate in the ACS (1%), these translate to 274 high-educated and 181 low-educated emigrants per year. Further multiplying these numbers by the average income of the high- and low-education groups, we estimate a total annual loss of $14.2 million for high-educated emigrants and $2.9 million for low-educated emigrants per county. In contrast, we estimate a total annual gain of $3.2 million for high-educated immigrants and $4.5 million for low-educated immigrants per county after the flood events. Therefore, the total annual net loss due to selective migration by education levels is estimated to be around $9.3 million per county. In a worse scenario (i.e., consider the upper bound of 95% CI for the changes in outflow migration and the lower bound of 95% CI for the changes in inflow migration), the net loss can be as large as $25.9 million. Similarly, we estimate the net loss due to flood-induced selective migration by age (Appendix Table A6).[5] A higher outflow of young people and an increased inflow of old people result in a net annual loss of over $4.7 million, after factoring the 1% sampling rate. This reduction in output has far-reaching consequences on economic growth as well, thus demonstrating another dimension of damage that these selective migration patterns cause.

Furthermore, we find that the selective migration patterns are further amplified when media improves information transparency on climate risks, because high- and low-potential individuals are likely to be nudged by good and bad news differently. On one hand, news reports increase awareness of flood exposure, causing agents to re-evaluate risk perception of their current location. This thus motivates current residents to leave, and dissuades potential migrants from moving in. However, this is a double-edged sword. Such news may also create a moral hazard effect for potential emigrants and immigrants. Firstly, reports on floods may promote the less well-off to move into affected counties, as they expect property prices to fall following a disaster. Secondly, these articles may provide information of governmental restoration policies to generate economic redevelopment in affected counties, which these agents may see as new opportunities. We highlight this as a concern as this moral hazard effect may endanger residents in flood-prone counties as they undervalue the risk of staying in these areas for potentially misconstrued economic gains. This is

[5]We skip the back-of-envelope analysis for selective migration by employment status, because unemployed migrants can be considered to have no personal income.

especially concerning for the older and poorer immigrants who are more vulnerable to damage caused by floods.

As info-communication technology continues to develop, news and general information will become increasingly available to the public. This not only increases the exposure that members of society have to information, but also makes it more transparent. Authorities have a better platform and thus greater responsibility to transmit information to their people. We support the push for greater transparency in news media, and hope that our results do not detract from this position. Instead, we would like our results to better inform authorities and news outlets how their portrayal of events can differentially affect different demographics.

Due to global warming and climate change, the number of flood-prone areas are expected to see a rise globally over the next century. Migration out of zones at risk is the best solution to such threats (Fagan, 2008). Our results shed light on how natural disasters influence migration, and beyond that, how governmental and media responses to them may further influence peoples' decisions. We propose that a more tempered response is key in minimizing exposure to environmental disasters.

Method

Data

Migration

We use the American Community Survey (ACS) waves from 2006-2019 to collect our migration data at the individual level. The ACS is an ongoing nationally representative survey with a sample rate of 1% of the total population, conducted over a 12-month period for geographic areas with at least 65,000 people. It collects individuals' information on economic characteristics (employment status, occupation, income, etc.), housing characteristics (facility, tenure, house structure, etc.), demographic characteristics (age, sex, race, etc.) and social characteristics (educational attainment, marital status, migration, previous and current residence).

Regarding migration, the ACS reports whether an individual has departed their original residence, and includes their previous and current house locations. Specifically, individuals were asked if they had lived in the "same house" (non-movers) or a "different house" (movers) one year ago. We group our data at county level, and redefine individuals as movers when they move out of their origin county to exclude people who move to different houses within the same county. Observations with missing data on mobility and location are excluded. We use their one-year-prior residence to aggregate the annual outflow for each county, and current residence to aggregate the annual inflow of counties.

The ACS also reports survey respondents' age, employment status, income and highest education qualification, which was used in our heterogeneity analysis. Appendix Table A1 reports the summary statistics for our migration data.

Flood Data

We rely on the data from the National Center for Environmental Information[6] to identify counties with a history of flooding across the United States. The dataset includes the location (states, counties and zones) of 48 different types of natural disaster events and detailed information on the events, which includes the start time, end time, number of injured victims, damages to property and cause of disaster, from January 1950 to October 2022. To identify the counties with flood history, we keep events related to floods and flash-floods, and calculate the number of flood events, average time of duration, time between two events each year at the county level. Finally, we obtain information from 2,218 flood events between 2006-2019 in 360 counties.

Media Awareness

To obtain our database of flood-related news articles, we scraped the online database Factiva for all news articles from the 5 most subscribed national newspapers (Wall Street Journal, New York Times, USA Today, Washington Post, and Los Angeles Times) that included keywords of "flood

[6]https://www.ncei.noaa.gov

damage", "flood recovery", "flood relief", "flood subsidy(ies)", "disaster funds" & "flood". We then ran a sentiment analysis through all the articles to obtain sentiment scores for each article, using the natural language processing package TextBlob [7]. Following this, we standardized the scores of each article, transforming them such that the mean and variance of sentiment scores are 0 and 1 respectively. We constructed the "state" dimension by identifying articles that mention any US state(s) and extracting the state name. In order to capture the strongest shocks from the news articles each year, our measure of media awareness is the minimum sentiment score (i.e., the strongest negative sentiment) of articles at the year-state level.

Housing Price

The housing price data we use in this study is obtained from Zillow, a real estate data vendor that estimates home values in approximately 8,000 neighborhoods in metropolitan areas across the United States. Specifically, we use the Zillow Home Value Index (ZHVI) for all property types at the county level, which is a smoothed, seasonally adjusted measure of the typical home value (i.e., homes in the 35th to 65th percentile range) in a given region, including newly constructed homes and/or homes that have not traded on the open market.

The housing rent data is also obtained from Zillow. Specifically, we use the Zillow Observed Rent Index (ZORI), a smoothed measure of the observed monthly market rent of typical homes (listed rents that fall into the 40th to 60th percentile range for all homes and apartments) in a given region. The ZORI has been adjusted using a repeat-rent method, which is weighted to the rental housing stock, so it is representative across the entire market, not just those homes currently listed for-rent. Both the housing price and rent data were collected at the county and year level. We match our flood event data with the Zillow housing price data between 2006 and 2019, while the rent data is available between 2015 and 2019 only.

[7] https://textblob.readthedocs.io/en/dev/

Other Control Variables

We consider the following macroeconomic control variables: GDP, population size, unemployment rate, and personal income data, which was obtained from the US Bureau of Economic Analysis [8]. The data provides annual information for a variety of macroeconomic indicators from 2006 to 2019 at the county level in the US. The public firm entry and exit data are obtained from the Augmented 10-X Header Database, [9] which collects the headquarter addresses from the annual financial reports of all public firms listed in the three stock exchanges in the US. The information on government subsidies provided for post-flood economic recoveries is collected from the OpenFEMA Dataset, an administrative data provided by the US Department of Homeland Security.[10].

Event Study Analysis

To verify that the assumption of parallel pre-trends holds in our DID estimation, we conduct an event study analysis to check the trends between treatment and control groups prior to treatment using equation 1 as follows:

$$\begin{aligned}
Y_{i,t} = {} & \alpha_1 Treat_{i,j,t} + \sum_{p=2,3} \beta_p Pre^p_{i,j,t} + \sum_{q=1}^{3} \beta_q Post^q_{i,j,t} + \sum_{p=2,3} \delta_p Treat_{i,j,t} \times Pre^p_{i,j,t} \\
& + \sum_{q=1}^{3} \delta_q Treat_{i,j,t} \times Post^q_{i,j,t} + X'_{i,t}\lambda_X + \omega_i + \theta_t + \mu_{s,t} + \rho_j + \epsilon_{i,j,t}.
\end{aligned} \tag{1}$$

Specifically, $Y_{i,t}$ is the outcome variable of interest for county i in year t, such as the total number of outflow migrants (in logarithmic form). $Treat_{i,j,t}$ is a dummy variable equal to 1 for county-year observations in the treatment group of flood event j. It equals 0 for the matched county-year observations in the control group. We use a set of dummy variables denoting each year in the [-3, -2] years window before the treatment ($\sum_{p=2,3} Pre^p_{i,j,t}$) and the [+1, +3] years window after the treatment ($\sum_{q=1}^{3} Post^q_{i,j,t}$). Using 1 year before treatment time t as the base group, we estimate the differences between the migration flows in the treatment and control groups in each

[8] https://www.bea.gov

[9] https://sraf.nd.edu/data/augmented-10-x-header-data/

[10] https://www.fema.gov/openfema-data-page/public-assistance-funded-projects-details-v1

year in the $[t-3, t+3]$ window, represented by the coefficients of the interaction terms (δ_p and δ_q). If the assumption of parallel pre-trends holds, we should expect that differences between the migration flows in the treatment and control group are not statistically different from zero in the pre-treatment period (δ_p), while the difference turns positive post-treatment (δ_q).

$X_{i,t}$ is a set of macroeconomic control variables at the county-year level, including GDP (in logarithmic form), unemployment rate, population size (in logarithmic form) and average income per capita (in logarithmic form). ω_i and θ_t denote the county and year fixed effects, respectively. $\mu_{s,t}$ represents the state times year fixed effects for state s which county i belongs to, which captures the time-variating confounding factors at the state level, such as state government subsidies. We also include the flood event fixed effects (ρ_j), which control for the potential unobserved variations in treatment sizes. $\epsilon_{i,j,t}$ is the error term. We cluster the standard errors at the flood event level.

Notably, the flood events shocked the treated counties in a staggered fashion. Recent literature has highlighted potential bias in traditional staggered DID estimates, with several alternative methods being proposed to address the issue (Baker et al., 2022).[11] Our baseline model and the event study model are close to one of these alternative methods—the stacked DID design with two-way fixed effects (TWFE), which creates event-specific data sets that are then stacked together (Cengiz et al., 2019). The difference here is that a standard stacked DID design requires the control group to be never treated, while we relax this requirement due to the lack of counties that are never flooded; instead, we select surrounding counties that have no floods in the $[t-3, t+3]$ window as the control group. To further alleviate the concern, we implement another alternative method—the Callaway and Sant'Anna DID strategy (CSDID)—as a robustness check (Callaway and Sant'Anna, 2021). This method estimates all possible good comparisons for the estimation and then aggregates them by putting more weight on larger and more precise estimators. We present the CSDID results in both our event study and baseline DID analysis. Moreover, we conduct Hon-

[11]The bias mainly originates from the variations in treatment timing and heterogeneous treatment effects. The parameter in the standard DID estimates the average difference by comparing the same unit across time, as well as by comparing different units with and without treatment at the same point in time. If the treatment effects are heterogeneous and the timing of the treatment varies across units, comparing later treated units with earlier treated units (control units) can be problematic (bad comparison). As control observations were already treated, the standard DID model would assign negative weights to these units.

est DID to assess the sensitivity to violations of parallel trends in a CSDID model (Rambachan and Roth, 2023).

Baseline Estimation

Our baseline analysis employs a DID strategy. Intuitively, this empirical design seeks to estimate the causal impact of flood events on migration by comparing the migration trends in control and treatment groups before and after treatment. The treatment group consists of counties with flood events occurring, and the control group consists of surrounding counties to treatment counties. For each flood event[12], we keep only the surrounding counties that did not have floods in the [-3, +3] year window as the control counties. In order to ensure that all sampled flood events have at least one eligible control country that satisfies this requirement, our main sample ends up using surrounding counties that are within the twelve nearest counties as the control group. Then, we collect the annual migration numbers in the treatment and control counties of each flood event around the [-3, +3] year window as the regression sample.

The baseline DID regression model is specified as follows:

$$Y_{i,t} = \beta_1 Treat_{i,j,t} + \beta_2 Post_{i,j,t} + \beta_3 Treat_{i,j,t} \times Post_{i,j,t} + X^{'}_{i,t}\lambda_X + \omega_i + \theta_t + \mu_{s,t} + \rho_j + \epsilon_{i,j,t}. \quad (2)$$

For the county-year observations in both the treatment and control groups of flood event j, the dummy variable $Post_{i,j,t}$ equals 1 if year t is after the occurrence of flood j. Otherwise, it equals zero. Therefore, the coefficient of the interaction term $Treat_{i,j,t} \times Post_{i,j,t}$ represents the average impact of floods on migration flow in the affected counties. Other variables have the same definitions as in Equation (1). Standard errors are clustered at the flood event level.

Our heterogeneous analyses across the different demographics also use Equation (2), imposing the relevant restrictions on the outcome variable $Y_{i,t}$. For example, in the analysis of outward migration of young residents, we use the number of young migrants, rather than the total number

[12] Since most of our outcome variables are available at the county-year level, in our staggered DID model, we consider multiple floods that occurred in a county within the same year as one flood event.

of migrants, as the outcome variable. The control variables and fixed effects remain the same as in Equation (2).

Media Sentiment

To study the effect of media awareness on migration choice, we ran a triple-difference identification strategy using the following model:

$$\begin{aligned} Y_{i,t+1} &= \beta_1 Treat_{i,j,t} + \beta_2 Post_{i,j,t} + \beta_3 Treat_{i,j,t} \times Post_{i,j,t} + \beta_4 Score_{i,t} + \beta_5 Score_{i,t} \times Treat_{i,j,t} \\ &+ \beta_6 Score_{i,t} \times Post_{i,j,t} + \beta_7 Score_{i,t} \times Treat_{i,j,t} \times Post_{i,j,t} + X'_{i,t}\lambda_X + \omega_i + \theta_t + \mu_{s,t} + \rho_j + \epsilon_{i,j,t}. \end{aligned} \tag{3}$$

This model is similar to the baseline Equation (2), but with key differences in the additional full interactions between the news sentiment ($Score_{i,t}$) and the dummy variables $Treat_{i,j,t}$ and $Post_{i,j,t}$. Thus, the coefficient of the triple interaction term (β_7) is the variable of our interest, which represents the differential treatment effect on the outcome variable in the post-treatment period resulting from the sentiment in news reports. Other control variables and fixed effects remain unchanged as in Equation (2). In particular, with the control of flood event fixed effects that captures the treatment intensity (i.e., severity) of each specific flood event, our specification identifies the variations in migration flow driven by additional information nudges beyond the severity of an individual environmental disaster.

Robustness Check

We reinforce our causal inference by conducting a battery of robustness checks. Appendix Table A3 reports the results of our robustness checks on the baseline analysis. Column (1) includes additional controls for public firm entry/exit, flood-related subsidies and housing prices. Column (2) uses a winsorized sample, removing the top and bottom 1% of the sample. Column (3) clusters standard errors at the county level. In Column (4), we restrict the sample to flood events that occurred in counties with high flood frequency (above the median), aiming to check if the impact attenuates in areas that often experience floods. Considering that adjacent control counties may be affected by the flood events in treatment counties, in Column (5), we exclude control counties within 100 km of the treatment counties in our sample. This eliminates concerns that our control counties may in fact be treated, as emigrants from treatment counties may be more likely to move into nearby regions post-flood. Column (6) uses the CSDID method to estimate the average treatment effect.

We also conduct a robustness check for the impact of media awareness on selective migration. In our main result, we include additional controls, such as housing prices, flood-related subsidies granted at the county-year level, and the number of firms entering and leaving each county in a given year, to shut off the alternative channels. Nevertheless, if these control variables are outcomes of flood events, the concerns of the bad control issue will arise. In our robustness check, we remove these additional controls in the regression (Appendix Table A4).

References

Baerlocher, D., Cortes, G., and Vinicios, S. (2023). Dust bowl migration and housing in Los Angeles. *Working Paper*.

Baker, A. C., Larcker, D. F., and Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2):370–395.

Bernstein, A., Billings, S. B., Gustafson, M. T., and Lewis, R. (2022). Partisan residential sorting on climate change risk. *Journal of Financial Economics*, 146(3):989–1015.

Bernstein, A., Gustafson, M. T., and Lewis, R. (2019). Disaster on the horizon: The price effect of sea level rise. *Journal of Financial Economics*, 134:253–272.

Bohra-Mishra, P., Oppenheimer, M., and Hsiang, S. M. (2014). Nonlinear permanent migration response to climatic variations but minimal response to disasters. *Proceedings of the National Academy of Sciences of the United States of America*, 111:9780–9785.

Borenstein, S. (2023). Study finds more people are moving into high flood zones, increasing risk of water disasters. *Associated Press News*.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.

Chen, J. J., Mueller, V., Jia, Y., and Tseng, S. K.-H. (2017). Validating migration responses to flooding using satellite and vital registration data. *American Economic Review*, 107(5):441–45.

Daepp, M. I., devin michelle bunten, and Hsu, J. W. (2023). The effect of racial composition on neighborhood housing prices: Evidence from hurricane katrina-induced migration. *Journal of Urban Economics*, 134.

Dahl, K., Cleetus, R., Spanger-Siegfried, E., Udvardy, S., Caldas, A., and Worth, P. (2018). Underwater: Rising seas, chronic floods, and the implications for us coastal real estate. *Cambridge, MA: Union of Concerned Scientists. Online at www. ucsusa. org/sites/default/files/attach/2018/06/underwater-analysis-full-report. pdf*.

Deng, Y., Lin, L., and Meng, L. (2023). Sea level rise and urban adaptation in jakarta. *Working Paper*.

Devitt, L., Neal, J., Coxon, G., Savage, J., and Wagener, T. (2023). Flood hazard potential reveals global floodplain settlement patterns. *Nature Communications*, 14:2801.

Fagan, B. M. (2008). *The Great Warming: Climate Change and the Rise and Fall of Civilizations*. Bloomsbury Press.

Fleming, E., Payne, J. L., Sweet, W. V., Craghan, M., Haines, J., Hart, J. A. F., Stiller, H., and Sutton-Grier, A. (2018). Chapter 8 : Coastal effects. impacts, risks, and adaptation in the united states: The fourth national climate assessment, volume ii. Technical report, U.S. Global Change Research Program.

Florida, R. (2019). The double whammy: Housing and income inequality. *Bloomberg*.

Gourevitch, J. D., Kousky, C., Liao, Y., Nolte, C., Pollack, A. B., Porter, J. R., and Weill, J. A. (2023). Unpriced climate risk and the potential consequences of overvaluation in us housing markets. *Nature Climate Change*, 13(3):250–257.

Heger, M. P. and Neumayer, E. (2019). The impact of the indian ocean tsunami on aceh's long-term economic growth. *Journal of Development Economics*, 141.

Hino, M. and Burke, M. (2021). The effect of information about climate risk on property values. *Proceedings of the National Academy of Sciences*, 118(17):e2003374118.

Hornbeck, R. (2012). The enduring impact of the american dust bowl: Short- and long-run adjustments to environmental catastrophe. *American Economic Review*, 102:1477–1507.

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

Huang, X. and Swain, D. L. (2022). Climate change is increasing the risk of a california megaflood. *Science Advances*, 8(32):eabq0995.

Jia, R., Ma, X., and Xie, V. W. (2022). Expecting floods: Firm entry, employment, and aggregate implications. Working Paper 30250, National Bureau of Economic Research.

Kocornik-Mina, A., McDermott, T. K., Michaels, G., and Rauch, F. (2020). Flooded cities. *American Economic Journal: Applied Economics*, 12:35–66.

Lee, S. (2022). Adapting to natural disasters through better information: Evidence from the home seller disclosure requirement. *Working Paper*.

Lindersson, S., Raffetti, E., Rusca, M., Brandimarte, L., Mard, J., and Di Baldassarre, G. (2023). The wider the gap between rich and poor the higher the flood mortality. *Nature Sustainability*, 6:995–1005.

Marsooli, R., Lin, N., Emanuel, K., and Feng, K. (2019). Climate change exacerbates hurricane flood hazards along us atlantic and gulf coasts in spatially varying patterns. *Nature Communications*, 10:3875.

Niu, D., Kok, N., and Eichholtz, P. (2023). Information provision and flood risk salience. *Working Paper*.

Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, page rdad018.

Rentschler, J., Avner, P., Marconcini, M., Su, R., Strano, E., Vousdoukas, M., and Hallegatte, S. (2023). Global evidence of rapid urban growth in flood zones since 1985. *Nature*, 622:87–92.

Richler, J. (2019). Nudge to flood insurance. *Nature Climate Change*, 9:906.

Rigaud, K. K., Sherbinin, A. D., Jones, B., Bergmann, J., Clement, V., Ober, K., Schewe, J., Adamo, S., Mccusker, B., Heuser, S., and Midgley, A. (2018). Preparing for internal climate migration groundswell. Technical report, World Bank.

Sisson, P. (2023). The other housing crisis: Too many sick, aging homes. *Bloomberg*.

Wing, O. E., Lehman, W., Bates, P. D., Sampson, C. C., Quinn, N., Smith, A. M., Neal, J. C., Porter, J. R., and Kousky, C. (2022). Inequitable patterns of US flood risk in the anthropocene. *Nature Climate Change*, 12:156–162.

Table 1: Impact of Floods on Migration

	(1) log(Outflow)	(2) log(Inflow)	(3) log(NetOutflow)	(4) Outflow	(5) Inflow	(6) NetOutflow
Treat × Post	0.027***	0.019***	0.010**	7.766***	3.552***	4.214***
	(0.005)	(0.007)	(0.004)	(1.106)	(1.281)	(1.601)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	16,405	16,405	16,405
R-squared	0.98	0.97	0.69	0.99	0.98	0.70
Mean Dependent Variable	4.83	4.80	5.92	169.85	167.09	2.76

Notes: Unreported macroeconomic control variables include unemployment rate, population, income per capita and GDP at the county-year level. In Column (3), before transforming the net outflow migration number into the logarithmic form, we subtract the minimum net outflow migration number in our sample from it to get a non-negative value. Standard errors are clustered at the flood event level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Impact of Media Sentiment on Selective Migration

Panel A. Outflow Migration

	(1) Total log(Outflow)	(2) High Education log(Outflow)	(3) Low Education log(Outflow)	(4) Employed log(Outflow)	(5) Unemployed log(Outflow)	(6) Young log(Outflow)	(7) Old log(Outflow)
Treat × Post × Score	-0.016***	-0.028***	-0.012**	-0.016***	-0.009	-0.018***	-0.013
	(0.005)	(0.007)	(0.006)	(0.005)	(0.008)	(0.005)	(0.008)
Treat × Post	0.022***	0.036***	0.013*	0.025***	-0.000	0.027***	0.012
	(0.006)	(0.009)	(0.007)	(0.006)	(0.011)	(0.006)	(0.010)
Subsidy	0.009	0.001	0.008	0.009	0.016	0.010	0.016
	(0.008)	(0.011)	(0.010)	(0.008)	(0.014)	(0.010)	(0.013)
log(Housing Price)	0.027	0.061	-0.010	0.102***	-0.197***	0.034	-0.031
	(0.030)	(0.045)	(0.035)	(0.031)	(0.053)	(0.032)	(0.053)
Firm Entry	0.016***	0.013**	0.017***	0.010***	0.025***	0.017***	0.023***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.007)	(0.005)	(0.006)
Firm Exit	0.010**	-0.011*	0.018***	-0.006	0.045***	0.007	0.018***
	(0.004)	(0.006)	(0.005)	(0.005)	(0.007)	(0.004)	(0.007)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	16,405	16,405	16,405	16,405
R-squared	0.98	0.96	0.96	0.97	0.93	0.97	0.93
Mean Dependent Variable	4.83	3.72	4.38	4.54	3.40	4.53	3.47

Panel B. Inflow Migration

	(1) Total log(Inflow)	(2) High Education log(Inflow)	(3) Low Education log(Inflow)	(4) Employed log(Inflow)	(5) Unemployed log(Inflow)	(6) Young log(Inflow)	(7) Old log(Inflow)
Treat × Post × Score	0.010*	-0.000	0.016**	0.004	0.024***	0.010*	0.016*
	(0.006)	(0.007)	(0.007)	(0.006)	(0.009)	(0.006)	(0.009)
Treat × Post	0.020***	0.009	0.030***	0.011	0.042***	0.011	0.042***
	(0.007)	(0.010)	(0.009)	(0.008)	(0.012)	(0.008)	(0.012)
Subsidy	0.002	0.011	-0.009	-0.000	-0.015	-0.008	0.038***
	(0.008)	(0.011)	(0.008)	(0.008)	(0.012)	(0.008)	(0.013)
log(Housing Price)	-0.216***	-0.071	-0.365***	-0.154***	-0.367***	-0.152***	-0.341***
	(0.036)	(0.050)	(0.041)	(0.038)	(0.059)	(0.040)	(0.055)
Firm Entry	0.009**	0.017***	0.007	0.006	0.032***	0.015***	-0.008
	(0.004)	(0.006)	(0.005)	(0.005)	(0.008)	(0.005)	(0.007)
Firm Exit	0.002	0.002	0.001	0.007*	0.002	-0.004	0.013*
	(0.004)	(0.006)	(0.005)	(0.004)	(0.007)	(0.005)	(0.007)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	16,405	16,405	16,405	16,405
R-squared	0.97	0.96	0.96	0.97	0.93	0.97	0.93
Mean Dependent Variable	4.80	3.68	4.34	4.51	3.34	4.48	3.42

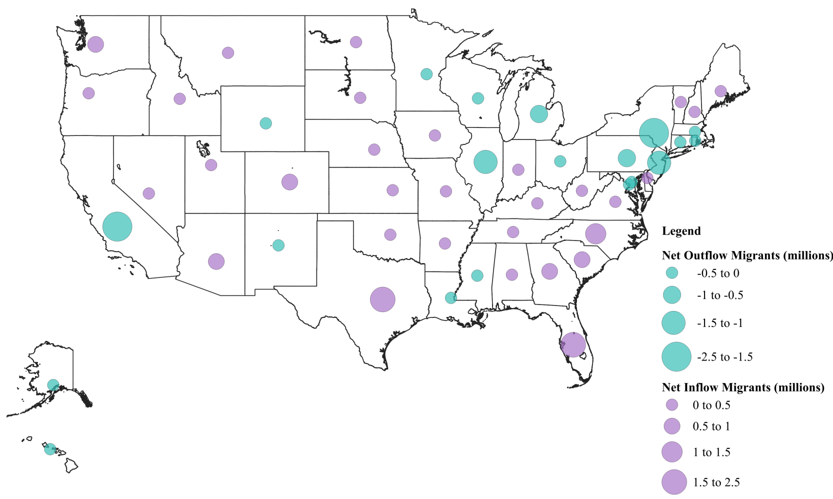
Notes: Unreported macroeconomic control variables include unemployment rate, population, income per capita and GDP at the county-year level. High (low) education refers to migrants with degrees at or above (below) the college level. Employed (unemployed) individuals are classified by their employment status in 1 year before the flood. Young (old) individuals are those under (in or above) the age of 40. Standard errors are clustered at the flood event level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Impact of Floods on Housing Market

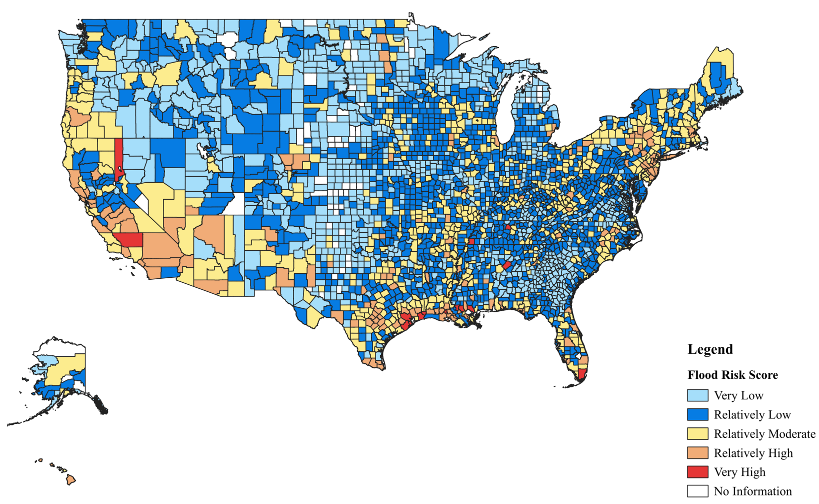
	(1) log(Housing Price)	(2) log(Housing Rent)
Treat × Post	-0.053***	0.074***
	(0.015)	(0.019)
Macroeconomic Controls	Yes	Yes
Flood Event Fixed Effects	Yes	Yes
County Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
State-year Fixed Effects	Yes	Yes
Observations	49,845	20,208
R-squared	0.36	0.26
Mean Dependent Variable	0.48	0.32

Notes: Unreported macroeconomic control variables include unemployment rate, population, income per capita, GDP and total housing supply at the county-year level. Standard errors are clustered at the flood event level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Correlation between Net Migration and Flood Risk in the US



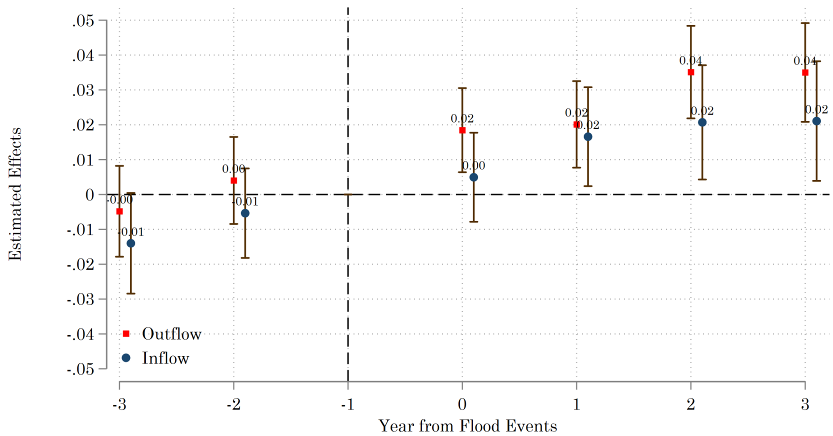
(a) Net Migration



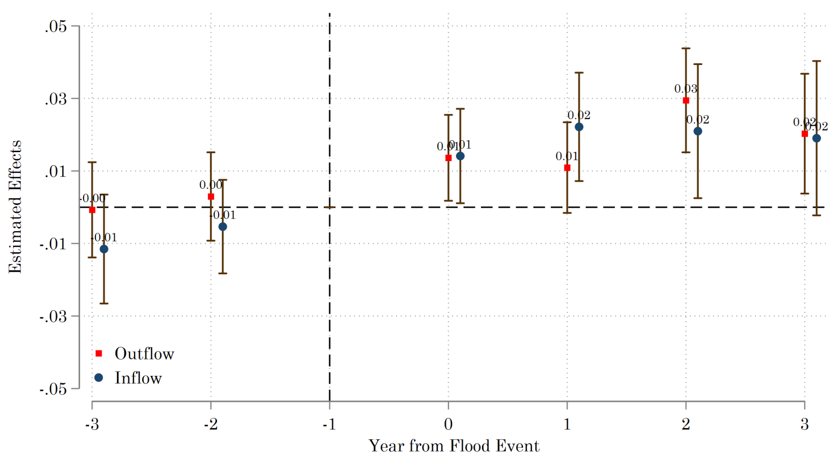
(b) Flood Risk Map

Notes: This figure shows (a) the net migration in the US at the state level between 2006 and 2019 and (b) the average flood risk in the same period at the county level. Data on state-to-state migration flows is obtained from United States Census Bureau; National flood risk index is obtained from Federal Emergency Management Agency.

Figure 2: Event Study Results



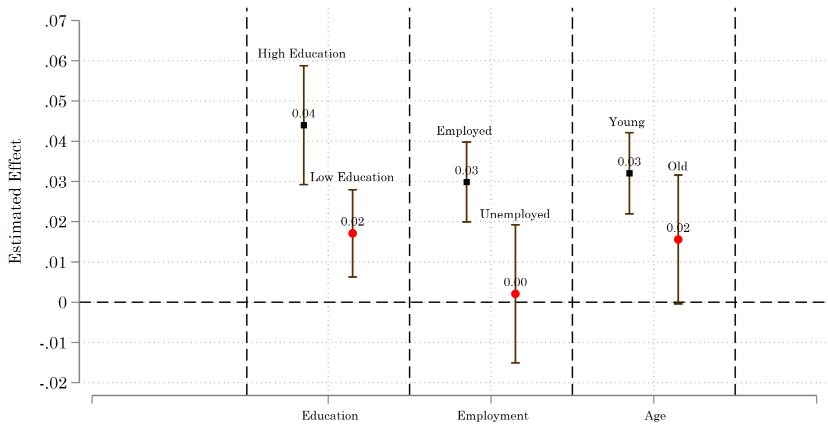
(a) TWFE



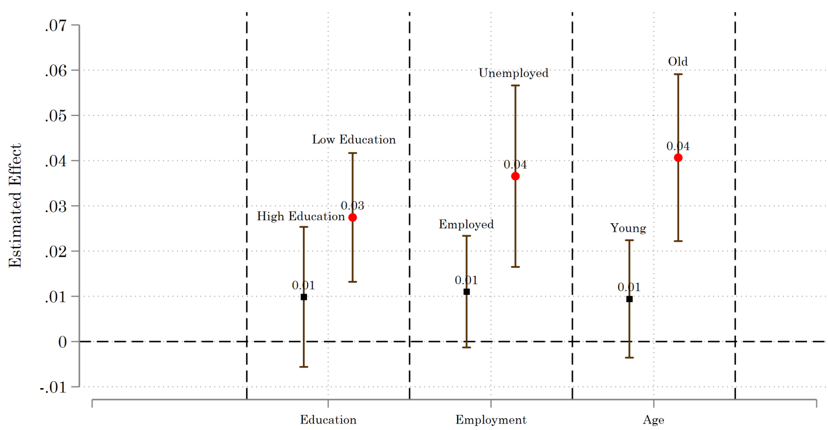
(b) CSDID

Notes: This figure plots the event-study results of the impacts of flood events on outward and inward migration, estimated using (a) the stacked TWFE method and (b) CSDID method. Error bars indicate 90% confidence intervals.

Figure 3: Selective Migration Patterns after Floods



(a) Outflow Migration



(b) Inflow Migration

Notes: The figure plots the flood-induced selective migration patterns in (a) outflow migration and (b) inflow migration. High (low) education refers to migrants with degrees at or above (below) the college level. Employed (unemployed) individuals are classified by their employment status in 1 year before the flood. Young (old) individuals are those under (in or above) the age of 40. Error bars indicate 90% confidence intervals.

Internet Appendix

Table A1: Summary Statistics

Panel A: Number of Outflow Migrants

	Obs	Mean	SD
Total	16,405	169.85	160.71
– Higher-educated (in or above college degree)	16,405	62.16	67.30
– Lower-educated (below college degree)	16,405	106.58	99.55
– Employed	16,405	127.71	120.72
– Unemployed	16,405	42.14	44.51
– Young (≥ 40 years old)	16,405	126.18	119.07
– Old (< 40 years old)	16,405	43.67	43.33

Panel B: Number of Inflow Migrants

	Obs	Mean	SD
Total	16,405	167.09	158.15
– Higher-educated (in or above college degree)	16,405	62.25	71.72
– Lower-educated (below college degree)	16,405	103.68	95.45
– Employed	16,405	126.54	120.43
– Unemployed	16,405	40.55	45.51
– Young (≥ 40 years old)	16,405	124.62	118.87
– Old (≥ 40 years old)	16,405	42.47	43.57

Panel C: Other Control Variables

	Obs	Mean	SD
Sentiment score (standardized)	16,405	-0.21	1.10
Unemployment rate (percent)	16,405	6.38	3.05
Annual personal income (thousand USD)	16,405	44.08	13.10
Population (thousand)	16,405	438.58	461.75
Annual housing price (thousand USD)	16,405	214.66	135.60
County subsidy for flood (thousand USD)	16,405	209.18	1794.51
Number of firms moving in	16,405	0.54	1.58
Number of firms moving out	16,405	0.57	1.76

Panel D: Housing Market Variables

	Obs	Mean	SD
Monthly housing price (thousand USD)	49,845	205.89	122.47
Monthly rent (thousand USD)	20,208	1.29	0.41

Table A2: Impact of Floods on Selective Migration Patterns

Panel A. Outflow Migration

	(1) High Education log(Outflow)	(2) Low Education log(Outflow)	(3) Employed log(Outflow)	(4) Unemployed log(Outflow)	(5) Young log(Outflow)	(6) Old log(Outflow)
Treat × Post	0.044***	0.017***	0.030***	0.002	0.032***	0.016
	(0.009)	(0.007)	(0.006)	(0.010)	(0.006)	(0.010)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	16,405	16,405	16,405
R-squared	0.96	0.96	0.97	0.93	0.97	0.93
Mean Dependent Variable	3.72	4.38	4.54	3.40	4.53	3.47

Panel B. Inflow Migration

	(1) High Education log(Inflow)	(2) Low Education log(Inflow)	(3) Employed log(Inflow)	(4) Unemployed log(Inflow)	(5) Young log(Inflow)	(6) Old log(Inflow)
Treat × Post	0.010	0.027***	0.011	0.037***	0.009	0.041***
	(0.009)	(0.009)	(0.007)	(0.012)	(0.008)	(0.011)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	16,405	16,405	16,405
R-squared	0.96	0.96	0.97	0.93	0.97	0.93
Mean Dependent Variable	3.68	4.34	4.51	3.34	4.48	3.42

Notes: Unreported macroeconomic control variables include unemployment rate, population, income per capita and GDP at the county-year level. High (low) education refers to migrants with degrees at or above (below) the college level. Employed (unemployed) individuals are classified by their employment status in 1 year before the flood. Young (old) individuals are those under (in or above) the age of 40. Standard errors are clustered at the flood event level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Robustness Check for the Baseline Analysis

Panel A. Outflow Migration

	(1) log(Outflow)	(2) log(Outflow)	(3) log(Outflow)	(4) log(Outflow)	(5) log(Outflow)	(6) log(Outflow)
Treat × Post	0.027***	0.026***	0.027***	0.015**	0.033***	0.018***
	(0.005)	(0.005)	(0.009)	(0.006)	(0.006)	(0.006)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	11,361	11,216	15,415
R-squared	0.98	0.98	0.98	0.96	0.98	-
Mean Dependent Variable	4.83	4.83	4.83	4.84	4.97	-

Panel B. Inflow Migration

	(1) log(Inflow)	(2) log(Inflow)	(3) log(Inflow)	(4) log(Inflow)	(5) log(Inflow)	(6) log(Inflow)
Treat × Post	0.019***	0.018**	0.019*	0.014*	0.038***	0.019**
	(0.007)	(0.007)	(0.011)	(0.008)	(0.008)	(0.008)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	11,361	11,216	15,415
R-squared	0.97	0.97	0.97	0.96	0.98	-
Mean Dependent Variable	4.80	4.79	4.80	4.82	4.96	-

Notes: Unreported macroeconomic control variables include unemployment rate, population, income per capita and GDP at the county-year level. Column (1) includes additional control variables of annual housing prices, government subsidies for post-flood recovery, and numbers of public firm entries and exits. Column (2) presents results after removing outliers in outflow/inflow migration at the top 1%. In Column (3), standard errors are changed to be clustered at the county level. Column (4) includes the subsample of treatment counties that frequently experience flood events (i.e., above the median frequency). In Column (5), we remove the control counties within 100km of treatment counties. Column (6) presents the estimates using the CSDID method. Standard errors are clustered at the flood event level (except for Column (3)). *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Robustness Check for the Impact of Media Sentiment on Selective Migration

Panel A. Outflow Migration

	(1) Total log(Outflow)	(2) High Education log(Outflow)	(3) Low Education log(Outflow)	(4) Employed log(Outflow)	(5) Unemployed log(Outflow)	(6) Young log(Outflow)	(7) Old log(Outflow)
Treat × Post × Score	-0.016***	-0.028***	-0.012**	-0.016***	-0.009	-0.018***	-0.013
	(0.005)	(0.007)	(0.006)	(0.005)	(0.008)	(0.005)	(0.008)
Treat × Post	0.022***	0.036***	0.013*	0.025***	0.001	0.027***	0.012
	(0.006)	(0.009)	(0.007)	(0.006)	(0.011)	(0.006)	(0.010)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	16,405	16,405	16,405	16,405
R-squared	0.98	0.96	0.96	0.97	0.93	0.97	0.93
Mean Dependent Variable	4.83	3.72	4.38	4.54	3.40	4.53	3.47

Panel B. Inflow Migration

	(1) Total log(Inflow)	(2) High Education log(Inflow)	(3) Low Education log(Inflow)	(4) Employed log(Inflow)	(5) Unemployed log(Inflow)	(6) Young log(Inflow)	(7) Old log(Inflow)
Treat × Post × Score	0.010*	-0.000	0.017**	0.004	0.025***	0.011*	0.016*
	(0.006)	(0.007)	(0.007)	(0.006)	(0.009)	(0.006)	(0.009)
Treat × Post	0.021***	0.009	0.031***	0.011	0.043***	0.012	0.043***
	(0.007)	(0.010)	(0.009)	(0.008)	(0.012)	(0.008)	(0.012)
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Flood Event Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,405	16,405	16,405	16,405	16,405	16,405	16,405
R-squared	0.97	0.96	0.96	0.97	0.93	0.97	0.93
Mean Dependent Variable	4.80	3.68	4.34	4.51	3.34	4.48	3.42

Notes: Unreported macroeconomic control variables include unemployment rate, population, income per capita and GDP at the county-year level. High (low) education refers to migrants with degrees at or above (below) the college level. Employed (unemployed) individuals are classified by their employment status in 1 year before the flood. Young (old) individuals are those under (in or above) the age of 40. Standard errors are clustered at the flood event level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Aggregate Changes in Income due to Selective Migration by Education

Panel A. Outflow Migration

	High Education			Low Education		
	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI
Change (%) in Migrants	2.64%	4.40%	6.16%	0.42%	1.70%	3.00%
× Average Migrant Population	6,216	6,216	6,216	10,658	10,658	10,658
Change (Number) in Migrants	164	274	383	45	181	320
× Average Income	51,806	51,806	51,806	15,962	15,962	15,962
Change in Aggregate Income	8,501,503	14,169,000	19,832,750	713,458	2,892,070	5,108,101

Panel B. Inflow Migration

	High Education			Low Education		
	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI
Change (%) in Migrants	-0.86%	1.00%	2.83%	1.05%	2.70%	4.44%
× Average Migrant Population	6,225	6,225	6,225	10,368	10,368	10,368
Change (Number) in Migrants	-53	62	176	109	280	460
× Average Income	51,332	51,332	51,332	16,154	16,154	16,154
Change in Aggregate Income	-2,739,790	3,195,430	9,047,500	1,753,823	4,521,970	7,437,860

Notes: This table presents a back-of-envelop estimation of aggregate changes in income due to flood-induced selective migration by education. High (low) education refers to migrants with degrees at or above (below) the college level. Panels A and B calculate the changes in aggregate income of outflow and inflow migration due to flood events, respectively. The average population of high-education and low-education migrants are equal to the corresponding mean values in Table A1, adjusted by the sampling rate of ACS (1%). The percentage changes in migrants are from Columns (1) and (2) in Table A2. The average income is obtained from the ACS dataset.

Table A6: Aggregate Changes in Income due to Selective Migration by Age

Panel A. Outflow Migration

	Young			Old		
	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI
Change (%) in Migrants	2.00%	3.20%	4.41%	-0.35%	1.60%	3.47%
× Average Migrant Population	16,218	16,218	16,218	4,367	4,367	4,367
Change (Number) in Migrants	325	519	715	-15	70	151
× Average Income	23,740	23,740	23,740	46,087	46,087	46,087
Change in Aggregate Income	7,712,487	12,320,600	16,964,170	-705,145	3,220,210	6,978,628

Panel B. Inflow Migration

	Young			Old		
	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI	Lower Bound 95% CI	Mean Point Estimate	Upper Bound 95% CI
Change (%) in Migrants	-0.61%	0.90%	2.49%	1.87%	4.10%	6.26%
× Average Migrant Population	12,462	12,462	12,462	4,247	4,247	4,247
Change (Number) in Migrants	-75	112	310	79	174	266
× Average Income	24,144	24,144	24,144	46,593	46,593	46,593
Change in Aggregate Income	-1,821,140	2,707,920	7,488,212	3,693,578	8,113,100	12,395,670

Notes: This table presents a back-of-envelop estimation of aggregate changes in income due to flood-induced selective migration by age. Panels A and B calculate the changes in aggregate income of outflow and inflow migration due to flood events, respectively. The average population of young and old migrants are equal to the corresponding mean values in Table A1, adjusted by the sampling rate of ACS (1%). The percentage changes in migrants are from Columns (5) and (6) in Table A2. The average income is obtained from the ACS dataset.