RESEARCH

Adapting to Climate Risk? Local Population Dynamics in the United States

Agustín Indaco[1] · Francesc Ortega[2]

Abstract

Using a new composite climate-risk index, we show that population in high-risk counties has grown disproportionately over the last few decades, even relative to the corresponding commuting zone. We also find that the agglomeration is largely driven by increases in the (white) working-age population. In addition, we show that high-risk tracts have typically grown more than low-risk tracts *within* the same county, suggesting the presence of highly localized amenities. We also document heterogeneous population dynamics by degree of urbanization, region and type of natural hazard. Specifically, population has been retreating from high-risk, low-urbanization locations, but continues to grow in high-risk areas with high residential capital. Net migration flows have contributed to the higher growth of high-risk counties in the South and Northeast of the country, but the opposite has happened in the West and Midwest. Last, we provide evidence of *microretreat* in the case of *coastal flooding*: tracts with high levels of this risk have grown significantly *less* than other tracts in the same county, suggesting that residents are willing to relocate within short distances to avoid predictably risky locations.

Keywords Climate risk · Agglomeration · Migration

Introduction

Over the last decades, the frequency and intensity of natural hazards in the United States (U.S.) has increased. According to the National Oceanic and Atmospheric Administration (NOAA), the U.S. experienced more than twice the number of billion-dollar disasters during

✉ Agustín Indaco
   aindaco@andrew.cmu.edu

   Francesc Ortega
   francesc.ortega@qc.cuny.edu

[1] Carnegie Mellon University in Qatar, Doha, Qatar

[2] Dina A. Perry Professor of Economics, Queens College, CUNY, New York, USA

2010-2020 than it did in the previous decade and, in fact, four of the five most costly natural disasters have occurred since 2010.[1] We illustrate this point in Fig. 1 using SHELDUS data (CEMHS, 2022). The chart identifies the most damaging event in each year between 1960 and 2020 on the basis of inflation-adjusted cost, separately for each of the main types of natural hazards. During the 1960s, the most damaging events were relatively benign (with costs mostly in the first two quintiles of the distribution). However, over the following decades, the most damaging events have become much more costly, with an increasing presence in the fourth and fifth cost quintiles.

In addition to the increased frequency of extreme natural disasters, the increase in damage over the last few decades appears to be intimately related to the increasing agglomeration of people and economic activity in high-risk areas.[2] Despite a few notable exceptions where hurricanes led to a persistent reduction in local population (Deryugina et al., 2018), there seems to be a general trend toward population agglomeration in hurricane-prone areas.[3] Previous studies have shown that, for several decades, coastal counties in the U.S. have grown disproportionately, including many counties that have been hit by large hurricanes over this period of time (Wilson and Fischetti, 2010; Lin et al., 2021).[4] Similar findings have been found regarding the pace of new construction in places with a high risk of wildfires (Radeloff et al., 2018) and heatwaves (Partridge et al., 2017). However, to the best of our knowledge, no study has done a comprehensive analysis of population dynamics that considers all major natural hazards, which also includes droughts, riverine flooding, tornados, hail, and so on.[5]

Our goal is to investigate population dynamics in areas that *currently* exhibit high climate risk, with a focus on examining whether population retreat is taking place or, rather, local population dynamics continue evolving along long-term trends. To do this, we introduce a novel composite measure of current climate risk (based on the historical frequency of climactic events) and merge it with population data at the county and sub-county levels over the last century. A one-dimensional measure of climate risk that incorporates all climate hazards is a useful construct. It provides a simple measure of the average climate risk associated with the distribution of population (or economic activity) at the desired level of geography, which may also be useful for the calibration of structural models featuring a large number of geographic units (e.g. Pang and Sun (2022)). It is also worth highlighting that our construction of an aggregate climate risk measure can be easily adapted to build analogous measures at lower levels of aggregation.[6]

---

[1] Hurricanes Sandy (2012), Harvey (2017), Irma (2017), and Maria (2017).

[2] Across the world, infrastructure investment in flood-prone coastal areas continues to rise, often ignoring sea-level rise projections (Balboni, 2021). Similarly, public expenditures on wildfire protection subsidize development in places with fire hazards (Baylis and Boomhower, 2023).

[3] Deryugina et al. (2018) analyzed the effects of hurricane Katrina on the population of New Orleans, along with the effects on their employment and income. They find a persistent reduction in population but only small and highly short-lived effects on labor market outcomes. Specifically, eight years after the storm, over a third of the displaced population had not returned to New Orleans.

[4] Over one third of the U.S. population lives in coastal counties. According to Wilson and Fischetti (2010), between 1960 and 2008, the share of population living in coastal counties along the Gulf of Mexico soared by 150%, more than double the national average.

[5] Jia et al. (2023) provide a recent review of the economic determinants of internal migration, including the role played by climate risk.

[6] For instance, one may want to create climate risk measures that aggregate all flooding-related hazards (coastal flooding, riverine flooding and hurricanes) or all heat-related hazards (heat waves and droughts).

Fig. 1 The costliest climate events between 1960 and 2020. **Notes:** The figure plots the costliest climate event, by type of hazard, in each year. Specifically, each year we identify the most damaging event for each of the 6 natural hazards considered. We then color-code it on the basis of its inflation-adjusted monetary cost. Green and yellow squares correspond to quintiles 1 and 2, respectively, of the all-time distribution of (inflation-adjusted) damage costs for each natural hazard. Orange and red squares correspond to quintiles 4 and 5 of the same distribution. The data source is SHELDUS (CEMHS, 2022)

Since climate risk discussions gained saliency during the 1990s, we are primarily interested in population dynamics during the period 1990-2020 .[7] However, we have assembled county-level population counts going back to 1900 in order to characterize long-term local population dynamics long before climate risk became a potentially relevant factor shaping mobility decisions. Equipped with our composite measure of climate risk, we estimate simple econometric models for the change in log population over time, which differences out all time-invariant local characteristics. These models allow us to estimate the gap in population growth between counties with currently high (or medium) climate risk and counties with low risk over a long period of time. We also examine whether *within-county* population dynamics mitigate or exacerbate cross-county population shifts. We use these estimates to test whether population is retreating from counties or census tracts with relatively high climate risk.

Our analysis delivers several findings. First, we find that in the last three decades, high-risk counties have grown about 2.9 log points more, per decade, than low-risk counties. Even after netting out the average growth in the commuting zone (which is typically considered a good approximation to the geographical scope of local economies), high-risk counties have grown disproportionately more than low-risk ones over the last few decades (with an excess of 0.5 log points per decade). These results suggest the presence of amenities in high climate-risk areas that operate at the county or sub-county levels (as opposed to county-level attributes or the gravitational pull of local economies). Additionally, we show that high-risk tracts typically grow *more* than low-risk tracts within the same county, which exacerbates the increase in climate exposure implied by the county-level analysis.

Our results also highlight that the effects of climate risk on population growth vary across several dimensions. We have found stark differences in the geographic sorting of different socio-demographic groups. More specifically, the increasing population agglomeration in high climate-risk counties appears to be largely driven by white, working-age individuals. Retirement-age and (less affluent) non-white populations appear to be retreating from counties with high climate risk.

We also documented differential local trends on the basis of the degree of urbanization. Specifically, we find population retreat from high-risk, low-urbanization locations, but increasing population agglomeration in high-risk, high-urbanization locations. We also find that in the South and Northeast of the country, the gap in population growth has been fueled by net migration into high-risk counties. In contrast, in the Midwest and West, over the last 3 decades, net migration flows are responsible for lowering the population growth in high-risk counties below the rate of growth for low-risk counties in the same region.

Lastly, we uncover evidence of *micro-retreat* in response to risk of coastal flooding. Namely, we show that tracts with high risk of coastal flooding grew less than other tracts in

---

[7] The United Nations officially recognized climate change as a serious global problem in 1992, with the Rio Earth Summit.

the same county. However, we do not find this pattern for other natural hazards. We argue that this is because coastal flooding is an easily predictable, highly localized risk, which allows residents to "insure" themselves by relocating to low-risk tracts while remaining in the same county.

All in all, our findings show increasing agglomeration in high climate-risk areas in the South and Northeast of the United States, likely driven by robust local economies. However, the rate of excess growth in high-risk areas at the national level seems to be decreasing since 1990. This reversal is due to changing demographic trends in the West and Midwest, where net migration flows have recently lowered the rate of population growth in high climate-risk counties below the rate of growth in low climate-risk counties in the same region.

Literature

The literature on climate risk and population dynamics is growing rapidly. Many studies have focused on the effect of extreme weather events and natural hazards on migration. Boustan et al. (2020) analyze the effect of a wide range of natural disasters on net-migration over the period 1920-2010 and find that severe disasters such as wildfires and hurricanes tend to trigger county-level out-migration. However, they find that flooding episodes tend to *attract* migrants.

The demographic effects of climatic events are also a function of population density and pre-existing demographic trends. For example, Fussell et al. (2017) document that hurricanes and tropical storms lower population growth only for the small subset of U.S. counties with high-density *and* growing populations, which only represent 2% of all US counties. This finding leads them to conclude that long-term local population trends overshadow the effects of episodic weather events. Other studies have also suggested that the effects of flooding on migration are heterogeneous in household and regional characteristics (as in the review by Hauer et al. (2020)).

Interestingly, other papers have studied the information content of natural hazards and whether residents in those affected areas do indeed update beliefs. For example, Petkov (2022) studies whether unexpected hurricanes lead to belief updating by locals and lead to larger population loss relative to more predictable hurricanes. His analysis shows that population growth declines more in counties that had not suffered hurricanes in the past, suggesting belief updating by residents exposed to large-scale climatic events for the first time.[8] Additional evidence in support of residents' belief updating in response to first-time experience of severe flooding is provided in Petkov and Ortega (2023). These authors analyze flood insurance take-up in the aftermath of a large hurricane in New York and show persistent increases in take-up among homeowners (located just outside the 100-year flood zone) that were likely exposed to severe flooding for the first time.

Our work is more closely related to studies examining local population dynamics on the basis of climate risk, rather than the effects of episodic climate events. Lin et al. (2021) document that, between 1990 and 2010, new residential construction in the Gulf of Mexico and Northeast regions of the U.S. was concentrated in high-density areas (Census blocks) with high projected risk of coastal flooding. The authors argue that urban agglomeration economies still overpower the risk associated with sea-level rise. Compared to their paper, our analysis includes both earlier data (going back to 1920) and more recent data (for 2020).

---

[8] Petkov (2022) also reports that unexpected hurricanes increase housing prices, but other studies find the opposite effect on housing values (e.g. Ortega and Tapınar (2018); Indaco et al. (2021)).

We also go beyond the analysis of coastal flooding risk and consider a wide range of climate hazards.

In the context of wildfire risk, Fussell et al. (2017) study the number of housing units built in the wildland-urban interface, an area prone to wildfires. The authors find that between 1990 and 2015, construction in the wildland-urban interface was the fastest-growing land use type in the United States. Similar trends have been found in regions at risk of droughts and heat waves. For instance, Partridge et al. (2017) document that, in the second half of the 20th century, Americans moved to locations that are predicted to experience severe heat waves and long-term droughts.

Social scientists have also used observed migration patterns and current climate projections to simulate future climate migration scenarios. These models make predictions of the demographic effects of climate change. Some studies emphasize that economically vulnerable populations may not be able to afford retreating to low-risk locations and may be *trapped* in high-risk locations (Black et al., 2011; Hauer et al., 2020, 2022). On their part, Black et al. (2011) point out that migration is already an important coping strategy in several countries, as is the case in Bangladesh in response to large-scale flooding episodes. They also predict that environmental factors will play an increasingly larger role in shaping international migration in many other areas of the world.

Other authors have focused on the impact on the geographical distribution of economic activity. Using a dynamic model of the world economy, Desmet et al. (2021) simulate the effects of sea-level rise on firms' location decisions, taking into account the effects of local agglomeration economies. Based on conservative sea-level rise projections, they estimate that by 2050 about 0.2% of the world's population (and firms) will have been displaced (reaching 1.5% in year 2100).[9] Interestingly, welfare losses are estimated to be larger than real GDP losses because the population endogenously retreats toward (non-coastal) areas with worse amenities. Importantly, their analysis implies highly heterogeneous geographical effects. For instance, while the U.S. as a whole is predicted to experience only a negligible reduction in real GDP (peaking at 0.01%), coastal areas in South Florida and Texas (and to a lesser extent in the Northeast) are predicted to suffer much larger output and population losses, which are offset by gains in neighboring inland locations.

Data Sources

Population by County

We use the *Surveillance, Epidemiology, and End Results Program* (SEER) dataset compiled by the National Cancer Institute. This dataset spans 1969-2020 and breaks down county population by 19 age groups, race (3 groups) and gender. We impose a few data restrictions: we drop Alaska and Hawaii due to the difficulty of linking counties over time for these states, and a few groupings of counties that were only used in the 1970 Census (FIPS 36910, New York City). As explained in detail in Appendix A, we used linear interpolation to impute population values for a handful of counties for years 1970 and 1980.

We extend the SEER dataset in two ways. First, we extend it backward by merging historical Census estimates for county population (overall) for the period 1900-1970. These data allow us to trace the evolution of population for the vast majority of counties for over a

---

[9] Their estimates only consider the costs associated with locations that will become permanently flooded and do not take into account the increased frequency of flooding episodes in other coastal areas.

century (1900 to 2020). For years prior to 1970 we use the county-level data as is. In addition, we also make use of the county-level dataset in Egan-Robertson et al. (2023), which provides county-level estimates for net migration for every decade from the 1960s to the 2010s. This data allow us to separate out natural population growth from growth driven by net migration into the county.

As shown in Table 1, population counts obtained aggregating our county dataset are fairly accurate.[10] As seen in Fig. 2, between 1920 and 2020 the country's population increased by about 220 million, corresponding to an average decadal growth of 11.3 log points (Table 1, column 3). Population growth has slowed down since 1970, averaging 10.4 log points per decade. Interestingly, the elderly and non-white populations have grown at much higher rates than the rest of the population over the period 1970-2020. Over this 50-year period, the population age 65 and above and the non-white population grew by an average of 21.3 and 23.3 log points per decade, respectively, more than twice the rate for the overall population. The higher growth rate among the elderly population in the last 50 years reflects both the aging of baby-boomers and the steady increase in longevity. The higher growth rates for non-whites might reflect the increase in immigration (from abroad) since the 1965 changes to US immigration policy (*Immigration and Nationality Act*), which opened the door to several decades of high immigration.

The top half of Table 2 presents summary statistics for the county data. The first set of variables reports the average decadal population growth (change in log population divided by the number of decades). Over the two last decades, population in the average county has grown by an average of 2.5 log points per decade (and solely 0.6 log points in the 2010s) but, obviously, there's a great deal of variation (ranging from a 31 log point reduction to a 51 log point increase). The table also shows that population growth has slowed down considerably. Between 1920 and 2020 the average population growth in the average county was 5.2 log points per decade, more than twice the value for the 2000-2020 period.

Population by Census Tract

The *Longitudinal Tract Data Base Census Dataset* (LTBD) provides Census-tract population data for the period 1970-2020. It combines data from the decennial Census and the ACS and, crucially, the tract boundaries have been harmonized to 2010 Census tract boundaries as described in Logan et al. (2014). We use the full-count (standard) dataset.[11]

The bottom half of Table 2 (and Fig. 9) presents summary statistics for the population data at the Census tract level. The first set of variables reports the average decadal population growth (change in log population divided by the number of decades). Over the last two decades, population in the average tract has grown by an average of 6.7 log points per decade. As expected, the variation in population growth across tracts is large, with population falling by 285 log points in some tracts and increasing by 804 log points in others. As shown before, population growth at the tract level has slowed down considerably. Between 1970 and 2020 the average population growth was 16.4 log points per decade, more than two times larger than the value for the period 2000-2020.

---

[10] According to the BLS, the U.S. population in 2020 was 331.4 (April 1 estimate). The SEER data report 327.3 million people. The 4-million disparity is due to the exclusion of Alaska, Hawaii and Puerto Rico.

[11] The data is freely available at https://s4.ad.brown.edu/Projects/Diversity/Researcher/Bridging.htm. See Appendix B for more details.

Table 1 Nationwide Trends

Demographic year	(1) All Pop(Mn)	(2) All 100*Dln	(3) All 100*AvDln	(4) Age65plus 100*AvDln	(5) Nonwhites 100*AvDln	(6) Med-Low All 100*AvDln	(7) High-Low All 100*AvDln	(8) High-Low Age65plus 100*AvDln	(9) High-Low Nonwhites 100*AvDln
1900	75.3	19.7	12.2			1.7	6.3		
1910	91.7	13.9	11.6			2.2	6.2		
1920	105.4	14.9	11.3			2.4	6.3		
1930	122.3	6.9	10.9			2.3	5.7		
1940	131.1	13.3	11.4			2.5	6.1		
1950	149.7	16.7	11.2			2.6	5.7		
1960	176.9	9.6	10.3			2.6	4.9		
1970	194.7	15.0	10.4	21.3	23.3	2.7	5.9	4.2	6.3
1980	226.2	9.2	9.2	19.2	21.2	3.1	5.0	2.3	3.2
1990	248.0	12.3	9.3	19.2	20.7	2.8	3.7	2.4	1.5
2000	280.3	9.2	7.8	23.0	18.5	2.7	3.8	3.1	1.8
2010	307.2	6.3	6.3	31.8	16.0	3.3	4.7	4.8	2.5
2020	327.3								
Average									
1900-2020	187.4	12.2							
1970-2020	264.0	10.4							

Notes: Column 1 is the population in millions, obtained by aggregating the county data. Column 2 reports the decadal change in log population (times 100), where each value corresponds to the following 10 years. Columns 3-10 report the average decadal growth (times 100) between the corresponding initial year and year 2020. Columns 6-9 net out the average growth rate for the low-risk category
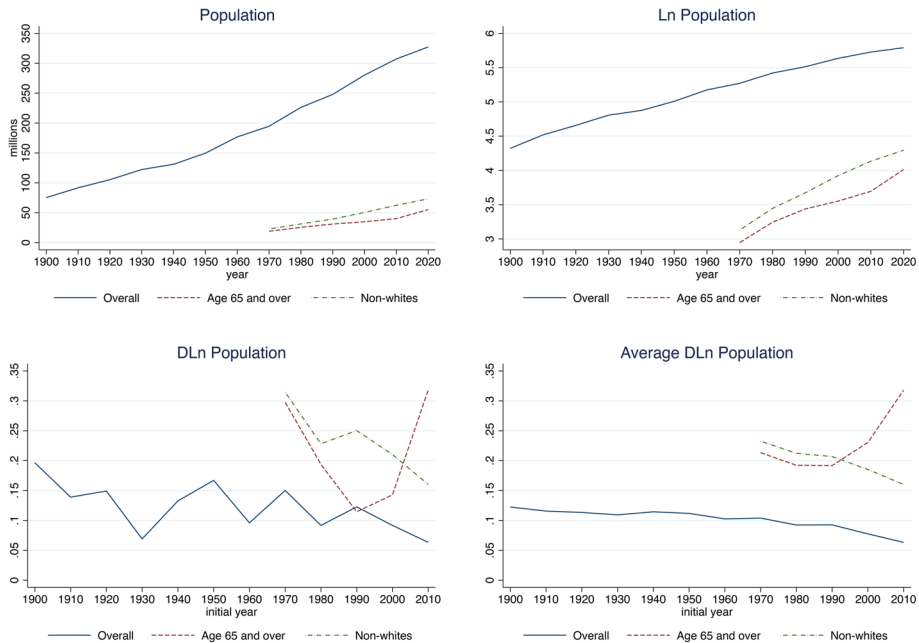
Fig. 2 Nationwide trends. All regions pooled. **Notes:** Population trends in United States. Top-left figure is in millions of individuals, top-right figure in logs, bottom-left figure is the decade-over-decade change in log population and the bottom-right figure is the average decadal growth between the year indicated in the horizontal axis and year 2020

Our Composite Climate Risk Index

We use natural hazard risk metrics provided by FEMA (November 2021 version).[12] Our starting point is the most comprehensive metric, which includes data for a large number of natural hazards and is a function of both the expected annual losses from each of the 18 hazards in each geographic area, and the area's social vulnerability and community resilience. This index combines information on 18 natural hazards, and takes values that range from 0 to 100.[13]

It is important to note that expected annual losses are a combination of the expected annual frequency of the climate events and the degree of exposure, which is a function of the area's population and its housing stock. Therefore there will be a mechanical correlation between this risk metric and population, both in levels and growth rates.

Given our interest in examining how climate risk impacts population growth, it is more appropriate to measure climate risk solely on the basis of *annual frequency*. Annualized frequency for each hazard is calculated as the number of historical occurrences (in counts of events or event-days) over the length of the time period, using a variety of primary sources that vary across each of the 18 specific hazards. The methodology to produce these estimates

---

[12] The data can be freely downloaded at https://hazards.fema.gov/nri.

[13] The natural hazards are: Avalanche, Coastal Flooding, Cold Wave, Drought, Earthquake, Hail, Heat Wave, Hurricane, Ice Storm, Landslide, Lightning, Riverine Flooding, Strong Wind, Tornado, Tsunami, Volcanic Activity, Wildfire and Winter Weather. We ignore volcanic activity and earthquakes, which are not directly related to climate.

Table 2 Descriptive Statistics (counties and tracts)

Variable Counties	Obs	Mean	Std. Dev.	Min	Max
(Dln 2020-2010) / 1	3,107	0.006	0.089	-0.401	0.866
(Dln 2020-2000) / 2	3,106	0.025	0.094	-0.308	0.512
(Dln 2020-1990) / 3	3,106	0.048	0.100	-0.242	0.589
(Dln 2020-1980) / 4	3,104	0.043	0.105	-0.270	0.661
(Dln 2020-1970) / 5	3,091	0.062	0.110	-0.219	0.747
(Dln 2020-1960) / 6	3,088	0.058	0.112	-0.240	0.719
(Dln 2020-1950) / 7	3,084	0.055	0.113	-0.252	0.662
(Dln 2020-1940) / 8	3,080	0.051	0.112	-0.241	0.619
(Dln 2020-1930) / 9	3,080	0.051	0.107	-0.215	0.623
(Dln 2020-1920) / 10	3,049	0.052	0.104	-0.189	0.617
FEMA NRI Risk Score	3,104	10.628	6.759	0.000	100.000
Z Risk Composite	3,114	0.000	0.257	-0.844	1.757
ZW Risk Composite	3,114	0.000	0.404	-0.917	2.031
Low ZW Risk	3,114	0.250	0.433	0.000	1.000
Med ZW Risk	3,114	0.500	0.500	0.000	1.000
High ZW Risk	3,114	0.250	0.433	0.000	1.000
Tracts					
(Dln 2020-2010) / 1	58,458	0.048	0.178	-3.924	4.369
(Dln 2020-2000) / 2	58,440	0.067	0.250	-2.853	8.041
(Dln 2020-1990) / 3	58,438	0.099	0.244	-2.111	6.363
(Dln 2020-1980) / 4	46,188	0.141	0.320	-1.637	4.939
(Dln 2020-1970) / 5	40,767	0.164	0.330	-1.336	4.564
FEMA NRI Risk Score	58,488	16.177	6.935	0.000	87.087
Z Risk Composite	58,488	0.000	0.232	-0.696	6.077
ZW Risk Composite	58,488	0.000	0.418	-0.764	4.766
Low ZW Risk	58,488	0.250	0.433	0.000	1.000
Med ZW Risk	58,488	0.500	0.500	0.000	1.000
High ZW Risk	58,488	0.250	0.433	0.000	1.000

Notes: Unweighted summary statistics. *Z risk score* is based on FEMA's annual frequency of climate events. We standardize the frequency of each event and compute a simple average (Z) and a weighted average (ZW). We define the 3 categories of composite climate risk (low, medium and high) as follows: below the 25th percentile, between 25th and 75th percentiles, or above the 75th percentile, respectively

differ somewhat for each hazard, depending on the nature of the hazard. In most cases, the frequency of hazards is recorded at the Census block level. Once the total number of recorded hazards is obtained, the annualized frequency is simply calculated as the number of recorded hazard occurrences within the recording period divided by the corresponding number of years. Once these measures are obtained at the Census block level, area-weighted aggregates are computed in order to obtain frequencies at the Census tract and county levels. Appendix C includes detailed information on the definition of occurrences and the calculation of annualized frequencies for each of the main hazard types.

Our frequency-based composite risk measure is built as follows. First, we standardize the annual frequency for each hazard (using the corresponding mean and standard deviation). Next, we average the standardized annual frequencies using hazard-specific weights, and denote the weighted composite index by ZW. The weights are meant to capture the disparity in the *economic consequences* of each hazard. Specifically, we compute each natural hazard's share in the expected annual loss due to property (buildings) damage and crop losses nationwide.[14] Because the main natural hazards in our composite risk index are geographically widespread across many counties and Census tracts, using their national dollar losses is unlikely to contaminate our county-level composite risk index. We will also examine the robustness of our results to the use of the weights in the calculation of our composite in two ways: by estimating our main models using the unweighted version of our composite index (which assigns equal weights to all hazards) and by repeating the analysis for each natural hazard separately (i.e. without combining them into a scalar index).

Last, we compute the 25th and 75th percentiles of the composite (weighted) index and classify a county as *low risk* if the composite annual frequency measure is below the 25th percentile, *medium risk* if it falls between the 25th and 75th percentiles, and *high risk* if it is above the 75th percentile.[15]

Housing Values and Residential Capital Data

In Heterogeneous Effects by Residential Capital, we will analyze local population growth (for the period 1990-2020) on the basis of county-level climate risk and economic density. The latter will be based on residential capital values obtained from the 2000 Census.[16] Specifically, we will partition all U.S. counties on the basis of below or above median values for each of the following three measures: (i) overall value of residential capital in the county, (ii) median value of homes in the county, and (iii) value of residential capital over the county's surface area.

Definitions and Nationwide Trends

This section defines our measures of population growth and examines both nationwide trends and the geographic distribution of climate risk. These exercises both provide an overview of the data and help assess their quality.

---

[14] For droughts the expected annual losses for building damages are not part of the dataset. We replaced them with expected annual agricultural losses. Ten out of the 16 hazards considered account for the vast majority of the nationwide economic damage caused by climate events and, in fact, the main 7 hazards account for 94% of all economic damage. We list them next in decreasing order, along with the corresponding shares: hurricanes (0.21), droughts (0.21), riverine flooding (0.18), tornados (0.13), wildfires (0.10), hail (0.06), coastal flooding (0.05), strong winds (0.04), ice storms (0.01) and winter weather (0.01).

[15] It is worth noting that the annual frequency distributions differ in the county and Census tract datasets, which delivers different threshold values for the risk categories.

[16] Given that our period of interest is 1990-2020, using pre-determined 1990 (or earlier) residential capital values would have been clearly better. Unfortunately, we were not able to obtain such data with the required geographic coverage. Nonetheless, the high persistence of housing stocks, and their value, over a 10-year period suggests that the results will probably not be affected much. We conduct some auxiliary analysis in Heterogeneous Effects by Residential Capital that shows this is a plausible assumption.

Population Growth

We begin by pooling all counties together and examining the evolution of population over time for the US as a whole. Figure 2 plots the evolution of population in levels (top-left) and in logs (top-right).[17] The top figures plot the evolution of the US population in levels and logs. The bottom left figure plots the *decadal population growth rates* for the 10-years beginning in year t. Namely,

$$g_t = ln\, Pop_{t+10} - ln\, Pop_t, \text{ for } t = 1900, 1910, ..., 2010. \qquad (1)$$

As illustrated in Fig. 2 (bottom-left figure), there is a downward trend in decadal growth rates, but there is also substantial variability, partly reflecting economic conditions. Specifically, decadal population growth was the lowest in the 1930s and 2010s, with 6.9 and 6.3 log points, respectively.

To smooth out fluctuations, it is helpful to define the *average decadal growth rate* over periods of time ranging from *initial year t* and final year 2020, or

$$\overline{g}_t = \frac{ln\, Pop_{2020} - ln\, Pop_t}{0.1 \times (2020 - t)}, \text{ for } t = 1900, 1910, ..., 2010, \qquad (2)$$

where the denominator simply counts the number of decades between initial year t and year 2020. A little algebra easily shows that $\overline{g}_t$ is simply the average of the decadal population growth rates (g_t) for the corresponding decades (beginning with years $\tau = t, ..., 2010$). Note also that $\overline{g}_{2010} = g_{2010}$.

The bottom-right figure in Fig. 2 clearly shows the downward trend in the growth rate for the overall population. Between the years 1900 and 2020, the average population growth rate has been around 12% per decade. In comparison, the corresponding rate fell to 10% for the 1970-2020 period and fell further to roughly 6% for the 2010-2020 decade.[18] To a large extent this trend reflects the reduction in fertility rates accompanying the secular increase in per-capita income. Despite large improvements in life expectancy and periods of high immigration, population growth has trended downward between 1900 and 2020.

Population Growth and Climate Risk

It is helpful to consider our county-level population datasets and partition all counties into 3 groups on the basis of our composite climate risk index (ZW). Specifically, we consider the three climate risk levels (indexed by r) defined in : low ($r = 0$), medium ($r = 1$) and high risk ($r = 2$). We then classify all counties by their composite risk category and *pool* all counties with the same risk category. Last, we compare the evolution of population across the three risk categories. In particular, we are interested in assessing whether population growth has been *lower* in high-risk areas, which would indicate population retreat.

We examine the trends in terms of the average decadal (10-year) growth rates $\{\overline{g}_t^r - \overline{g}_t^0\}$, for $r = 1, 2$. As can be seen in Fig. 3 (bottom right), up until the 1970s, the average growth differential between high-risk and low-risk areas was high and relatively stable, roughly 6 percentage points per decade. Since then, the gap in growth rates appears to have fallen gradually: over the last 20 years the average growth rate has been about 4 percentage-points

---

[17] The figures also plot the evolution of two demographic groups: the population age 65 and over and the non-white population. Both groups have grown rapidly, relative to the overall population, since 1970.

[18] The figures also plot the data for the population age 65 and over and the non-white population. Since 1970 these groups have increased at much higher rates, on average, than the overall population.
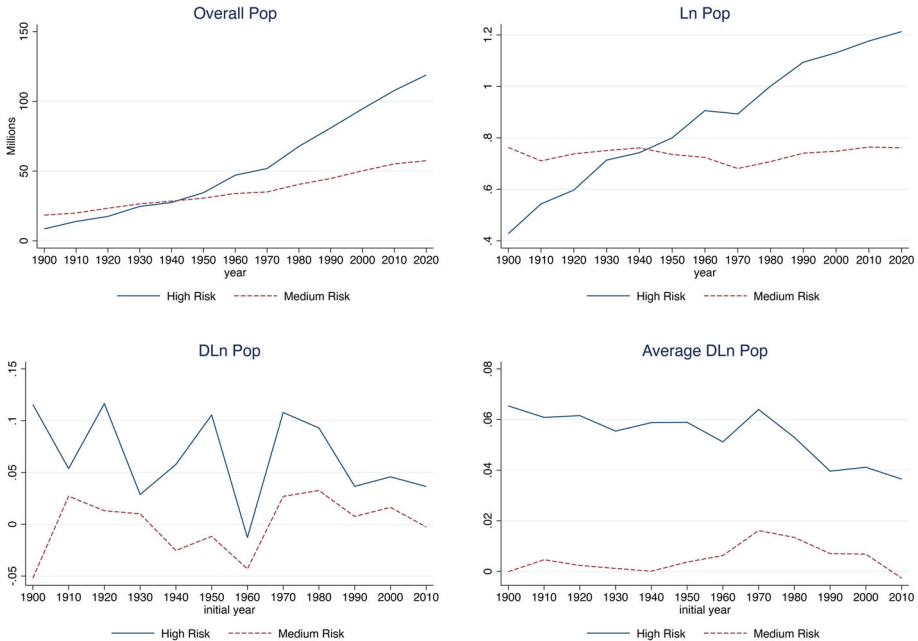
Fig. 3 National trends. By climate risk (weighted composite). **Notes:** Composite risk with hazard weights. Population trends in United States by climate risk, net of the corresponding value for the low-risk category. Top-left figure is in millions of individuals, top-right figure in logs, bottom-left figure is the 10-year log change between the year indicated in the horizontal axis and 10 years later, and the bottom-right figure is the average decadal growth between the year indicated in the horizontal axis and year 2020

higher in high-risk areas than in low-risk ones. In comparison, medium-risk areas have grown at similar rates as low-risk areas, except for the 1970-2000, period when medium-risk areas grew at somewhat higher rates than low-risk areas.

In conclusion, the data indicate that population growth remains much higher in high-risk areas than in low-risk areas, even though the gap appears to have been closing slowly in the last few decades. In other words, nationwide population is *not* retreating from high-risk counties. Rather, these counties continue to grow disproportionately, albeit at a decreasing rate.

The Geographic Distribution of Climate Risk

To understand the geographical variation of our composite risk measure, we map it at the county level. As shown in Fig. 4, there is substantial geographical heterogeneity in climate risk measure (ZW). The higher values of the composite risk measure are found in the Southern half of the country, particularly in the South east and South west. This pattern is also found in a recent study by Amornsiripanitch and Wylie (2023) who document the highest climate risk exposure in the Gulf of Mexico and South Atlantic coast. More specifically, our climate risk index shows that high climate risk in the Northeast and in the South is more prevalent among coastal counties. It is also worth noting that many counties in northern Texas and Oklahoma also exhibit moderate levels of climate risk. In the West the highest climate risk area falls in counties on both sides of the California-Arizona border. Last, the Midwest is
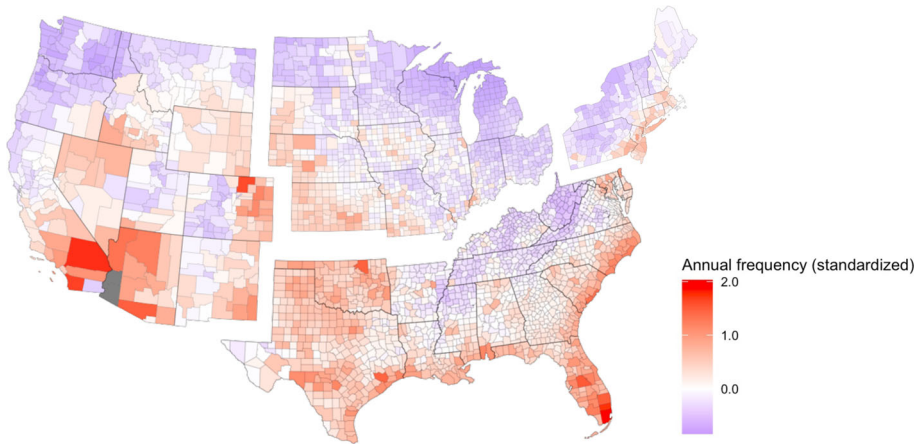
Fig. 4 Composite risk index (ZW) at county level. **Notes:** Map plots composite risk measure for each county. Map separates counties by Census region (Northeast, Midwest, South and West). Heat-map shows counties with more risk in red and counties with lower risk in purple

generally a region with low climate risk and only some counties in Nebraska and Kansas exhibit moderately high climate risk.

Naturally, there is regional variation in the specific climate hazards concerning the population. Obviously, landlocked counties are not exposed to coastal flooding and tornadoes are much more frequent along the Tornado Alley (which includes parts of Texas, Louisiana, Oklahoma, Kansas, South Dakota, Iowa and Nebraska). Figure 5 plots county-level risk levels (based on estimated annual frequency) for the 10 natural hazards with positive weight in our composite measure (see Population Growth and Climate Risk).[19]

Droughts are a serious concern in many counties in the western half of the United States. In contrast, the eastern and southern coastal counties face moderate to high risk of hurricanes. We also note that wildfire risk correlates with risk of droughts, whereas coastal flooding risk largely coincides with risk of hurricanes (particularly in southern counties in Texas, Louisiana and Florida).[20]

In later sections we will also examine whether different natural hazards affect population growth differently, possibly due to differences in the availability of mitigation technologies or other factors.

Is There Population Retreat from High Climate Risk Locations?

County-level Analysis

The findings in the previous section show that population is not retreating from high climate-risk areas. At best, we observe a recent reduction in the gap between growth rates in regions

---

[19] The figures are sorted (top to bottom) in decreasing weight in our composite index. The hazards with the highest weights are droughts (0.21), hurricanes (0.21), riverine flooding (0.18) and tornadoes (0.13).

[20] Note though that several counties in the northwest are at high risk of coastal flooding but are not exposed to hurricanes.
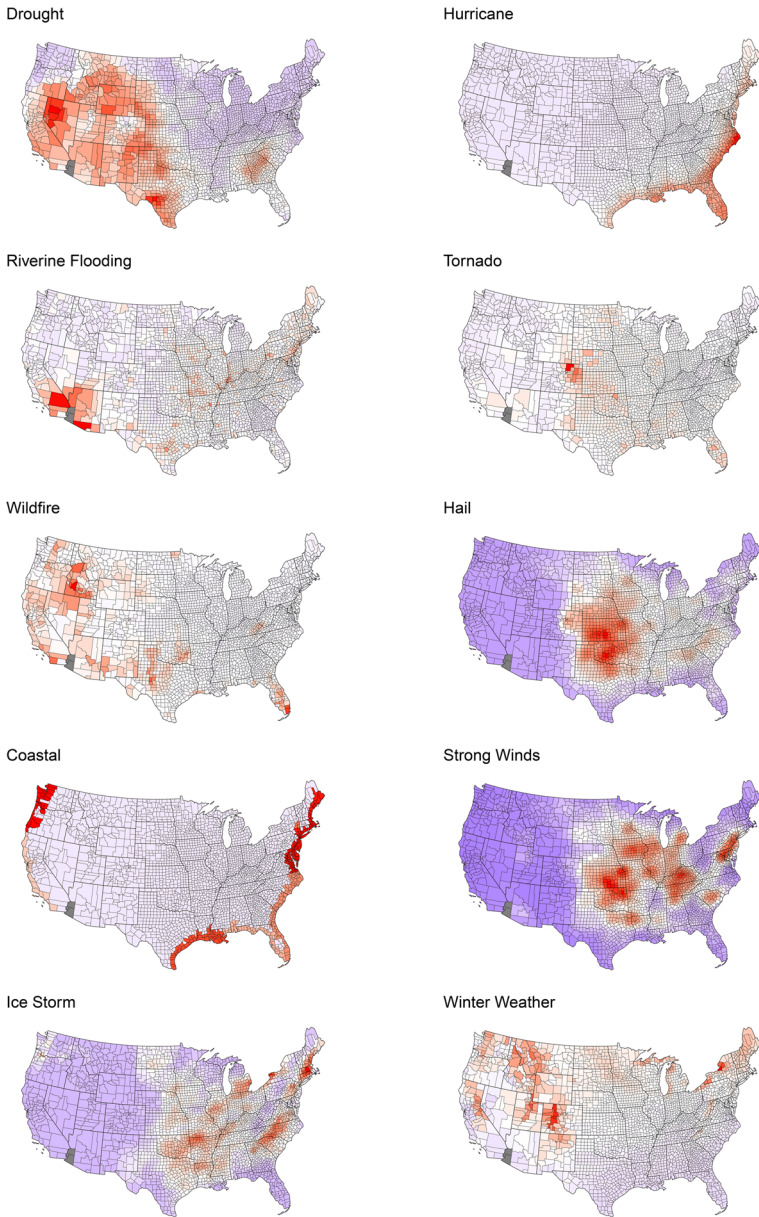
Fig. 5 Risk for individual hazards at county level. **Notes:** Risk score for each hazard at the county level, based on the standardized average annual frequency of each hazard in each couty. We only include hazards who have a positive weight in our composite measure. Order of hazards in this figure is determined by the hazard's weight in our composite measure. Heat-map shows counties with more risk in red and counties with lower risk in purple

with high and low climate risk. This section will offer a more formal test of the *retreat hypothesis* exploiting cross-county variation.

Naturally, if population is growing faster in high-risk areas it must be because the pull factors in those areas outweigh the expected losses associated to climate risk. The pull factors may differ in terms of geographic scope.[21] Some may spread across whole states (e.g., low taxation), other pull factors may better coincide with commuting zones (e.g., strong labor markets), yet others may operate at the county or sub-county level (e.g., nice views or proximity to nature).

We hypothesize that, if we were able to condition on all relevant pull factors, we would be able to observe population retreat from high-risk areas. In other words, individuals currently living in an area with high climate risk would be willing to relocate to lower climate-risk areas *with the same attributes*. We refer to this as the *conditional retreat hypothesis* and we will also test it below.

We analyze these questions exploiting cross-county variation to estimate differences in population growth on the basis of climate risk, where growth will sometimes be defined relative to the neighboring counties to net out the effects of region-specific factors. Our primary interest is on the period 1990-2020, when climate risk has become increasingly salient, but we provide estimates for a longer time period in order to examine if there has been a departure from long-term population trends.[22].

We consider a series of cross-sectional models that differ in their dependent variable. To fix ideas, denote the average (decadal) change in log population in county c between years 1990 and 2020 by $\overline{g}_c$. We posit that

$$\overline{g}_c = \alpha + \beta_1 RiskMed_c + \beta_2 RiskHigh_c + u_c, \tag{3}$$

where $RiskMed_c$ and $RiskHigh_c$ are dummy variables taking a value of one for medium or high-risk counties, respectively. The omitted category are counties with low (or non-existing) risk. Coefficients β_1 and β_2 estimate the excess mean population growth in medium-risk and high-risk counties relative to low-risk counties nationwide. We cluster standard errors at the level of commuting zones. This clustering allows for arbitrary spatial correlation patterns across counties (or tracts) within commuting zones.

It is also interesting to ask if counties with higher climate risk grow more (or less) than neighboring counties located in the same commuting zone. Appropriately demeaning the dependent variable allows us to address this question. In this case, we estimate the model

$$\overline{g}_c - \overline{g}_z = \alpha + \tilde{\beta}_1 RiskMed_c + \tilde{\beta}_2 RiskHigh_c + u_c, \tag{4}$$

where the dependent variable is the average population growth in county c net of the average population growth among all counties in the same commuting zone z. To the extent that commuting zones characterized by higher climate risk grow systematically more (less) than commuting zones with low risk, the estimates for $\tilde{\beta}_1$ and $\tilde{\beta}_2$ will be lower (higher) than the analogous estimates obtained in Eq. 3.[23] Note also that Eq. 4 neutralizes the effect of factors

---

[21] For example, Rappaport and Sachs (2003) show that US economic activity is overwhelmingly concentrated in coastal counties and argue that this is a result of the opportunities that stem from proximity to the ocean in terms of productivity and quality of life. Glaeser et al. (2001) argue that the inherent economic success of a city hinges on its consumption value, which is closely related to the amenities it offers.

[22] Beeson et al. (2001) show that migration decisions in the late 1800's were largely driven by natural characteristics of the counties, such as access to water transportation and mineral resources.

[23] Note that $\beta_2 = E(g_c|HighRisk) - E(g_c|LowRisk)$, whereas in Eq. 4, $\tilde{\beta}_2 = \beta_2 - (E(g_z|HighRisk) - E(g_z|LowRisk))$.

Table 3 Estimates average population growth by composite climate risk

Period 1990-2020 Climate index Sample cty AvDlnPop	(1) NRI All	(2) ZW All	(3) ZW All net State	(4) ZW All net CZ	(5) ZW HM CZ net CZ	(6) ZW All FE CZ
HighRisk	0.071***	0.029***	0.010	0.005**	0.007*	0.023***
	[0.008]	[0.008]	[0.007]	[0.002]	[0.004]	[0.007]
MedRisk	0.023***	0.006	0.002	-0.001	0.001	0.003
	[0.005]	[0.006]	[0.005]	[0.002]	[0.004]	[0.005]
Constant	0.041***	0.038***	-0.004	-0.001	-0.002	0.041***
	[0.002]	[0.005]	[0.004]	[0.001]	[0.003]	[0.004]
Observations	3,103	3,106	3,106	3,106	2,584	3,103
R-squared	0.027	0.013	0.002	0.002	0.002	0.006
Mean Dep. Var.	0.048	0.048	0	0	0	0.048

Notes: The dependent variable is the change in the log of population between 2020 and 1990, divided by 3 (decades). In column 1 the measure of climate risk is FEMA's NRI. In columns 2-6, climate risk is measured using our index based on the (weighted) aggregation of the standardized annual frequencies of all hazards, where the weights are based on the monetary value of the nationwide damage due to each hazard. The dependent variable in columns 3 and 4 nets out the average population growth in the State and Commuting zone, respectively. Column 5 restricts the sample to counties in commuting zones with above average proportion of medium or high climate-risk counties. In column 6, the dependent variable is the (gross) change in the log of population but the model includes (722) commuting-zone fixed effects. In all models the omitted category is low risk. Standard errors are clustered at the level of commuting zones. P-values: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

that make a commuting zone more (or less) attractive, on average, than other commuting zones. Examples of such factors are cross-state (or cross-city) differences in taxation, weather, or the robustness of their local economies during the period of consideration. Hence, this model provides a test of the conditional retreat hypothesis.[24]

Main Results

We now turn to the estimation of Eq. 3. Table 3 reports the results. Before turning to our composite climate risk index, we employ FEMA's National Risk Index (NRI). As seen in column 1, there is a strong positive association between high-risk counties (on the basis of the NRI) and population growth. However, this index is constructed on the basis of the frequency of natural disasters *and* a measure of exposure, which includes building values that are obviously correlated with population. As a result, there is a nearly mechanical relationship between high values of the NRI and a county's population growth. Primarily for this

---

[24] An alternative approach to netting out factors that affect all counties in a commuting zone equally would be to include commuting-zone fixed-effects in the estimation of Eq. 3. However, in that case, the interpretation of the coefficients of interest is less straightforward. When we include fixed-effects, identification is based on the correlation between the transformed (demeaned) population growth and the transformed risk dummy. Note that the transformed risk dummy becomes a continuous variable, so it is no longer the difference in the mean for high-risk versus low-risk counties. The demeaned dummy variable then becomes a measure of relative risk (vis-a-vis the corresponding commuting zone) whose variation is entirely driven by the fraction of high-risk counties in the commuting zone. In our opinion, this 'local' measure of county risk (which depends on the mean risk among the counties in the commuting zone) is less helpful than using a 'global' measure of county risk. At any rate, we shall also report estimates obtained by including fixed-effects.

reason, we built a composite index that is purely based on the average annual frequency of natural hazards in each county. Instead, columns 2-6 employ our climate risk index (ZW). As expected, the association between population growth and climate risk at the county level is considerably weaker in column 2 than in column 1. Nonetheless, we still find evidence of higher population growth (over the last 3 decades) in high-risk counties. We estimate the growth gap between low and high-risk counties to be 2.9 log points (i.e., about 3%) per decade. In contrast, medium-risk counties have grown, on average, at the same rate as low-risk counties over the last 30 years. Thus, we reject the retreat hypothesis, confirming the findings in Fig. 3. In other words, high-risk counties continue to gain population, presumably because the pull factors in these locations offset the expected losses associated with climate risk.

Columns 3 and 4 examine whether higher-risk counties have grown disproportionately *relative* to their neighbors. Respectively, the dependent variables in these columns net out the average population growth in the state and commuting zone where each county is located. The point estimates fall in value, indicating that high-risk counties tend to be located in high-risk areas (states or commuting zones). However, the estimates show that high-risk counties have grown at a higher rate than the commuting zone (or state) where they are located. Based on column 4, we estimate the high-low *net* gap in population growth to be 0.5 log points per decade.

Column 5 restricts the sample to commuting zones with an above average proportion of medium-risk or high-risk counties, which increases the net population growth gap between high-risk and low-risk counties. Last, column 6 reports estimates from a model that includes commuting-zone fixed-effects (where the dependent variable is the average change in log population). Intuitively, this model correlates deviations in population growth relative to each county's commuting zone with a measure of relative risk. The estimates entail a larger gap in population growth between high-risk and low-risk counties.

In sum, our estimates show that high-risk counties have grown substantially more than low-risk counties over the last 3 decades, even when the comparison is restricted to counties in the same commuting zone, which is commonly considered as a fair approximation of the geographical scope of local economies. This result points to the presence of important pull factors at the county or sub-county levels and imply a rejection of both the unconditional and conditional retreat hypotheses.

Flexible Relationship

Let us now examine a more flexible model than Eq. 3 using local linear regression. This analysis will be informative regarding the functional form for the relationship between our composite index (as a continuous variable) and the average population growth. The results are depicted in Fig. 6. The top figure plots average decadal population growth and our frequency-based composite risk index at the county level. The figure shows a positive association between climate risk and population growth across the whole range of the composite index, with the exception of the first bin. The bottom figure is the conditional counterpart of the previous figure, where each county's average population growth rate has been demeaned using the corresponding commuting-zone value. In this case, the relationship is both closer to a linear function and exhibits a smaller slope.[25]

---

[25] Conservatively, our regression models do not assign extra weight to more populated counties, which tend to have the largest values for the composite climate risk index.
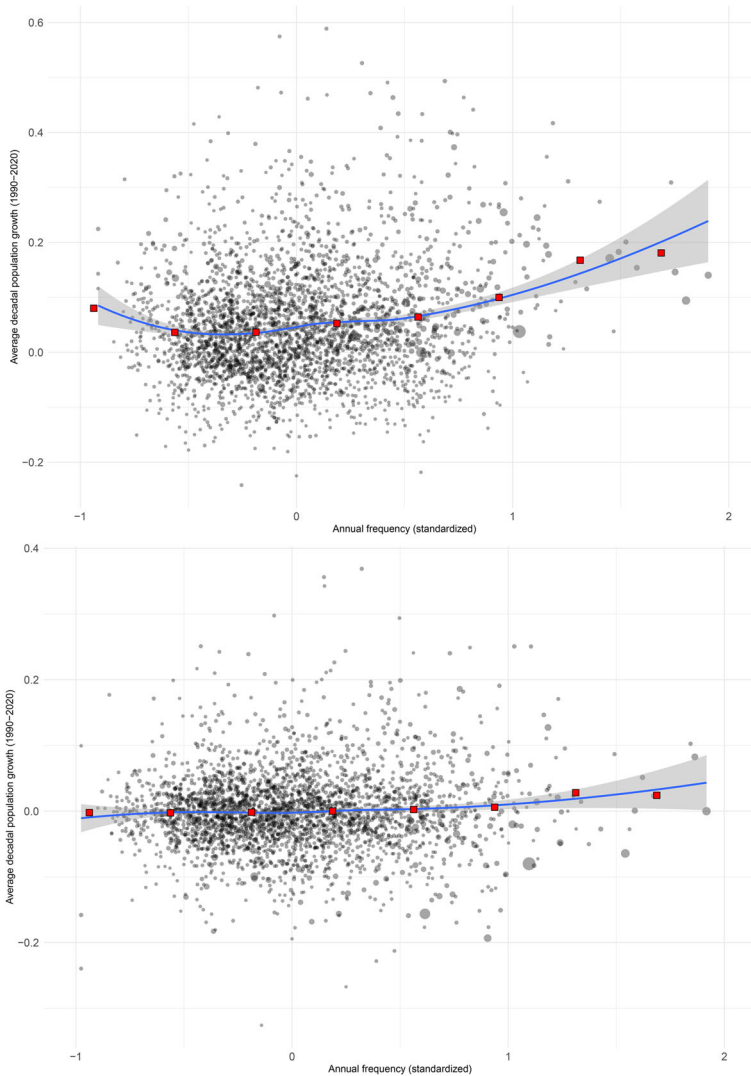
Fig. 6 Population growth and composite climate risk. Flexible functional form. **Notes:** Each point is a county and the horizontal axis correspond to our main composite climate risk index (weighted average of annual frequency of each natural hazard). The top figure plots the average decadal population growth in the period 1990-2020 for each county. The bottom figure is analogous but the data for each county have been demeaned using the average value in the corresponding the commuting zone. Each red square is the local linear regression estimate for the corresponding bin. The shaded band depicts the 95% confidence interval

Evolution Over Time

It is also interesting to examine the evolution of the growth differentials between high (and medium) risk counties and low risk counties over time.[26] The results are collected in Fig. 7.

---

[26] An important caveat is that county boundaries have only been harmonized for years 1980-2020 so as to be stable over time. As we move back in time, there will be an increasing number of boundary changes, which reduces the reliability of the estimates.
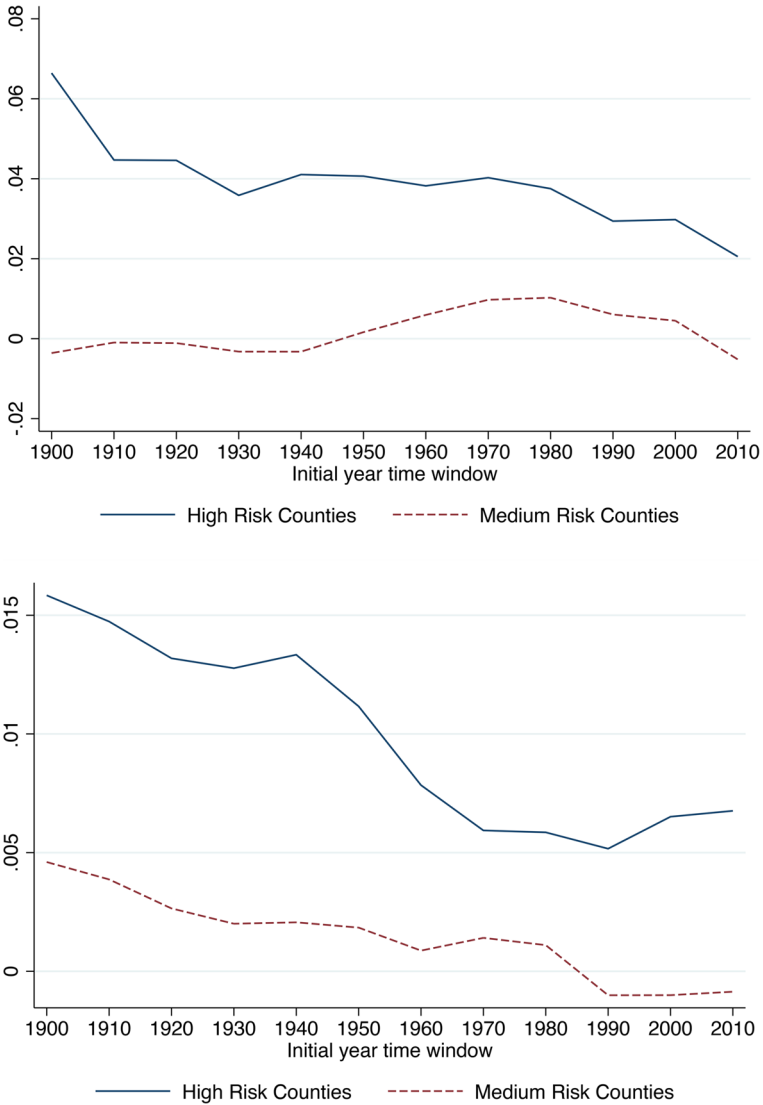
Fig. 7 Evolution population growth differential. County population growth (top); demeaned by CZ average growth (bottom). **Notes:** Risk categories based on weighted composite risk index (ZW). Each point estimate refers to the average decadal change in the log of population between the corresponding initial year and final year 2020. Point estimates obtained from models for the average decadal population growth (top) and for the same variable but demeaned using average growth in the corresponding commuting zone (bottom). In all cases the omitted category are counties with low risk

The top figure is based on models where the dependent variable is the average population growth in the county, whereas in the bottom figure the dependent variable has been demeaned using the average population growth in the corresponding commuting zone.

Both figures indicate a secular reduction of the excess growth of high risk-counties relative to low-risk counties, but they diverge in regard to the recent trends. Over the last 30 years, the

excess growth of high-risk counties (relative to low-risk counties) has trended down when we do not net out the growth rate of the corresponding commuting zone (top figure), but this is not the case when we consider county growth relative to each county's commuting zone (bottom figure). Tentatively, this finding suggests that in recent times commuting zones exposed to high climate risk may be losing gravitational pull.

It is also interesting to examine the robustness of these findings to the weights used in the construction of our composite risk index. Accordingly, Fig. 10 plots the estimated excess population gaps based on the unweighted composite index (Z), which assigns equal weights to all hazards. The patterns we obtain are qualitatively similar: throughout the whole period, high-risk counties exhibit excess population growth relative to low-risk counties, regardless of whether we net out population growth in the corresponding commuting zone. However, the estimated excess growth is much lower when we use the unweighted composite index. Specifically, between 1990 and 2020, the excess growth of high-risk counties is estimated to be around 1 log point per decade when we rely on the unweighted composite index. This is substantially lower than the 2.9 log point excess obtained using the weighted composite index. Interestingly, the estimated excess growth in relation to the corresponding commuting zone is very similar whether we use the weighted or unweighted versions of the composite index. As we shall see later (in Heterogeneity by Climate Hazard), the quantitative discrepancies between the two versions of the composite risk index reflect the diverging population-risk dynamics for some individual natural hazards.

Summing Up

Our county-level analysis offers two main conclusions. First, we find no evidence of population retreat from areas with high climate risk. In other words, we reject the *unconditional retreat hypothesis*. Over the last three decades, on average, high-risk counties have grown more than low-risk counties (by 2.9 log points per decade). In addition, the same qualitative pattern is found when considering each county's population growth relative to the growth of the corresponding commuting zone, which neutralizes the effect of state and commuting-zone characteristics (such as differences in taxation or strong local labor markets). Thus, we also reject the *conditional retreat hypothesis*, suggesting that the factors that attract people to high climate-risk areas operate at the county or sub-county levels.

Micro Retreat: tract-level Analysis

There is an important caveat to the conclusion of no retreat from high climate risk locations, even after controlling for state and commuting-zone pull factors. It might be the case that retreat takes place at the *sub-county* level. In other words, while population in high-risk counties has been growing disproportionately, it is conceivable that the growth is concentrated in low-risk towns or neighborhoods *within* those counties. If this were the case, the outlook would be much more optimistic. We refer to the disproportionate growth of low-risk sub-county locations as the *micro retreat hypothesis*.

In order to assess whether *micro-retreat* is taking place, we switch to tract-level data. There are about 70,000 Census tracts in the United States. The main implementation challenge is the changing tract boundaries between each decennial Census. We use the LTBD dataset (Logan et al., 2014), which contains harmonized tract boundaries for the 1970 through 2010

Table 4 Estimates population growth by composite climate risk. Tracts Analysis

1990-2020 Tracts AvDlnPop	(1) All	(2) All net State	(3) All net CZ	(4) All net County	(5) HM Cty net County	(6) FE Cty
HighRisk	0.090***	0.016	0.017***	0.015***	0.031***	0.071***
	[0.023]	[0.011]	[0.005]	[0.003]	[0.006]	[0.005]
MedRisk	0.028	0.015*	0.011*	0.008**	0.023***	0.039***
	[0.018]	[0.009]	[0.006]	[0.004]	[0.007]	[0.004]
Constant	0.062***	-0.011*	-0.010**	-0.008***	-0.023***	0.061***
	[0.019]	[0.006]	[0.004]	[0.002]	[0.006]	[0.003]
Observations	57,931	57,931	57,931	57,931	47,291	58,438
R-squared	0.018	0.001	0.001	0.001	0.001	0.003
Number FE						2,821
Mean Dep.Var.	0.099	0	0	0	0	0.099

Notes: Dependent variable is the average decadal change in the log of population between 1990 and 2020. In all columns the climate risk categories are defined on the basis of the weighted composite index. Column 5 restricts the estimation to counties with an above-average proportion of medium-risk or high-risk tracts. In all models the omitted category is low risk. Standard errors clustered by commuting zone in all columns, except column 6 where clustering is at the county level. P-values: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Censuses. When we merge these data with the FEMA NRI dataset, we obtain 59,030 tracts for years 1990, 2000, 2010 and 2020.[27]

Our empirical specifications are analogous to those used in our county-level analysis; the only changes are that observations are now defined at the tract level (indexed by r) and that we use county-level averages to compute relative tract-level growth. Namely, the models we consider are:

$$\overline{g}_r = \alpha + \beta_1 RiskMed_r + \beta_2 RiskHigh_r + u_r \tag{5}$$

$$\overline{g}_r - \overline{g}_c = \alpha + \beta_1 RiskMed_r + \beta_2 RiskHigh_r + u_r \tag{6}$$

The bottom panel in Table 2 describes the main variables in the tract-level dataset. Roughly, our merged dataset (which excludes Hawaii and Alaska) contains 58,500 tracts. Over the last 5 decades, the average tract has grown by 16.4 log points per decade, which is much higher than the corresponding value in the counties dataset (6.2 log points).[28] The growth rate for the average county has also declined over time. In the last decade this value was 4.8 log points (compared to 0.6 log points in the county-level data).

The estimates of the relationship between current climate risk and population growth over the last 3 decades are collected in Table 4. The first column estimates Eq. 5. The estimates show that high climate-risk tracts have grown at a much higher rate than low-risk tracts nationwide (by a differential of 9 log points per decade). In columns 2-4 we demean the dependent variable using the average growth rate in the corresponding state, commuting zone and county. As expected, the high-low relative gap decreases in size, but remains almost

---

[27] For years 1970 and 1980 the number of tracts is significantly lower (around 49,000 tracts on average) so we exclude these years from the main estimation sample. Census tracts in the LTBD data we use are harmonized to 2010 boundaries.

[28] It is worth noting that population growth at the tract level is censored. When a tract reaches a certain threshold (around 4,000 individuals), the tract is split into two separate tracts. However, this is not the case in our harmonized dataset, which keeps boundaries stable at their 2010 values.

unchanged across the three columns. Namely, high-risk tracts have grown about 1.5 log points more than low-risk tracts in the same state/CZ/county. Furthermore, column 5 shows that the excess growth in high-risk tracts is even larger in counties with high climate risk (defined as counties with above average proportion of medium-risk or high-risk tracts). Last, column 6 shows that the results are qualitatively similar when employing a model that includes tract-level fixed-effects (though the estimated high-low excess growth is much larger than in our preferred specification).

In sum, our estimates entail a clear rejection of the *micro retreat hypothesis* stated above. In fact, not only high-risk *counties* are growing more than low-risk ones (within the same commuting zones). Our results here show that that high-risk *tracts* are also growing more than low-risk (and medium-risk) tracts within the same county. Thus, the sub-county population dynamics imply that the degree of exposure to climate risk is underestimated in the county-level analysis. Furthermore, our estimates suggest that the pull factors that make high-risk tracts attractive are highly localized in scope (at the sub-county level).

Heterogeneous Effects by Residential Capital

Our estimates based on the national sample have failed to provide evidence of population retreat from high-risk locations, even after neutralizing the effects of state-level, commuting-zone and county-level pull factors. This suggests the presence of powerful localized pull factors that still outweigh the costs associated to exposure to climate risk.

However, these findings could vary on the basis of local characteristics, such as the degree of urbanization. In particular, counties with robust local economies and highly concentrated physical assets may invest more in resiliency measures to protect from climate shocks, whereas capital-poor regions may not be able to afford such investments. As a result, population dynamics may differ substantially across high climate-risk locations on the basis of the value of their residential capital stock. In fact, in the context of coastal flooding risk, Lin et al. (2021) show that residential construction in the United States is increasingly concentrated in high-risk *and* high-density coastal areas, but it is not known if these dynamics apply more generally to other climate hazards.[29]

To analyze these questions, we partition counties on the basis of (i) overall value of their residential capital, (ii) median value of homes, and (iii) economic density (defined as overall value per unit of surface). We measure housing values using the 2000 Census (100% sample, Census Table) and extend our previous empirical model to allow for heterogeneous effects of climate risk on population growth for counties above and below the median value of the corresponding discriminating variable.[30] The overall housing stock in the median U.S. county in year 2000 had a value of $0.9 billion; the median home value in the median county was $75,600 (and the median homeownership rate was 80.1%). We use these cutoff values to

---

[29] Relatedly, Balboni (2021) estimates large costs from coastal favoritism in deciding the location of public infrastructure works.

[30] So far we have not located comprehensive data on housing values for all U.S. counties for year 1990 so we use year-2000 values. We do have the 1990 data for Census tracts based on the 5% Census sample. However, aggregation of the median tract housing values to the county level results in 1,877 counties, well short of the approximately 3,100 plus counties in the United States. As we show below, the stock of residential capital is very persistent at the county level over a 10-year period, resulting in the partition of counties being practically the same. Hence, relying on year-2000 values to analyze population growth over the 1990-2020 period is a fairly safe choice.

Table 5 Heterogeneity by residential capital stock

Het by Net AvGPop Since	(1) ValueH 1990	(2) ValueH net CZ 1990	(3) ValueH net CZ 2000	(4) MedValH 1990	(5) MedValH net CZ 1990	(6) ValH/Area 1990	(7) ValH/Area net CZ 1990
Constant	0.006	-0.010***	-0.011***	-0.018***	-0.010***	0.009	-0.008***
	[0.005]	[0.002]	[0.002]	[0.004]	[0.002]	[0.006]	[0.002]
Hval	0.059***	0.017***	0.019***	0.095***	0.016***	0.051***	0.013***
	[0.008]	[0.003]	[0.003]	[0.008]	[0.003]	[0.008]	[0.003]
HighRisk	-0.003	-0.004	-0.004	0.027***	-0.003	0.009	-0.002
	[0.008]	[0.003]	[0.003]	[0.007]	[0.003]	[0.009]	[0.003]
MedRisk	-0.000	-0.004	-0.005*	0.011**	-0.006**	-0.004	-0.006**
	[0.006]	[0.003]	[0.003]	[0.005]	[0.003]	[0.007]	[0.003]
Hval × HighRisk	0.058***	0.017***	0.018***	0.030**	0.021***	0.052***	0.018***
	[0.012]	[0.006]	[0.006]	[0.012]	[0.006]	[0.013]	[0.006]
Hval × MedRisk	0.024**	0.010**	0.012***	0.014	0.015***	0.029***	0.013***
	[0.010]	[0.004]	[0.004]	[0.010]	[0.004]	[0.010]	[0.005]
Observations	3,106	3,106	3,106	3,106	3,106	3,106	3,106
R-squared	0.203	0.054	0.069	0.314	0.064	0.176	0.045
Mean DepVar	0.048	0.048	0.025	0.048	0.048	0.048	0.048

Notes: The header of each column defines the variable used to partition counties as being above or below the median value for that variable (value of housing stock, median value of residential properties, or value of the housing stock over area of the county). The data for housing values is from the 2000 Census (summary tables). Dependent variable is the change in the log of population between 2020 and 1990 (except for column 3 where the change is relative to year 2000) divided by the number of decades between the beginning and endpoint of the time interval. In all columns the climate risk categories are defined on the basis of the weighted composite index defined at the county level. In all models the omitted category is low risk. Standard errors clustered by commuting zone. P-values: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

partition counties according to whether their year-2000 values for these variables are above or below the corresponding mean.[31]

Our dependent variable is the 1990-2020 average change in log population. As before, we present both estimates of models where the dependent variable is the gross population growth rate of counties and models where we net out the mean value for neighboring counties.

Table 5 collects the results. Column 1 estimates the model for the average change in population growth. This specification includes interactions terms that allow for heterogeneous coefficients for counties with low (versus high) overall residential capital stock, where the cutoff is given by the median value of residential capital across all counties (in year 2000). The estimates in column 1 show that in counties with low residential capital stock, climate risk is not related to population growth. Instead, the picture is very different in counties with high residential capital stock. First, population growth in these counties is uniformly higher in these counties (by 5.9 log points per decade) regardless of climate risk. But, additionally, high-capital, high-risk counties have grown more than low-risk counties that also have a large residential capital stock. The estimates in column 2 show that the excess growth in high-risk, high-capital counties is also observed after netting out the growth of the corresponding commuting zone. However, as before, this population agglomeration in high-risk counties is not happening in counties with lower residential capital. Column 3 focuses on population growth between years 2000 and 2020, which is better aligned with the year in which we measure housing values. The results are practically identical to those in column 2, confirming that the correlation for county housing values for years 2000 and 1990 is very high.

Columns 4-5 repeat the analysis but, this time, counties are partitioned on the basis of median housing values (among homeowners). The estimates confirm the agglomeration of population in high-risk counties with high median housing values. In regard to low-value counties, we now find higher population growth in high-risk counties (column 4), but this is largely due to the relatively high population growth in the corresponding commuting zones. In fact, high-risk counties with low median housing values have grown less than their neighboring counties (in the same commuting zone).

Columns 6 and 7 partition counties by economic density, defined as the value of the stock of residential capital divided by the area of the county. The results are also in line with what we found in the previous columns of the table.

In conclusion, our analysis in this section clearly indicates that the agglomeration of population in high-risk areas is a phenomenon taking place in economically dense, urban areas, with large stocks of residential capital and high median values. This finding echoes the conclusions in Lin et al. (2021). In contrast, population is not booming in high-risk counties in less urbanized areas. If anything, these counties are growing disproportionately less than otherwise similar low-risk counties. All in all, these findings suggest that to live in high-risk areas, local residents require a compensating differential, possibly associated with high residential capital or robust local economies.

Regional Heterogeneity and Net Migration

This section investigates if the finding of higher population growth in areas with higher climate risk found in the national samples is also present in regional subsamples (defined as

---

[31] Our results also relate to Fussell et al. (2017) who analyze the effects of hurricanes (and tropical storms) on population growth and also find heterogeneous effects (on the basis of prior population trends).

Census divisions). In addition, we will assess whether the findings are driven by disparities between high-risk and low-risk areas in natural population growth or in net migration.

Regional Heterogeneity

Table 7 estimates gaps in average decadal population growth at the county level for the period 1990-2020 on the basis of climate risk. Column 1 simply reproduces our earlier finding: nationwide high-risk counties have grown more than low-risk counties (by 2.9 percent per decade). Columns 2 through 5 provide estimates for each census division separately. We find substantial regional heterogeneity in the growth gap between high-risk and low-risk counties. As seen in columns 2 and 4, between 1990 and 2020, in the Northeast and South, population has increased much more rapidly in high-risk counties than in low-risk counties (by approximately 3.7 percent and 4.9 percent, respectively).

Interestingly, the pattern is markedly different in the Midwest and West (columns 3 and 5). In these regions, high-risk counties have grown *less* than low-risk counties between 1990 and 2020. In fact, in the Midwest, the average high-risk county actually experienced a population decline (by 1.1 log point per decade) whereas low-risk counties actually gained population. In the West, the average county experienced robust population growth between 1990 and 2020 regardless of climate risk, but high-risk (and medium-risk) counties grew 1 to 2 log points less per decade than high-risk counties.

The Role of Net Migration

Next, we examine whether the the heterogeneous regional patterns in the differential growth of counties with high climate risk is driven by diverging patterns of net migration. Specifically, the estimates in Table 7 suggest that high-risk counties in the Northeast and South have been net recipients of migrants whereas high-risk counties in the Midwest and the West have experienced negative net migration.

To address this question, we rely on a recent dataset by Egan-Robertson et al. (2023). This dataset contains decadal county level net migration from 1950 to 2020. Net migration for each decade is estimated as a residual, computed as the overall population change minus the *counterfactual* population growth driven purely by natural growth. In turn, the latter is estimated by aging forward the population at the beginning of the decade, subtracting deaths, and adding births. Importantly, we computed the counterfactual population change in the absence of net migration over, say, the period 1990-2020 as the 2020 population resulting purely from natural growth over the 3 decades minus the overall population in 1990.

Figure 8 reports the results. Let us consider first the nationwide estimates. As we did earlier (Fig. 7), the solid (blue) line reports the gap in the average decadal change in the log population of high-risk versus low-risk counties, between each initial year and 2020. As was the case in Fig. 7 (top panel), the average total population growth gap (between high-risk and low-risk counties) in the net migration dataset (Egan-Robertson et al., 2023) is estimated to be around 4 percentage points per decade higher in the high-risk counties for the time windows starting in 1960 through 1980, but narrowing substantially for time windows starting from 1990 onward. In fact, Fig. 8 implies that the population growth gap during the 2010s has fallen almost to zero. This decline was also displayed in Fig. 7 (top panel), but quantitatively less drastic. The discrepancy between the two datasets is largely due to the exclusion of the population age 75 and over in the dataset by Egan-Robertson et al. (2023), as explained in Appendix A.
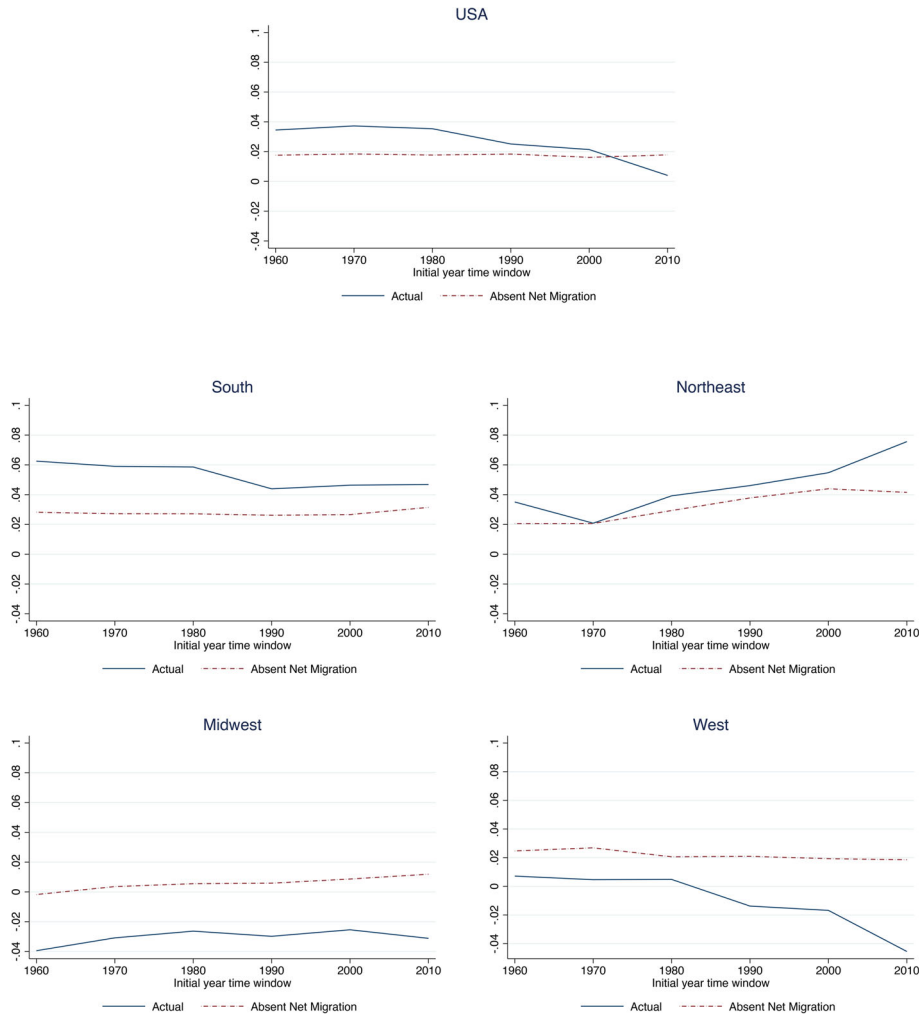
Fig. 8 Population growth gaps, inclusive and exclusive of net migration. **Notes:** County-level population estimates (overall or absent net migration) are from Egan-Robertson et al. (2023) and exclude population age 75 or older. The solid (blue) line reports point estimates of the gap in average decadal change in the log of total population in high-risk counties relative to low-risk counties between the corresponding initial year and final year 2020. The dashed (red) line reports analogous estimates but for the change in log population **absent net migration** into the county from the initial year onward

The dashed (red) line in the top subfigure in Fig. 8 plots the gap in the *counterfactual* population growth absent net migration (throughout the full period of interest) between high-risk and low-risk counties for the national sample.[32] The data show that high-risk counties would have grown about 2 percent more per decade than low-risk counties uniformly between 1960 and 2020. Hence, since 1980, the gap between total and counterfactual population growth has been narrowing, turning negative in the 2010s. In other words, since 1980 net

---

[32] Importantly, the dashed (red) line in Fig. 7 plotted the overall population growth gap between *medium-risk* and low-risk counties.

migration into high-risk counties has been falling in relative terms and, since 2010, net migration in high-risk counties has fallen below net migration in low-risk counties.

Let us now turn to examine the estimated growth gaps by census division. Consider first counties in the South of the country. Absent net migration, the population in high-risk counties would have steadily grown by about 3 percent per decade *more* than in low-risk counties. But net migration increased the excess growth in high-risk counties to around 5-6 percent per decade. The pattern is similar in the Northeast, with net migration contributing to the higher overall population growth of high-risk counties.

In contrast, in the Midwest and the West, net migration has had the opposite effect, as indicated by the uniformly lower solid line in the figures as compared to the dashed line. Specifically, since 1960 in the Midwest, population growth would have been very similar in high-risk and low-risk counties in the absence of net migration, but net migration resulted in a substantially *lower* overall population growth in high-risk counties (by about 3 percent per decade). In the West, in the absence of net migration, high-risk counties would have grown more (by about 2 percent per decade) than low-risk counties. But, similarly to what took place in the Midwest, net migration flows reversed the sign of the gap in overall population growth, which turned negative (i.e. in favor of low-risk counties) since 1990.

It is likely that several factors explain these regional disparities in the sign of net migration into high-risk counties. One such factor is probably the prevalence of economically vibrant local economies on flood-prone coastal areas in the South and Northeastern coast. Differences in the specific hazard mix affecting each region may also play a role.[33] Seeking further clues, next we compare counties that experienced population growth (over the period 1990-2020) to those that suffered a decline in terms of their climate risk exposure. The top panel in Table 6 reports this information for the U.S. as a whole.[34] Among counties with negative population growth, the composite risk index takes a value of negative 0.04. In contrast, the mean value for growing counties is 0.02. The difference between the two values is small but already indicates that growing counties tend to have (slightly) higher exposure to climate risk.

Let us now turn our attention to the Northeast region (in the second panel). The first row of the panel already indicates that the Northeast is heavily exposed to hurricanes (0.43), riverine flooding (0.68) and coastal flooding (0.88). Moreover, growing counties have a very high exposure to these hazards, even relative to the rest of the region (with values of 0.73, 0.79 and 1.43, respectively). In contrast, counties with falling population have substantially lower exposure to these hazards. Rather similarly, hurricanes are the most prominent natural hazard in the South (0.47 for all counties in the region), and growing Southern counties are characterized by high risk of hurricanes (0.58 for growing counties). In contrast, the main natural hazard in the Midwest is hail (0.47) and growing counties have relatively low exposure to this particular hazard (0.36). In turn, the main exposure in the West is to drought (0.96) and wildfires (0.94). Interestingly, shrinking counties exhibit a very high risk of drought (1.37), consistent with our retreat hypothesis. However, wildfire risk is higher in growing counties (1.01) than in those losing population (0.60). All in all, these observations underscore the presence of regional differences in exposure to each type of natural hazard. The following sections will try to shed some light on the nature of this distinction.

---

[33] As illustrated in Fig. 5, the West is mainly exposed to droughts, wildfires and, to a lesser extent, coastal flooding. Coastal areas in the South and the Northeast are at high risk of coastal flooding and hurricanes. In contrast, the Midwest has a relatively low exposure to all climate hazards. We explore this dimension in the next section.

[34] The average value across *all counties* for all risk measures in the table is zero because we standardized the annual frequencies of each hazard (and the composite) to have a zero mean and a unit standard deviation.

Table 6 Regional heterogeneity in climate risk and population change (1990–2020)

US/Region	ΔPop 1990–2020	Composite risk	Drought	Hurricane	Riverine flooding	Tornado	Wildfire	Hail	Coastal flooding
USA		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Negative	-0.04	0.00	-0.22	-0.10	0.21	-0.07	0.24	-0.21
	Positive	0.02	0.00	0.11	0.05	-0.10	0.04	-0.11	0.10
Northeast		-0.14	-0.86	**0.43**	**0.68**	-0.69	-0.46	-0.68	**0.88**
	Negative	-0.29	-0.90	-0.05	0.51	-0.61	-0.48	-0.69	-0.03
	Positive	0.02	-0.84	0.73	0.79	-0.74	-0.45	-0.68	1.43
South		0.15	0.17	**0.47**	-0.08	0.08	0.07	0.09	0.13
	Negative	0.10	0.26	0.17	-0.23	0.2	0.28	0.16	-0.05
	Positive	0.17	0.14	0.58	-0.02	0.04	0	0.07	0.19
Midwest		-0.16	-0.42	-0.49	0.07	0.20	-0.37	**0.47**	-0.34
	Negative	-0.11	-0.22	-0.51	-0.05	0.39	-0.36	0.59	-0.34
	Positive	-0.2	-0.61	-0.48	0.18	0.03	-0.38	0.36	-0.34
West		-0.04	**0.96**	-0.58	-0.25	-0.39	**0.94**	-1.12	-0.02
	Negative	-0.01	1.37	-0.58	-0.58	-0.12	0.60	-0.87	-0.34
	Positive	-0.04	0.89	-0.58	-0.19	-0.43	1.01	-1.17	0.04

Notes: For each region in the table (USA as a whole, Northeast, Midwest, South and West), we report average climate risk (using the composite or hard-specific measures) for the whole region (first row of each panel), the counties in the region with negative net migration (second row) and the counties in the region with positive net migration (third row). Counties with negative (positive) net migration lost (gained) population between 1990 and 2020. We focus on the six individual hazards with the highest weights in our composite index

Heterogeneity by Demographic Group

This section examines if the population trends described above differ along two demographic dimensions: age and race. For ease of comparison, the top panel in Table 8 simply reproduces results from the previous section (Table 7).

The middle panel of the table focuses the analysis on the growth of the population age 65 and above. The estimated intercepts in columns 1-5 show that this demographic group has grown substantially more than the whole population over the 1990-2020 period (15.6 versus 3.8 log points per decade), fueled by the aging of the baby boom. However, the excess growth of the 65-and-older population in high climate-risk counties has been *smaller* than the excess growth for the overall population. In other words, the attraction power of high climate-risk locations appears to be linked to considerations that are less important to older individuals, suggesting that job opportunities may be the driving factor behind the increasing population agglomeration in high climate-risk areas. It is also worth noting that the South stands out from the other regions because the excess growth of the 65-and-older population in high-risk areas is almost as large as the excess growth for the population as a whole. Namely, the factors attracting the older and younger populations to high-risk locations are much more aligned in the South than elsewhere in the United States.

The bottom panel of the table examines the association between climate risk and the local growth in the non-white population. Once again, the growth of this demographic group over the 1990-2020 period has been much larger than that of the overall population (41.9 log points per decade, as shown in column 1). But, as was the case for the population age 65 and above, the excess growth for this group in the high climate-risk counties has been much smaller than for the overall population. In fact, our estimates suggest that, except for the Midwest, the non-white population has grown less in high-risk counties than in low-risk ones. This pattern suggests that the non-white population may have been priced out of rapidly growing high-risk areas.

Before concluding the section, it is worth turning to column 6, where the dependent variable has been demeaned using the average population growth rate in the commuting zone. This transformation is meant to remove the attraction power of the commuting zone, helping isolate the role of county-level pull factors. As discussed earlier, the estimate in the

Table 7 Regional heterogeneity. Avg. decadal county population growth 1990-2020

Since 1990 Region avGPop	(1) All	(2) NE	(3) MW	(4) South	(5) West
HighRisk	0.029***	0.037***	-0.030**	0.049***	-0.014
	[0.008]	[0.008]	[0.012]	[0.014]	[0.017]
MedRisk	0.006	0.015*	-0.011	0.027**	-0.022*
	[0.006]	[0.008]	[0.008]	[0.012]	[0.013]
Constant	0.038***	0.006	0.019***	0.034***	0.113***
	[0.005]	[0.007]	[0.006]	[0.011]	[0.011]
Observations	3,106	217	1,054	1,421	411

Notes: The dependent variable in the top panel is the average decadal change in the log of population between 2020 and 1990. Column 1 pools all counties. Columns 2-5 restrict samples to the corresponding Census region. In all models the omitted category is low risk. Standard errors are clustered at the level of commuting zones. P-values: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8 Demographic heterogeneity. Avg. decadal population growth 1990-2020. County analysis

Region Demean Pop	(1) All	(2) NE	(3) MW	(4) South	(5) West	(6) All CZ
HighRisk	0.029***	0.037***	-0.030**	0.049***	-0.014	0.005**
	[0.008]	[0.008]	[0.012]	[0.014]	[0.017]	[0.002]
MedRisk	0.006	0.015*	-0.011	0.027**	-0.022*	-0.001
	[0.006]	[0.008]	[0.008]	[0.012]	[0.013]	[0.002]
Constant	0.038***	0.006	0.019***	0.034***	0.113***	-0.001
	[0.005]	[0.007]	[0.006]	[0.011]	[0.011]	[0.001]
Observations	3,106	217	1,054	1,421	411	3,106
R-squared	0.013	0.079	0.014	0.021	0.008	0.002
Age$\geq$ 65						
HighRisk	0.011	-0.014	-0.066***	0.040**	-0.029	0.005*
	[0.011]	[0.012]	[0.016]	[0.017]	[0.026]	[0.003]
MedRisk	-0.005	0.006	-0.047***	0.033**	-0.006	-0.000
	[0.009]	[0.014]	[0.012]	[0.016]	[0.021]	[0.003]
Constant	0.156***	0.149***	0.124***	0.145***	0.266***	-0.001
	[0.007]	[0.012]	[0.010]	[0.013]	[0.018]	[0.002]
Non-whites						
HighRisk	-0.101***	-0.131***	0.049	-0.031	-0.018	-0.030***
	[0.022]	[0.048]	[0.040]	[0.040]	[0.044]	[0.006]
MedRisk	-0.075***	-0.030	0.031	-0.075*	-0.023	-0.025***
	[0.019]	[0.043]	[0.027]	[0.038]	[0.029]	[0.007]
Constant	0.419***	0.385***	0.474***	0.282***	0.462***	0.020***
	[0.016]	[0.039]	[0.022]	[0.036]	[0.025]	[0.005]

Notes: The dependent variable in the top panel is the average decadal change in the log of population between 2020 and 1990, except in the last column where the variable has been demeaned using the commuting-zone average. Analogously, the middle and bottom panels refer to the average decadal change in the log of the population age 65 or older and the non-white population, respectively. Columns 1 and 6 pool all counties in the United States. Columns 2-5 restrict samples to the counties within the corresponding Census region. In all models the omitted category is low risk. Standard errors are clustered at the level of commuting zones. P-values: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

top panel suggests that high-risk counties have more attraction power than other counties in the same commuting zone. The analogous estimate in the middle panel shows that this is also the case for the population age 65 and older, who also seem willing to accept the higher risk of some counties in order to enjoy the local attributes, such as proximity to the coast or to wooded areas. In contrast, the falling non-white population in high-risk relative to low-risk counties reveals that the attributes found in high climate-risk counties are not strong enough to attract this population, or that the average individual in this group cannot afford to live in those counties.

To sum up, our analysis in this section highlights stark differences in the geographic sorting of different socio-demographic groups. More specifically, the increasing population agglomeration in high climate-risk counties appears to be largely driven by white, working-age individuals. Retirement-age and (less affluent) non-white populations appear to be retreating from counties with high climate risk.

This finding is in some sense contrary to that in Amornsiripanitch and Wylie (2023), who find that residents in high-risk areas have lower household incomes than those in low-risk areas. Nonetheless, their study is a reflection of the stock of residents in these areas while our finding refers to the change in that stock. Thus, the presumption that low-income families are *trapped* in high-risk areas may need some qualification.

Heterogeneity by Climate Hazard

This section starts by constructing hazard-specific risk categories, also based on average annual frequencies, but using different thresholds than the composite index that account for the low frequency for some natural hazards. Next, we will examine the conditional and unconditional retreat hypotheses separately for each natural hazard.

There are reasons to suspect that local population dynamics will vary across different natural hazards. For instance, the geographic scope of a natural hazard may be an important aspect shaping residents' adaptation (or the feasibility of resiliency investments). Namely, while some natural hazards impact a whole county with similar intensity (e.g., hurricanes), others are much more localized and affect only a small subset of the county (e.g., coastal flooding). We will refer to the latter as *micro-hazards* and identify them in the data on the basis of within-area variability. Importantly, individuals can easily adapt to *micro-hazards* by simply relocating to nearby towns or neighborhoods with relatively lower climate risk, while still enjoy certain elements that operate at higher geographic levels (such as a strong labor market).

Hazard-specific Risk Categories

Some climate hazards are very infrequent: for 6 of the hazards in our data, the 25th percentile of annual frequency is zero.[35] Thus, the definitions for our categories of low, medium and high risk need to account for this feature of the data. Accordingly, in our definition the *Low* risk category includes locations (counties or Census tracts) with zero or below the 10th percentile of annual frequency. The medium (*Mid*) risk category includes locations with an annual frequency higher than the 10th percentile (hence, strictly positive) but lower than the 50th percentile *conditional* on positive annual frequency. Naturally, the *High* category contains the locations with an annual frequency above the conditional 50th percentile.

Table 9 reports the resulting classification (for counties).[36] As can be seen in column 2, avalanches, coastal flooding, and to a lesser extent, cold waves, hurricanes and heat waves are infrequent hazards. The low frequency partly reflects that some locations have zero exposure to that particular hazard, such as counties in the interior with zero risk of coastal flooding. Our hazard-specific partition of counties into low, medium and high-risk categories can be seen in columns 3-5. Infrequent hazards, such as coastal flooding, entail a high concentration of counties in the low-risk category (88% of counties). In contrast, widespread events, such as lightning, entail a higher concentration of counties in the medium and high-risk categories.

---

[35] In fact for 3 hazards (avalanches, coastal flooding and tsunamis) even the 75th percentile is zero.

[36] As reported in column 6, just 7 hazards account for 94% of the economic cost of natural hazards in the U.S., which determines the weights in our composite risk index (defined as the importance-weighted annual frequency for each hazard). In decreasing order of importance (and weights in parentheses): drought (0.21), hurricane (0.21), riverine flooding (0.18), tornados (0.13), wildfires (0.10), hail (0.06) and coastal flooding (0.05).

Table 9 Risk categories individual hazards (counties)

Hazard	Counties	Freq. Zero	Freq. Low	Freq. Med.	Freq. High	Weights
Composite NRI	3,106		0.79	0.17	0.05	
Z Composite	3,116		0.25	0.50	0.25	
ZW Composite	3,116		0.25	0.50	0.25	
Drought	3,116	0.10	0.10	0.45	0.45	0.21
Hurricane	3,116	0.29	0.29	0.36	0.36	0.21
Riverine flooding	3,116	0.01	0.11	0.40	0.49	0.18
Tornados	3,116	0.00	0.10	0.40	0.50	0.13
Wildfires	3,116	0.00	0.10	0.40	0.50	0.10
Hail	3,116	0.00	0.10	0.40	0.50	0.06
Coastal flooding	3,116	0.88	0.88	0.06	0.06	0.05
Strong winds	3,116	0.00	0.10	0.40	0.50	0.04
Ice storm	3,116	0.04	0.12	0.40	0.48	0.01
Winter weather	3,116	0.03	0.10	0.43	0.47	0.01
Avalanche	3,116	0.93	0.93	0.04	0.03	0.00
Cold wave	3,116	0.37	0.37	0.32	0.31	0.00
Heat wave	3,116	0.26	0.26	0.38	0.35	0.00
Landslide	3,116	0.00	0.81	0.10	0.10	0.00
Lightning	3,116	0.00	0.10	0.40	0.50	0.00
Tsunamis	3,116	0.99	0.99	0.01	0.01	0.00

Notes: Specifically, we define the *Low* risk category to include locations (counties or Census tracts) with zero or below the 10th percentile of annual frequency. The medium (*Mid*) risk category includes locations with annual frequency higher than the 10th percentile (hence, strictly positive) but lower than the 50th percentile *conditional* on positive annual frequency. Naturally, the *High* category contains the locations with annual frequency above the conditional 50th percentile. The last column reports the weights given to each hazard in our composite risk index, computed on the basis of each hazard's share in the expected annual loss at the national level (adding the monetary value of buildings and people). Adding the weights in the last column results in a number lower than one due to rounding

Unconditional Retreat

We now turn to the estimation of (average decadal) population growth gaps on the basis of climate risk based on Eq. 3, but this time we consider each natural hazard separately. The results are collected in the top panel of Table 10. Column 1 reproduces the estimates using the composite index, which show substantially higher population growth in high-risk counties than in low-risk ones over the period 1990-2020 (by about 2.9 log points per decade). The following columns consider all major climate hazards separately (defined as those with the highest weights in the composite index). Clearly, population growth is significantly higher in high-risk counties (relative to low-risk ones) for droughts, hurricanes, wildfires, coastal flooding and, to a lesser extent, riverine flooding. The only exceptions to this pattern are counties with high risk of tornadoes and counties with high risk of hail. In sum, we reject the unconditional retreat hypothesis for 5 out of the 7 main natural hazards.

These estimates also shed light on why the estimated excess population growth exhibited by high-risk counties is substantially *lower* when we rely on the (unweighted) composite risk index that assigns equal weights to all natural hazards (Evolution Over Time). Among the 7 natural hazards with the highest weight in our composite index (which account for 94% of

Table 10 Heterogeneity by climate hazard

AvGpop 30 year Counties AvGPop	(1) ZW	(2) Dght	(3) Hren	(4) RFld	(5) Tornds	(6) Wildfir	(7) Hail	(8) CFld
HighRisk	0.029***	0.043***	0.024***	0.012	-0.028***	0.050***	-0.074***	0.031***
	[0.008]	[0.008]	[0.008]	[0.009]	[0.009]	[0.008]	[0.009]	[0.010]
MedRisk	0.006	0.016**	0.004	0.002	0.006	0.023***	-0.068***	0.063***
	[0.006]	[0.007]	[0.008]	[0.008]	[0.009]	[0.008]	[0.009]	[0.011]
Constant	0.038***	0.022***	0.038***	0.042***	0.060***	0.014**	0.113***	0.043***
	[0.005]	[0.006]	[0.006]	[0.008]	[0.009]	[0.007]	[0.007]	[0.004]
Obs.	3,106	3,106	3,106	3,106	3,106	3,106	3,106	3,106
Counties net CZ								
HighRisk	0.005**	0.000	-0.001	0.013***	0.008**	0.006***	0.001	-0.004*
	[0.002]	[0.002]	[0.001]	[0.003]	[0.003]	[0.002]	[0.001]	[0.002]
MedRisk	-0.001	-0.001	0.001	0.006*	0.009**	0.005*	-0.001	-0.002
	[0.002]	[0.002]	[0.001]	[0.003]	[0.004]	[0.003]	[0.001]	[0.003]
Obs.	3,106	3,106	3,106	3,106	3,106	3,106	3,106	3,106
Tracts net County								
HighRisk	0.015***	0.000	0.000	0.030***	0.030***	0.062***	0.003*	-0.008*
	[0.003]	[0.001]	[0.000]	[0.006]	[0.005]	[0.013]	[0.001]	[0.005]
MedRisk	0.008**	-0.000	-0.000	0.025***	-0.013***	0.026**	-0.001	0.008**
	[0.004]	[0.001]	[0.000]	[0.005]	[0.005]	[0.012]	[0.002]	[0.003]
Obs.	57,931	57,931	57,931	57,931	57,931	57,931	57,931	57,931

Notes: The header of each column labels the natural hazard considered. The dependent variable is the average decadal change in the log of population between 1990 and 2020. The top and middle panels use the county-level dataset and the bottom panel uses the tract-level dataset. In all models the omitted category is low risk. Standard errors clustered by commuting zone. P-values: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 11 Within tract and county variability by natural hazard

	Commuting zones (1) Share of CZ where all tracts have 0 risk	Counties (2) Share of counties where all tracts have 0 risk	(3) CoV (counties with mean> 0)
Composite NRI	0.00	0.00	16.5
Composite (unweighted)	0.00	0.00	620.6
Composite	0.00	0.00	497.8
Drought	0.07	0.11	26.1
Hurricane	0.35	0.27	9.5
Riverine flooding	0.06	0.01	14.4
Tornados	0.00	0.00	**103.2**
Wildfires	0.00	0.00	66.4
Hail	0.00	0.00	6.1
Coastal flooding	0.87	0.88	**111.2**
Strong winds	0.00	0.00	7.1
Ice storm	0.03	0.03	14.0
Winter weather	0.02	0.02	6.5

Notes: These calculations are based on the tract-level dataset. The first (second) column computes the share of commuting zones (counties) for which all tracts have zero risk for the corresponding natural hazard. The third column reports the within-county dispersion of each hazard's risk. Specifically, we compute the coefficient of variation (CoV) for each county, based on the county-specific standard deviation and mean of the average annual frequency, and then average across all counties with positive mean average annual frequency. The table only includes hazards with positive weight in our composite index

all economic damage nationwide), for only 2 hazards with relatively low weights (tornados and hail) do we estimate *negative* excess growth for high-risk counties. These two hazards play an outsized role in the unweighted composite index, relative to the version that weighs each hazard on the basis of its nationwide economic damage.

Growth Relative to the Commuting Zone

We now turn to the estimation of models for county population growth net of the average for the commuting zone. By construction, this comparison neutralizes the effect of pull factors that operate at the level of commuting zones (or a higher geographical level).

The estimates are reported in the middle panel of Table 10. Two main observations stand out. First, we do not estimate any excess population growth in counties at high risk of drought, hurricane and hail. Furthermore, the sign for the coefficient for *coastal flooding* turns negative and the point estimate implies a (marginally statistically significant) 0.4 log point *lower* decadal population growth in high-risk counties relative to low-risk counties within the same commuting zone.

The vanishing of the excess population growth in counties with high risk of droughts, hurricanes and coastal flooding when using the corresponding commuting zones as benchmark indicates that the pull factors that drive population growth operate at the geographic level of commuting-zones (or higher), or that the scope of these natural hazards encompasses entire commuting zones. This could plausibly be the case for droughts and hurricanes, but does

not explain the reversal of the sign for coastal flooding, which is much more geographically localized.

Tract-level Data and Growth Relative to the County

We now turn to our tract-level dataset to examine sub-county population dynamics by natural hazard, which will allow us to investigate if micro-retreat is taking place. In other words, it will reveal whether sub-county population shifts exacerbate or mitigate the increasing exposure of high-risk counties.[37]

The bottom panel in Table 10 presents the estimates for population growth net of the county average. Two results stand out. First, we find a negative (and statistically significant at a 10% level) coefficient for the high-risk dummy variable for *coastal flooding* risk. Namely, over the last 3 decades, these tracts have grown less than other tracts within the same county (by about 0.8 log points per decade). Secondly, this is not the case for any of the other natural hazards: on the basis of within-county comparisons, tracts at high risk of droughts or hurricanes grew at the same rate as low-risk tracts, and tracts with high-risk of riverine flooding, tornadoes, wildfires or hail grew more than tracts with low risk levels for those specific hazards.[38]

In sum, for most natural hazards, high-risk counties have grown disproportionately more than low-risk counties (with the exceptions of counties with high risk of tornados or hail). When we turn to within-county, cross-tract comparisons, we find that these agglomeration dynamics are reinforced in counties with high risk of riverine flooding, tornadoes, wildfires and hail. In contrast, we do not find within-county variation in population growth on the basis of risk of drought or hurricanes. In the case of coastal flooding risk, our estimates suggest that high-risk tracts have grown less than low-risk tracts within the county.

The Micro-retreat Hypothesis

What explains the differential sub-county population dynamics in areas exposed to coastal flooding relative to other types of climate risk? We hypothesize that residents of areas with high risk of coastal flooding can reduce their risk exposure by relocating within the same county, which allows them to continue enjoying many of the same attributes. In contrast, this type of *micro-retreat* may not be feasible for residents exposed to other natural hazards.

The first step toward investigating the micro-retreat hypothesis is to determine which natural hazards entail high variation in exposure across tracts within a given county. Additionally, this variation should be easily predictable; otherwise, county residents cannot determine which low-risk tracts can provide "insurance" against that specific climate risk.

To measure the degree of cross-tract, within-county variability of each natural hazard, we follow the following 3 steps. For each county c, we first compute the mean and standard deviation (across tracts) of the average annual frequency of the climactic event, which we denote by (m_c, s_c). We then compute the coefficient of variation specific to each county c as $CoV_c = s_c/m_c$. Last, we average CoV_c across all counties (with $m_c > 0$). For instance, we expect high variability in exposure to coastal flooding within *coastal* commuting zones or counties, but much lower variability in exposure to hurricanes, which tend to impact whole counties to a similar degree (and even commuting zones).

---

[37] Table 12 shows the classification of risk categories for individual hazards at the tract level.

[38] Estimates for additional model using the tract-level dataset are reported in Table 13.

The resulting cross-tract variability measures are reported in Table 11, which considers the natural hazards used in the construction of our composite climate risk index. Column 1 reports the share of commuting zones where *all tracts* have zero risk (for the corresponding natural hazard). We observe that 87% of the commuting zones have zero exposure to coastal flooding (followed by 35% with zero risk of hurricanes), reflecting that only coastal areas are exposed to coastal flooding. In contrast, all commuting zones are exposed to tornados, wildfires, hail and strong winds. Similarly, column 2 reports on the share of counties that have no exposure to the corresponding natural hazard, meaning that all tracts in the county have zero risk. Both qualitatively and quantitatively, the results resemble column 1: 88% of counties are not exposed to coastal flooding, and 27% have no exposure to hurricanes. Clearly, only coastal counties (mostly in the Northeast and South of the country) are exposed to coastal flooding, and while hurricanes have a much larger geographical scope, large areas in the interior and north of the country have zero exposure. We turn next to column 3, which reports the coefficient of within-county variation for each of the natural hazards, which averages only the counties with positive exposure to the corresponding natural hazard. Two natural hazards stand out in terms of their within-county variability: coastal flooding ($CoV = 111$) and tornados ($CoV = 103$).

It is worth noting that there is a fundamental difference in the nature of the within-county variability for coastal flooding and for tornados. For coastal flooding, the high variability reflects the large disparity in risk for tracts on the coast and tracts in the interior of the same county. Thus, it is fairly obvious to any county resident which tracts provide "insurance" against the risk of coastal flooding. In contrast, the within-county variability of tornados has to do with the randomness of their path, which implies that no tracts in the county can be considered entirely risk-free. As a result, *micro-retreat* is only an effective way to mitigate climate risk, without losing access to county-level attributes, in the case of risk of coastal flooding. More colloquially, residents of counties with high risk of coastal flooding can 'have it both ways', that is, they can reside in low-risk tracts within attractive counties. It is worth noting that this finding is consistent with the results in Lin et al. (2021). Their analysis of residential construction in U.S. coastal areas shows that building density peaks at 2.5 km from the coast (and declines asymmetrically, falling more rapidly as we approach the waterfront).

In sum, the feasibility of micro-retreat is a plausible explanation for the pattern of estimates in the bottom panel of Table 10, which entails that county-level estimates of climate risk *overestimate* the actual risk in areas at high risk of coastal flooding, but *under-estimate* risk in locations highly exposed to some of the other natural hazards (such as riverine flooding, tornados or wildfire).

Conclusions

Our paper introduces a new composite climate-risk index designed to analyze the relationship between current climate risk and local population dynamics. The composite index has both high geographic granularity and includes all major natural hazards. While it is closely related to FEMA's National Risk Index (or NRI), cross-county (or cross-tract) variation in our composite index stems exclusively from differences in the average annual frequency of each hazard and is not mechanically related to local population levels.

On the basis of our climate risk index, we find that population is not retreating from the average county with high climate risk. In fact, since 1990, we find that high-risk counties

have grown *more* than low-risk ones (by about 2.9 log points per decade), although there are signs pointing to a recent decline in the excess population growth of high-risk counties.

Importantly, the disproportionate growth of high-risk counties remains even after netting out the average growth in the commuting zone: over the past three decades, high-risk counties grew about 0.5 percent more, per decade, than low-risk counties *within the same commuting zone*. This finding also implies that the factors that attract people to high climate-risk areas operate at more narrow (i.e., county or sub-county) geographical levels. Further, we show that the increasing population agglomeration in high climate-risk counties appears to be largely driven by white, working-age individuals. We also analyzed population dynamics at a more granular geographical level and found that high-risk Census tracts have typically grown more than low-risk tracts within the same county. This finding indicates that the county-level analysis *underestimates* the degree of population agglomeration in high-risk areas.

We also conducted heterogeneity analysis along various dimensions. First, we investigated the role of urbanization in mediating the relationship between climate risk and local population growth. We did not find increasing agglomeration in high-risk areas with *low* residential capital (in terms of scale, median value and density). In contrast, we clearly rejected the retreat hypothesis in more urbanized areas. Our finding of increasing agglomeration in high-risk, high-urbanization areas implies that the conclusions in Lin et al. (2021) extend to natural hazards other than coastal flooding. Our analysis is too descriptive in nature to fully identify the factors responsible for shifting population toward high-risk, high-urbanization local areas. However, our findings suggest that these areas may be highly productive and offer high wages. As a result, people gravitate toward those areas despite the high exposure to disruptive climate shocks. In a way, suffering damaging climate shocks once in a while is viewed as "the cost of doing business" in those locations. In a recent study, Pang and Sun (2022) provide a dynamic model with endogenous migration across multiple locations (across the Texas coastline) that differ by flood risk and productivity. The previous interpretation of our findings is consistent with the equilibrium dynamics in his setup.[39]

Our analysis has also uncovered important regional heterogeneity. In the South and Northeast of the country, high-risk countries have consistently grown more than low-risk counties due to net migration into high-risk counties. In contrast, in the Midwest and West, over the last 3 decades, net migration flows are responsible for lowering the population growth in high-risk counties below the rate of growth of low-risk counties.

Last, we analyzed whether our results vary by hazard type. For most individual hazards, we find that population growth is higher in counties with high climate risk than in counties with low risk (except for tornados and hail). Thus, the unconditional retreat hypothesis is rejected for most natural hazards. On the other hand, within commuting-zone comparisons reveal that excess population growth in high-risk counties disappears for droughts, hurricanes, hail, and coastal flooding. This implies that commuting-zone pull factors, such as strong labor markets, may explain the vigorous population growth in areas with high exposure to these types of climate risk.[40] We also find evidence of micro-retreat, namely lower population growth in high-risk Census tracts (relative to the corresponding county), in the case of coastal flooding risk, but not for the other natural hazards. We argue that this might be because counties with

---

[39] Our findings are also in line with the implications of the modeling framework in Rafols (2023), which evaluates the implications of projected warming temperatures and flood risk in the Philippines. In her model, individuals tend to migrate to locations with high consumption amenities and high productivity.

[40] Desmet et al. (2021) analyze the costs of sea-level rise using a model where individuals and firms make location decisions taking into account both agglomeration economies and mobility costs. In their setting, firm-level spillovers slow down retreat from high-risk locations.

high-coastal risk are characterized by predictable, highly localized risk. As a result, residents can 'have it all', that is, they can reside in low-risk tracts within attractive counties.

All in all, our findings show increasing agglomeration in high climate-risk areas in the South and Northeast of the United States, likely driven by robust local economies and possibly reinforced by inertia in public investments in densely population risky locations (Balboni, 2021). However, our analysis also tentatively suggests that the excess population growth in high-risk counties may be shrinking in the 2010s (Fig. 8), largely driven by intense out-migration from high-risk counties in the West. Perhaps the recent increase in insurance costs (and the exit of some of the major insurance providers in California) may have begun reshaping the mobility decisions of residents in some parts of the country.

Appendix A

Adjustments to County Population Dataset

Our main dataset is the Surveillance, Epidemiology, and End Results Program (SEER) provided by the National Cancer Institute. To address county boundary changes, we make the following adjustments:

1. Alaska and Hawaii are dropped from the dataset due to difficulties in linking county-level data over time.
2. We drop two groupings of counties (with a FIPS code) that were only used in the 1970 Census: 36910 - New York City and 51918 - Arling/Alexan/Fairfax/Falls Church.
3. We impute some county population values for a few counties using linear interpolation. In particular, we use the 1960 and 1990 values to impute values for 1970 and 1980 for counties with FIPS 35061, 51095, 51153, 51830. And we use the 1960 and 1980 value to impute the missing 1970 value for counties with FIPS 51165, 51177, 51179, 51199, 51510, 51580, 51610, 51630, 51660, 51690.
4. For a few counties in Virginia (FIPS 51683, 51685, 51735 and 51830), the 1982 data is available, but the 1980 is not. We use the 1982 value for 1980.
5. Last, we note that there are missing values for a few counties in Virginia that cannot be interpolated because we lack values after 1970, but we keep them in the data. In any case, they have small populations and they will not appear in the estimation sample.
6. More details on the construction of the SEER data can be found at: https://seer.cancer.gov/seerstat/variables/countyattribs/ruralurban.html.

We obtained the county-level net migration data and the counterfactual population counts in the absence of net migration from Egan-Robertson et al. (2023). The data contain estimates of net migration for the 1950s through the 2010s, and can be disaggregating by gender and (5-year) age groups. The data are freely available and detailed documentation is available online at https://netmigration.wisc.edu.

Net migration is estimated as a residual, by subtracting the observed population (in the Census) at the end of the decade from the counterfactual ("expected") population in the absence of migration. In turn, the latter is estimated as follows: the actual population at the beginning of the decade is aged forward, subtracting out deaths and adding in births, to generate an "expected population" at the end of the decade. The deaths and births are obtained from vital statistics records. Next, we highlight a few important points regarding this dataset:

1. At the time we downloaded the dataset (December 2023), the most updated version of the net migration dataset is based on a preliminary version of the 2020 population Census.

As explained in Egan-Robertson et al. (2023), the authors of the net migration dataset felt that the estimates for the 2010-2020 change in the population age 75 and older at the county level were unreliable and decided to eliminate this age group, which introduces a discontinuity with the data for the previous decades. To address this issue and maintain comparability, we removed this age group from all previous decades in our analysis of the role of net migration.

2. We verified that the population counts in the SEER and Net Migration datasets were consistent with each other. We were able to match exactly the values in 1970. For other years, the values were fairly close, though not exactly the same. More details can be provided upon request.

3. The county-to-county merge between the SEER and net migration datasets was successful for over 99 percent of the counties. Practically all counties that did not merge belonged to Hawaii, Alaska or Virginia, because we had dropped them already from the SEER dataset (as explained earlier in this section).

Appendix B

Adjustments to Tract Population Data

Our Census-tract population data is the Longitudinal Tract Data Base Census Dataset (LTBD) available at https://s4.ad.brown.edu/Projects/Diversity/Researcher/Bridging.htm. It combines data from the decennial Census and the ACS, harmonized to 2010 Census tract boundaries as described in Logan et al. (2014). The data covers the period 1950-2020.

We use the full-count (standard) dataset. The specific sources for the population counts (and age and race variables) in our analysis are as follows: Total population for the 1970-2020 period is obtained from 1970 Census (Count 2, 100% Data, T1), 1980 (STF1, 100% Data, T1), 1990 (STF1, 100% Data, P1), 2000 (SF1, 100% Data, P1), 2010 (SF1, 100% Data, P1), 2020 (PL94, 100% Data). While the geographic coverage of the LTDB is very comprehensive, some tracts are missing for the following reasons:

1. In 1970, some areas did not belong to any Census tracts. They were only included in block numbering areas. These areas were not included in the LTDB crosswalk.

2. Some tracts had zero population in one (or more) of the following years: 1970, 1980, 1990, 2000. These tracts were not omitted from the LTDB crosswalk to year-2010 tracts.

3. Some tracts were populated entirely by crews of vessels in one (or more) of the following years: 1970, 1980, 1990. These tracts were not omitted from the LTDB crosswalk to year-2010 tracts.

Appendix C

FEMA data on Hazard Occurrence and Annual Frequency

As part of their efforts to construct a National Risk Index, FEMA calculates the *annual probability* with which a given natural hazard will occur for each Census block in the country. This is done separately for each of the 18 natural hazards in order to adapt the methodology to the nature of each hazard. In most cases, the annualized block-level frequency is based on the count of hazard occurrences in polygons that intersect the corresponding block. However,

for widespread hazards, like strong winds, hurricanes and ice storms, a $49km^2$ fishnet grid is cast to count the number of hazard occurrences in each cell. Each Census block contained in a cell gets assigned the same total number of occurrences for these widespread hazards.

What is considered a *hazard occurrence* differs across types of hazard. For some hazards, the hazard occurrence count is based on the number of distinct events, whereas for other hazard types, the hazard occurrence is based on the count of days a hazard has lasted. Table 14 collects the information for each hazard type and further below we discuss the detailed calculation of occurrences and annual frequencies for the main natural hazards in our composite index.

Once the tally of the total number of occurrences has been calculated for each hazard in each Census block, it is straightforward to calculate the annualized frequency (AF):

$$AF = \frac{\text{Number of Recorded Hazard Occurrences}}{\text{Recording Period Years}} \tag{C.1}$$

To produce annual frequencies at the Census tract or county levels, FEMA aggregates the Census block values using area weights:

$$\text{Census Tract AF} = \frac{\sum(\text{Census Block AF} \times \text{Area of Census Block})}{\text{Area of Census Tract}} \tag{C.2}$$

$$\text{County AF} = \frac{\sum(\text{Census Block AF} \times \text{Area of Census Block})}{\text{Area of County}} \tag{C.3}$$

In what follows, we go into more detail in terms of the process for calculating the annualized frequency for the hazards with the most significant weight in our Composite Risk Index measure. The following five natural hazards make up 83% of the weighted Composite risk index.

Hurricanes The total aggregate number of Hurricane occurrences is taken from the HURDAT2 dataset. This dataset registers all storms between 1851-2020 for counties along the Atlantic ocean and between 1949-2020 for those along the Pacific ocean. Based on its associated wind speeds, each registered Hurricane is assigned a category.[41] Using tracking data, a multi-segment line is created for each storm that tracks its geographical spatial location as well as the storm category in each segment. These hurricane event paths are then intersected with Census block polygons to help determine the number of hurricanes suffered by each Census block.

The aggregate count of hurricanes is then divided by the total number of years for the period over which the hurricanes occurred in order to calculate the annualized frequency for each Census block. These values are then used to construct the area-weighted annualized frequencies for both Census tracts and counties using Eqs. C.2 and C.3, respectively.

Droughts The University of Nebraska-Lincoln National Mitigation Center (NDMC) U.S. Drought Monitor provides historical data on the areas that have experienced droughts on a weekly basis since year 2000. Each drought event is then categorized by the severity of the event, with an indicator that goes from abnormally dry (D0) to exceptional drought (D4). The sum of the annualized recorded Drought occurrences in a given Census tract, in event-days, over the number of years in the recording period constitutes the annualized frequency for a particular Census tract. These Census tract measures are then used to construct the area-weighted annualized frequency values at the County level using Equation refeq:AFcounty.

Riverine Flooding FEMA relies on the NCEI Storm Events Database to record each riverine flooding event for the period 1996-2019. The number of days in which riverine

---

[41] These range from Topical Storms, which are the least violent, to Category 5 hurricanes which have wind speeds above 157 mph.

flooding events occurred for counties (and Census tracts) that intersect a 1% annual chance riverine floodplain (as determined by FEMA's National Flood Insurance Program) are then tallied. The annualized frequency for each county is the calculated as the number of riverine flooding occurrences, in event-days, over the period of record. Census tracts simply inherit the annualized frequency occurrence from the county they belong to.

Wildfires FEMA use a series of raster datasets created by the US Forest Service Missoula Fire Sciences Laboratory that assess both the burn probability (BP) and the conditional fire intensity level (FIL) at different locations throughout the US. These raster files work in parallel; each of them at a 270-meter grid spatial resolution. On one hand, the BP raster dataset assesses the probability that a particular area is exposed to a large fire, defined as a fire that "escapes initial fire suppression and spreads". On the other hand, the FIL dataset contains six independent raster files, each of which determines the portion of simulated fires in a particular area that reach a specified intensity. FEMA developed a a custom raster-vector intersect tool to determine the intersections of the raster cells with Census blocks. The annualized frequency is then calculated as the area-weighted BP for every Census block at a given year. These Census block estimates are then aggregated to the Census tract and County levels using area-weighted Eqs. C.2 and C.3, respectively.

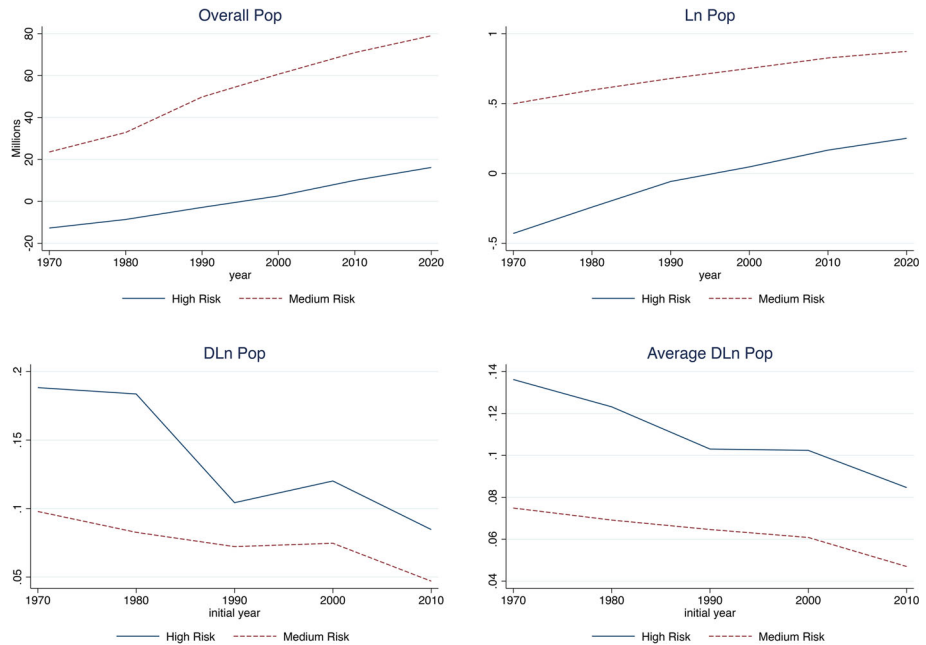## Appendix D

Figures and Tables

Fig. 9 National trends (pooling tracts) by climate risk. **Notes:** Census tract dataset. Composite risk with hazard weights. Population trends in United States by climate risk, net of the corresponding value for the low-risk category. Top-left figure is in millions of individuals, top-right figure in logs, bottom-left figure is the 10-year log change between the year indicated in the horizontal axis and 10 years later, and the bottom-right figure is the average decadal growth between the year indicated in the horizontal axis and year 2020
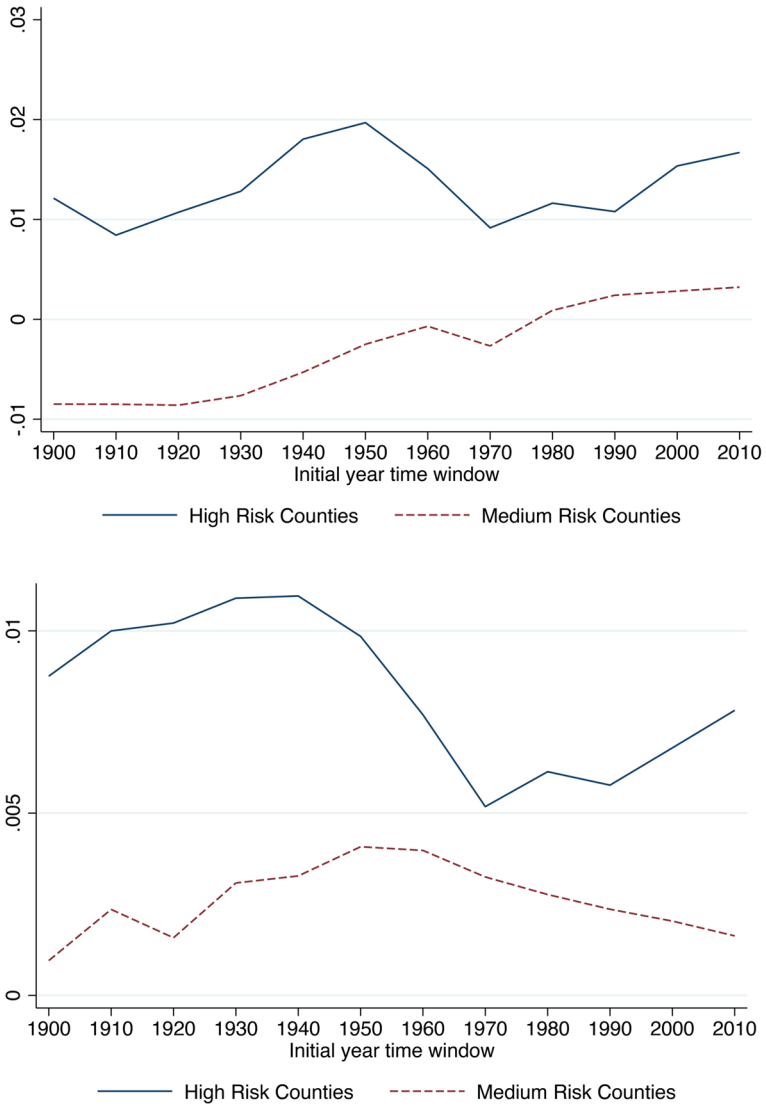
Fig. 10 Evolution population growth differential. Unweighted composite risk index. County population growth (top); demeaned by CZ average growth (bottom). **Notes:** Risk categories based on unweighted composite risk index (Z). Each point estimate refers to the average decadal change in the log of population between the corresponding initial year and final year 2020. Point estimates obtained from models for the average decadal population growth (top) and for the same variable but demeaned using average growth in the corresponding commuting zone (bottom). In all cases the omitted category are counties with low risk

Table 12 Risk categories individual hazards (tracts)

Climate hazard	Tracts	Zeros	Low	Mid	High	Weights
Composite NRI	58,488		0.69	0.20	0.11	
Composite (unweighted)	58,488		0.25	0.50	0.25	
Composite	58,488		0.25	0.50	0.25	
Drought	58,488	0.20	0.20	0.40	0.40	0.21
Hurricane	58,488	0.14	0.14	0.44	0.42	0.21
Riverine flooding	58,488	0.20	0.20	0.40	0.40	0.18
Tornados	58,488	0.00	0.12	0.38	0.50	0.13
Wildfires	58,488	0.15	0.15	0.42	0.42	0.10
Hail	58,488	0.00	0.10	0.41	0.49	0.06
Coastal flooding	58,488	0.75	0.75	0.12	0.12	0.05
Strong winds	58,488	0.00	0.10	0.40	0.50	0.04
Ice storm	58,488	0.07	0.12	0.42	0.46	0.01
Winter weather	58,488	0.08	0.11	0.44	0.45	0.01
Avalanche	58,488	0.95	0.95	0.03	0.02	0.00
Cold wave	58,488	0.43	0.43	0.32	0.24	0.00
Heat wave	58,488	0.22	0.22	0.40	0.38	0.00
Landslide	58,488	0.00	0.98	0.00	0.02	0.00
Lightning	58,488	0.00	0.10	0.40	0.50	0.00
Tsunamis	58,488	1.00	1.00	0.00	0.00	0.00

Notes: Specifically, we define the *Low* risk category to include locations (counties or Census tracts) with zero or below the 10th percentile of annual frequency. The medium (*Mid*) risk category includes locations with annual frequency higher than the 10th percentile (hence, strictly positive) but lower than the 50th percentile *conditional* on positive annual frequency. Naturally, the *High* category contains the locations with annual frequency above the conditional 50th percentile. The last column reports the weights given to each hazard in our composite risk index, computed on the basis of each hazard's share in the expected annual loss at the national level (adding the monetary value of buildings and people)

Table 13 Heterogeneity by climate hazard. Tracts-analysis

30 year AvGPop	(1) ZW	(2) Dght	(3) Hrcn	(4) RFld	(5) Tornds	(6) Wildfir	(7) Hail	(8) CFld
HighRisk	0.090***	0.122***	-0.052	0.049**	-0.022	0.117***	-0.133***	0.006
	[0.023]	[0.021]	[0.036]	[0.020]	[0.046]	[0.019]	[0.050]	[0.014]
MedRisk	0.028	0.037**	-0.095***	0.035	-0.089*	0.038**	-0.146***	0.015
	[0.018]	[0.014]	[0.037]	[0.023]	[0.046]	[0.017]	[0.049]	[0.017]
Constant	0.062***	0.035***	0.162***	0.065***	0.143***	0.033*	0.224***	0.096***
	[0.019]	[0.013]	[0.034]	[0.023]	[0.046]	[0.020]	[0.048]	[0.012]
Observations	57,931	57,931	57,931	57,931	57,931	57,931	57,931	57,931
net CZ								
HighRisk	0.017***	0.002	-0.000	0.049***	0.037***	0.087***	0.004	-0.012*
	[0.005]	[0.001]	[0.001]	[0.009]	[0.006]	[0.017]	[0.002]	[0.006]
MedRisk	0.011*	0.000	0.000	0.041***	-0.026***	0.043**	-0.001	0.004
	[0.006]	[0.001]	[0.001]	[0.010]	[0.006]	[0.017]	[0.003]	[0.005]
net County								
HighRisk	0.015***	0.000	0.000	0.030***	0.030***	0.062***	0.003*	-0.008*
	[0.003]	[0.001]	[0.000]	[0.006]	[0.005]	[0.013]	[0.001]	[0.005]
MedRisk	0.008**	-0.000	-0.000	0.025***	-0.013***	0.026**	-0.001	0.008**
	[0.004]	[0.001]	[0.000]	[0.005]	[0.005]	[0.012]	[0.002]	[0.003]

Notes: The header of each column labels the natural hazard considered. The dependent variable is the average decadal change in the log of population between 1990 and 2020. All panels based on the tract-level dataset. In all models the omitted category is low risk. Standard errors clustered by commuting zone. P-values: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 14 Geographic Level and Count Determination of Hazard Occurrence

Hazard Type	Hazard Occurrence Basis	Geographic Level of Occurrence
Avalanche	Distinct events	County
Coastal Flooding	No event count	No event count
Cold Wave	Event-days	Census block
Drought	Event-days	Census tract
Hail	Distinct events	$49km^2$ fishnet
Heat wave	Event-days	Census block
Hurricane	Distinct events	$49km^2$ fishnet
Ice storm	Event-days	$49km^2$ fishnet
Landslide	Distinct events	Census tract
Lightning	Distinct events	$4km^2$ fishnet
Riverine Flooding	Event-days	County
Strong Wind	Distinct events	$49km^2$ fishnet
Tornado	Distinct events	$49km^2$ fishnet
Tsunami	Distinct events	Census tract
Wildfire	No event count	No event count
Winter Weather	Event-days	Census block

Notes: Our own elaboration based on FEMA (2020)

Declarations

Competing interests The authors declare no competing interests.

References

Amornsiripanitch N, David W (2023) "Who bears climate-related physical risk?," FRB of Philadelphia Working Paper No. 23–29

Balboni C (2021)"In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities," Technical Report, Mimeo MIT

Baylis P, Boomhower J (2023) The Economic Incidence of Wildfire Suppression in the United States. Am Econ J: Appl Econ 15(1):442–73

Beeson PE, David ND, Werner T (2001) Population growth in U.S. counties, 1840–1990. Regional Sci. Urban Econ. 31(6):669–699

Black R, Bennett SRG, Thomas SM, Beddington JR (2011) Migration as adaptation. Nature 478(7370):447–449

Boustan, LP, Kahn ME, Rhode PW, Yanguas ML (2020) "The effect of natural disasters on economic activity in US counties: A century of data." J Urban Econ 118(C)

CEMHS (2022) "The Spatial Hazard Events and Losses Database for the United States," ASU Center for Emergency Management and Homeland Security, Version 20.0 Online Database. Arizona State University

Deryugina T, Kawano L, Levitt S (2018) The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns. Am Econ J: Appl Econ 10(2):202–233

Desmet K, Kopp RE, Kulp SA, Nagy DK, Oppenheimer M, Rossi-Hansberg E, Strauss BH (2021) Evaluating the Economic Cost of Coastal Flooding. Am Econ J: Macroecon 13(2):444–486

Egan-Robertson D, Curtis KJ, Winkler RL, Johnson KM, Bourbeau C (2023) Age-specific net migration estimates for US counties, 1950–2020. University of Wisconsin-Madison, Applied Population Laboratory

FEMA (2020) National Risk Index: Primer. Technical Report, FEMA

Fussell E, Curran SR, Dunbar MD, Babb MA, Thompson L, Meijer-Irons J (2017) Weather-Related Hazards and Population Change: A Study of Hurricanes and Tropical Storms in the United States, 1980–2012. Ann Am Academy Political Soc Sci 669(1):146–167

Glaeser EL, Kolko J, Saiz A (2001) "Consumer city," J Econ Geography 01 , 1(1):27–50

Hauer ME, Fussell E, Mueller V, Burkett M, Call M, Abel K, McLeman R, Wrathall D (2020) Sea-level rise and human migration. Nature Rev Earth Environ 1(1):28–39

Hauer ME, Sunshine Jacobs, and Scott Kulp, "Climate Migration Amplifies Demographic Change and Population Aging." Technical Report, Mimeo Florida State University February

Indaco A, Ortega F, Taspınar S (2021) Hurricanes, flood risk and the economic adaptation of businesses. J Econ Geography 21(4):557–591

Jia N, Molloy R, Smith C, Abigail W (2023) The Economics of Internal Migration: Advances and Policy Questions. J Econ Literat 61(1):144–80

Lin Y, McDermott TKJ, Guy M. (2021) "Cities and the Sea Level," IZA Discussion Papers 14243, Institute of Labor Economics (IZA)

Logan JR, Xu Z, Stults B (2014) "Interpolating U.S. Decennial Census Tract Data from as Early as 1970 to 2010: A Longtitudinal Tract Database." The Professional Geographer : the Journal of the Association of American Geographers 07 , 66(3):412–420

Ortega F, Tapınar S (2018) "Rising sea levels and sinking property values: Hurricane Sandy and New York's housing market." J Urban Econ 106(C):81–100

Pang X, Sun P (2022) Moving Into Risky Floodplains: The Spatial Implication of Flood Relief Policies. Penn State University, Mimeo

Partridge MD, Bo F, Rembert M (2017) Improving Climate-Change Modeling of US Migration. Am Econ Rev 107(5):451–455

Petkov I (2022) "Weather Shocks, Population, and House Prices: the Role of Expectation Revisions." Economics of Disaster and Climate Change, Forthcoming

Petkov I, Ortega F (2023) Flood risk and Insurance Take-up in the Flood Zone and its Periphery. Technical Report, Mimeo CUNY, Queens College

Radeloff VC, Helmers DP, Kramer HA, Mockrin MH, Alexandre PM, Bar-Massada A, Butsic V, Hawbaker TJ, Martinuzzi S, Syphard AD, Stewart SI (2018) Rapid growth of the US wildland-urban interface raises wildfire risk. Proceedings of the national academy of sciences 115(13):3314–3319

Rafols R (2023) Economic consequences of climate change: evidence from the Philippines. Syracuse University, Mimeo

Rappaport J, Sachs JD (2003) The United States as a Coastal Nation. J Econ Growth 8(1):5–46

Wilson SG, Fischetti TR (2010) "Coastline Population Trends in the United States: 1960 to 2008 ." Current Population Reports P25–1139, U.S. Census Bureau