Climate Emigration: Evidence from the United States from 2011-2021

Akshay Yeddanapudi

School of Global Policy and Strategy, UC San Diego

April 9, 2024

Abstract

I study the effects of climate-related shocks on emigration within the US. I develop a theoretical framework under which a Bayesian hypothetical migrant will update their expectation about the number of events they will experience in a given location and will choose to emigrate as a function of these changes. I empirically show that, in a destination ambivalent setting, there is a statistically significant and positive relationship between the number of emigrant households and the number of climate shocks, of any type, in the origin county. In a dyadic setting I find that a similar statistically significant and positive relationship exists between the number of households emigrating to a given destination and the number of climate shocks to the origin county. Curiously, I also find a positive, increasing and statistically significant relationship between the number of destination county climate shocks, prior to emigration, and the number of emigrant households. I subsequently utilize migration as an instrument by which climate shocks to the origin county can influence mean annual wages in the destination county, in the year post emigration, and find that the results are inconclusive. The inclusion of origin-destination fixed effects eliminating any statistically significant relationship while the omission of such fixed effects lead the relationship to be statistically significant and positive.

1 Introduction

Recent literature has demonstrated that climate change has altered the frequency and intensity of precipitation patterns as well as the frequency and intensity of tropical storms and cyclones (Xi, Lin, and Gori 2023, Knutson et al. 2021). Other work has highlighted the effects of climate change on the increasing risks of wildfires (Westerling 2016, Turco et al. 2023) and on changes to snow cover patterns (Mitterwallner et al. 2024). Concurrently, there exists an extensive literature analyzing the effects of climate shocks on domestic and interntional migration. The majority of these studies utilize temperature or precipitation fluctuations to proxy for climate related shocks before then examining the effect on migration (Cattaneo and Peri 2015). Recent work focused on the United States has shown that emigration is responsive prolonged variations in temperature (Mullins and Bharadwaj 2021). Given the reliance of temperature or precipitation fluctuations to proxy for climate related shocks, the rational question of whether aggregate storm or weather event shocks, disparate from precipitation or temperature measurements, can explain climate change induced migration. Moreover migration, in response to climate change, has also been shown to be damped, the US, in response to wage feedback loops (Fan, Fisher-Vanden, and Klaiber 2018). This raises a secondary question of whether climate-shock induced migration alters this feedback systematically over time, by altering wages in the destination county via an expanding labor pool in counties experiencing net immigration. I tackle both these problems.

I first construct a theoretical process detailing how a Bayesian individual might update their beliefs about the expected number of climate-related shocks of type m they would experience in a finite time period in a given location. I then define a condition on the rationality of migration and link the individual's perception of the expected number of climate-related shocks to the place based utility that they derive from a given location. Under the assumptions of the propositions I put forth, decreases in place based utility should increase the set of "acceptable" locations to which an individual could migrate to, thus making emigration more likely. Conversely increases in place based utility lead to decreased likelihoods of emigration. To empirically test this claim and to answer the first of the 2 questions raised earlier, I construct both a destination ambivalent and dyadic, gravity-style model of migration. I consider only domestic emigration within the US to minimize dampening effects of temporal budget constraints on migration patterns. Unlike much of the existing literature, I utilize National Oceanic and Atmospheric Data on the occurrence of storm and other weather events as an indicator of a climate-related shock and map migration using the IRS US Population Migration Dataset. Under the destination ambivalent framework, I specific a pooled, binned and type disaggregated set of models to test for the effects of origin county shocks on total emigration. I then utilize a similarly defined pooled and binned model in the dyadic setting to test for the effects of origin and destination county shocks on dyadic emigration from the origin to destination county. To answer the second question raised - namely the effect of origin county climate shocks on destination county wages - I utilize migration as an instrument and estimate a 2 stage least squares model. As I consider year on year changes, the time horizons are sufficiently short to suppose that labor has enough time to update beliefs about the place based utility and emigrate while firms do not due to their relatively higher fixed costs. Moreover, a consideration of year on year changes, as opposed to a shorter time horizons, allows for the "averaging-out" of fluctuations in wages due to damages from climate-related shocks. Working at the year on year time horizon, I thus avoid a maximal amount of bias introduced by both hyper-short run fluctuations and long run trends, thus providing the best possible environment for an estimation of the causal effect of climate-related shocks on destination county wages.

In the broader economic context, I make 2 contributions to literature. First, I show how, in a partial equilibrium setting, Bayesian updating with regards to the expected number of climate related shocks can alter place based utility and thus alter the rationality of migration between any pair of counties. Under this setting, I demonstrate that the probability of emigration, from any given origin county, is decreasing with regards to the expected number of climate related shocks which arise from individual level Bayesian updating. Second, I provide evidence that is not solely related to temperature or precipitation fluctuations, that climate shocks have a statistically significant and mostly positive effect on emigration such that a county which experiences a greater number of climate shocks experiences a greater outflow of migrants. This work most closely follows the path set by Mullins and Bharadwaj 2021 as I restrict the focus of this analysis to be domestic emigration. This allows me to minimize biases arising from varied barriers to migration, such as national borders, capital constraints, institutions, labor considerations etc. I also contribute to existing literature in the vein of Catteneo and Perri 2015 and Chen, Oliva and Zhang 2017 who respectively show the effects of temperature and air pollution shocks on migration in a international context (Catteneo and Perri 2015, Chen, Oliva and Zhang 2017).

The remainder of this paper is structured as follows. Section 2 provides a brief discourse about the rationality of migration given a Bayesian potential migrant who updates their beliefs about the expected number of climate shocks they will experience over a fixed interval in a given location. Section 3 introduces my data and details all data cleaning, the methodology for the construction of the dyadic and destination ambivalent working data sets and provides brief descriptive statistics. Section 4 specifies the models I use to empirically test the effect of climate-related shocks on migration and the effect on destination county wages though an emigration instrument. Section 5 discusses the results in the destination ambivalent setting. Section 6 does the same in the dyadic setting. Section 7 details the results of robustness checks for heterogeneous treatment effects and false detection rate corrections. Section 8 concludes.

2 Emigration Given Updated Climate Beliefs

I begin by constructing a theoretical processes detailing how an individual, given prior beliefs about the number of climate-related shocks they will experience in a given location, "learns" as they experience new climate-related shocks, thus updating their belief about the number of climate-related events they will experience, in the same location in the future. I show mathematically that under a model of migration wherein the individual only migrates if there is at least null change in utility, this affects the locations to which they would rationally choose to migrate to.

2.1 Framework

Suppose that a representative individual l, in location i at time t has a prior belief about the distribution of the frequency λ_m with which location i experiences a climate-related shock of type m such that $\lambda_m \sim Gamma(s_m, r_m)$. Suppose also that individual l remains in location i over some discrete time interval $[t, t+1)$ which can be decomposed into n disjoint sub-intervals over which individual l observes following set of climate shocks $\{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\}$ such that $y_{m,k}$ is the number of climate-related shocks of type m that occurs during the k-th sub-interval. at time $t+1$, individual l updates their prior belief about the distribution of λ_m such that their posterior distribution on λ_m, at $t+1$ is then

$$f_1(\lambda_m | \{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\}) \propto g(\{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\} | \lambda_m) f_0(\lambda_m)$$

where each of the $y_{m,k}$'s are i.i.d draws such that $\forall 1 \leq k \leq n,\ y_{m,k}|\lambda_m \sim Poisson(\lambda_m)$ and where $f_0(\lambda_m)$ is the corresponding probability density function to the Gamma distribution. Note then that over the interval $[t+1, t+2)$, $f_1(\lambda_m|\{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\})$ is individual l's prior.

Proposition 1: *Let $\tau > t+1$ and take $\lambda_m \sim Gamma(s_m, r_m)$ where r_m and s_m are endogenous hyperparameters on the distribution of λ_m. Then the total number of events, of all types, that individual l expects at location i between $[t+1, \tau]$ is denoted $\mathcal{N}_{i,t+1}$ such that*

$$\mathcal{N}_{i,t+1} = (\tau - (t+1)) \sum_{m=1}^{p} \left(\frac{s_m + \sum_{k=1}^{n} y_{m,k}}{r_m + n} \right) \tag{1}$$

Proof: See Appendix A.1

Proposition 1 then illustrates how exposure to types of climate events m influence the belief that individual l has about the number of events of type m that will strike location i in $[t+1, \tau]$ as well as the total number of events which will occur. As each event bears a cost on the individual, it is rational to suppose the existence of a place-based utility function for individual l, denoted $u_t(i)$ which is implicitly a function of the expected number of events that will occur in a given location, i. Take $u_t(i)$ to define l's utility of living in location i at time t such that

$$u_t(i) = A_t(i)c_t(i)z_t(i)$$

where $A_t(i)$ denotes the amenity value of location i at time t, $c_t(i)$ denotes the utility of consumption and $z_t(i)$ denotes wages. Let $A_t(i)$ be linearly decomposed into a climate and non-climate component, such that

$$A_t(i) = A_t^c(i) + A_t^{nc}(i)$$

where $A_t^c(i)$ denotes the climate influenced component of the amenity value and $A_t^{nc}(i)$ denotes the non-climate component. In expectation, changes in $\mathcal{N}_{i,t+1}$ that then arise due to the individual updating their beliefs should then alter $u_t(i)$ only through changes in the climate-related amenity value $A_t^c(i)$. Given this construction of $u_t(i)$, Proposition 2 then links the change in l's expected number of events to a change in the number of possible locations to which l will rationally choose to migrate to.

Proposition 2: *Take $\mathcal{N}_{i,t}$ as defined in proposition 1 and suppose that individual l's perception of climate-related amenity values of location i evolve such that*

$$A_{t+1}^c(i) = A_t^c(i) \cdot h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) \tag{2}$$

where $h : \mathbb{R}^2 \to \mathbb{R}$ is any continuous and differentiable function with the following properties:

$$h(\mathcal{N}_{i,t} = k, \mathcal{N}_{i,t-1} = j) \geq 0$$

$$h(\mathcal{N}_{i,t} = \mathcal{N}_{i,t-1} = k) = 1$$

$$\frac{\partial h}{\partial \mathcal{N}_{i,t}} < 0$$

$\forall j, k \in \mathbb{R}$. Let L denote the set of all locations to which l can migrate to. Suppose that at each time interval $t+k$ for $k \in \mathbb{Z}^+$, individual l randomly samples location $j \in L$ and chooses to migrate if

$$u_t(i) \leq u_t(j) - C(i,j) \tag{3}$$

where $C(i,j)$ is the constant cost of moving from i to j. As $A_t(i), c_t(i), z_t(i) > 0$ and assuming constant consumption and wages in location i between t and $t+1$, the set of locations to which it is rational to for individual l to migrate to at time $t+1$, $\mathcal{F}_{i,t+1} \subset L$, will decrease if

$$h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1 \tag{4}$$

Conversely, $\mathcal{F}_{i,t+1} \subset L$, will increase if

$$h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1 \tag{5}$$

Proof: See Appendix A.2

The assumption of constant consumption and wages in location i, between t and $t+1$, while strong, is reasonable in the context of developed countries given the existence of robust social safety nets. Likewise, so is the assumption of relatively constant costs of migration between counties. A direct prediction of Proposition 2 is then the idea that as shocks, in location i will lead individual l to update $\mathcal{N}_{i,t-1}$ at time $t+1$ to be $\mathcal{N}_{i,t}$ and that if $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1$, the individual will be have a fewer number of "acceptable" locations to move to. Moreover, note that in the frictionless setting, $C(i,j) = 0$ and thus individuals will always move to the utility maximizing location. The introduction of non-negative frictions $C(i,j)$ implies that some individuals may not utility maximize since, at at a given time t', $u_{t'}(j) - u_{t'}(i) \leq C(i,j)$ thus making migration irrational.

Proposition 2, while discussing the conditions surrounding the rationality of migration, does not describe the probability with which a representative agent will emigrate. I turn now to this problem. To model the migration decision, suppose that individual l chooses some $j \in L$ with probability $\mathbb{P}(j)$, at each time $t+k$ $k \in \mathbb{Z}^+$, subject to the constraint that $u_{t+k}(j) > u_{t+k}(m) \Rightarrow \mathbb{P}(j) \geq \mathbb{P}(m)$. Define the sets $S_{1,t+k}, S_{2,t+k}, ..., S_{|L|,t+k}$ such that

$$S_{1,t+k} := \arg\max_{j \in L} u_{t+k}(j)$$

$$S_{2,t+k} := \arg\max_{j \in L - S_{1,t+k}} u_{t+k}(j)$$

$$\vdots$$

$$S_{|L|,t+k} := \arg\min_{j \in L} u_{t+k}(j)$$

where $\mathbb{P}(s \in S_{1,t+k}) = p_1,\ \mathbb{P}(s \in S_{2,t+k}) = p_2, ..., \mathbb{P}(s \in S_{|L|,t+k}) = p_{|L|}$ are constant and sum to 1.

This construction implies that while the probability of any fixed location, $j \in L$, being randomly selected as the migration destination can change discontinuously over integer time intervals, the magnitude ordered probabilities remain the same across time. As such $\mathbb{P}(s \in S_{1,t+k}) = p_1$, $\mathbb{P}(s \in S_{2,t+k}) = p_2, \ldots$ are time invariant with only the elements of the sets $S_{1,t+k}, S_{2,t+k}, \ldots$ change as a function of time. Moreover note that the constraint that $u_{t+k}(j) > u_{t+k}(m) \Rightarrow \mathbb{P}(j) \geq \mathbb{P}(m)$ implies that individual l is more likely to select locations with higher utilities. This does not preclude locations with a utility $u_{t+1}(j) < u_{t+1}(i)$, for $j \in L$, from being selected by individual l. Rather, allowing for some $j \in L$, for which $u_{t+1}(j) < u_{t+1}(i)$, to be selected with some probability $\mathbb{P}(j) > 0$ allows for uncertainty in the information that individual l has about the place-based utility of each location.

Now take $i \in L$ to be origin location for individual l and suppose that $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1$. Since $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1 \Rightarrow |\mathcal{F}_{i,t}| \geq |\mathcal{F}_{i,t+1}|$ and $\mathcal{F}_{i,t+1} \subseteq \mathcal{F}_{i,t} \subset L$, note that if $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1$ then

$$\sum_{j \in \mathcal{F}_{i,t+1}} \mathbb{P}_i(j) \leq \sum_{j \in \mathcal{F}_{i,t}} \mathbb{P}_i(j) \leq \sum_{j \in L} \mathbb{P}_i(j) = 1$$

For simplicity of notation, let

$$\mathbb{P}_i(\mathcal{F}_{i,t+k}) = \sum_{j \in \mathcal{F}_{i,t+k}} \mathbb{P}_i(j)$$

for all $k \in \mathbb{Z}^+$. Then $\mathbb{P}_i(\mathcal{F}_{i,t+1})$ denotes the probability of individual l selecting any location to emigrate to for which migration is rational, as defined in Proposition 2. Thus $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1 \Rightarrow \mathbb{P}_i(\mathcal{F}_{i,t+1}) \leq \mathbb{P}_i(\mathcal{F}_{i,t})$ and thus the individual is less likely to migrate from location i at time $t+1$. Since the statement is symmetric for all $k \in \mathbb{Z}^+$, a similar statement holds true for all times $t+k$. Alternatively, suppose that $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1$. Then by Proposition 2, $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1 \Rightarrow |\mathcal{F}_{i,t}| \leq |\mathcal{F}_{i,t+1}|$ and thus $\mathcal{F}_{i,t} \subseteq \mathcal{F}_{i,t+1} \subset L$. Then, given $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1$, note that

$$\mathbb{P}_i(\mathcal{F}_{i,t}) \leq \mathbb{P}_i(\mathcal{F}_{i,t+1}) \leq \sum_{j \in L} \mathbb{P}_i(j) = 1$$

and thus the individual is more likely to migrate from location i at time $t+1$, relative to at time t. Consequently, it follows that changes in the expected number of climate shocks then alters the probability of an individual emigrating from a given location wherein an increase in the number of shocks expected imply a greater probability of migration while a decrease in the number of shocks expected implies the converse. This constitutes the major theoretical prediction of the the model presented here. Note that this is independent of any functional specification of the the probability measure $\mathbb{P}$, on the probability space $(L, \mathcal{G}, \mathbb{P})$, beyond the assumption that the probability measure $\mathbb{P}(j)$ is monotonically increasing with respect to $u_t(j)$.

The above claim can be taken further to consider changes in the expected number of emigrants from the initial fixed location i. In particular, since individual l is assumed to be a representative agent, suppose that there exist $M_{i,t}$ other similar individuals at location i and at time t who also make a similar migration decision given similar preferences and having undergone a similar update to $\mathcal{N}_{i,t}$. As the migration decision across $M_{i,t}$ similar individuals can then be modeled using a *Binomial* $(M_{i,t}, \mathbb{P}_i(\mathcal{F}_{i,t+1}))$ distribution the expected number of individuals who choose to emigrate from location i at time $t+1$ is $M_{i,t}\left(\mathbb{P}_i(\mathcal{F}_{i,t+1})\right)$. Proposition 3 then discusses 2 special cases where the time evolution of the population, in given location i, is monotonic.

Proposition 3: *Let $i \in L$ be an arbitrary fixed location. The following claims hold true.*

1. *If $M_{i,0} \geq M_{i,1} \geq ... \geq M_{i,t}$, where $M_{i,0}$ is exogenously given, and if $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1$ for all $t > 1$ then the number of expected emigrants from location i is weakly decreasing.*

2. *If $M_{i,0} \leq M_{i,1} \leq ... \leq M_{i,t}$, where $M_{i,0}$ is again exogenously given, and if $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1$ for all $t > 1$, then the number of expected emigrants from location i is weakly increasing.*

Proof: See Appendix A.3

Proposition 3 then claims that conditioned on a representative individual l having a decreasing expected number of shocks to location i and conditional on the size of population in location i being bounded above by the initial population, the number of emigrants should, in expectation, monotonically decrease. Conversely, if the expected number of shocks is increasing over time and the population at any time $t > 0$ is lower bounded by its initial condition, then there will be an increase in the expected number of emigrants. Note that this is concordant with expectations of aggregate behavior as it is rational to expect that locations where the climate related amenity value increases, all else fixed, should experience a decrease in net emigrants. Likewise locations where the climate related amenity value decreases, all else fixed, should experience an increase in net emigrants.

3 Data

To empirically test the effect of climate-related shocks on migration and wages, I compile my main environmental data from the National Oceanic and Atmospheric Administration's (NOAA 2019) National Center for Environmental Information and migration data from the IRS's US Population Migration Dataset. Additional data is sourced from the IRS County, Metropolitan and Micropolitan SOI Tax Dataset, the USDA Rural-Urban Continuum Codes and the Bureau of Economic Analysis' Economic Profiles by County. I discuss these sources here and detail the construction of the compiled data products used in my analysis.

3.1 US Population Migration Data

I filter the IRS US Population Migration dataset to include only the years in 2011-2021. Outflow data is dyadic and is recorded at the county level with source and destination counties by year. The data records both the number of households and the approximate number of individuals which migrate from county i to county j. An emigrant household is recorded as residing in county i in year t and in county j in year $t+1$ for $t \in \{2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020\}$. The migration dataset which I utilize is constructed from personally identifiable information submitted to the IRS via a Tax Form 1040. Form 1040 submissions are aggregated by the IRS, using the filing address of an individual as an identifier, to generate non-personally identifiable information on household migration between counties in adjacent year pairs. The IRS Population Migration Data does not record instances of origin-destination migration where the number of households emigrating from a given origin county, i, to a destination county, j, were less than 20 due to privacy concerns. The approximate number of individuals is constructed aggregated Form 1040 submissions by considering the number of dependents per migrant household. The data further contains state level aggregates and county-to-foreign migration measures, at both the household and approximate individual level. As I focus on domestic inter-county emigration, I exclude all non-county-to-county migration flows from the data. I likewise exclude all non-migrants from the dataset by eliminating flows where the origin and destination counties are the same. This reduces my sample to consist of 604,073 dyadic county-county emigration observations between 2011 and 2021. The dyadic data only contains information on county-county pairs where individuals migrate from the former to the latter and does not record cases of 0 county-to-county migration flows.

3.2 NOAA Storm Events Data

The NOAA Storm Events Database denotes unusual weather phenomenon in a given geographic location. Events are deemed unusual, and thus included, if the event is defined as (1) having sufficient intensity to cause significant damage or other major harms; (2) being rare in a given location relative to the expected type of event; or (3) representing a local extrema in the type of event (NOAA 2019). Typical or expected events are therefore omitted. As such, the NOAA Storm Events Database represents an underestimate of the total number of events that occurred in a given geographical location in a given year. The individual weather and climatic events (i.e. Hail, Wildfires, Tropical Storms, Rain, etc.) are recorded at county, county-equivalently, zone and marine levels. To match the county-year level unit of analysis used throughout this paper, I consider only county and county-equivalent data points and discard any marine or zone level events. Between 2011 and 2021 there were 515,200 observed events across 3,245 unique counties and county equivalent levels in the US. Each event satisfies at least 1 of the 3 aforementioned necessary conditions to be recorded in the dataset. Additional descriptors by event in the NOAA Storm Events Database include (1) the monetary magnitude of crop and property damage, (2) start time, (3) end time, (4) start date, (5) end date, (6) magnitude (for selected events), (7) reporting agency, and (8) a short description of the event. To aggregate the event to the year level, I discard all minute, hour, day and month information from the start and end times. I define an event to have occurred in a given year on the basis of the start time. The end time of the event is not considered. I construct event counts per county-year with regards to both the total number of climate-related events and the number of events grouped by type to serve as county-level measures of exposure to a climate shock. Data is aggregated to the county-year level.

3.3 Tax Data and Economic Covariates

Additional economic covariates, which may influence the migration decision of households are sourced from IRS County, Metropolitan and Micropolitan SOI tax data. As was the case with the NOAA storm data, this is filtered down to the years 2011-2021. The raw data is decomposed by county-year and adjusted gross income level. I aggregate it up to the county-year level to match the coarseness of the IRS US Population Migration Data. The dataset is filtered down to only keep variables describing (1) the number of single returns, (2) number of joint returns, (3) the total adjusted gross income in the county and in a given year, (4) the number of returns with a positive income, (5) the total income by county and year, (6) the number of returns with business income, (7) the amount of net business income in a county and year, (8) the number of farm returns, (9) the number of returns with unemployment compensation and (10) the amount of unemployment compensation in the county in a given year.

Rural-Urban Continuum Codes are sourced from the USDA Economic Research Service. These codes classify each county on a discrete scale, from 1 to 9, by the degree by which they are considered metropolitan. A county with a continuum code of 1 is classified as the most metropolitan while a county with a continuum code of 9 is the most non-metropolitan county. This classification is done every 10 years thus prompting me to use the 2013 iteration of the classification as it lies within the sample time frame of 2011-2021. For those counties which did not exist prior to 2013, and thus were not included in the 2013 Rural-Urban Continuum Codes, I use the updated 2023 version. This allows me to classify all 3,245 unique counties on a scale of metropolitan to non-metropolitan.

The Bureau of Economic Analysis' Economic Profile by County provides wage information by county-year. The data exists for all years between 1969 and 2022 and provides information on (1) total personal income, (2) net earnings by place of residence, (3) personal current transfer receipts, (4) per capita personal income, (5) earnings by place of work, (6) proprietor's income and (7) total employment. As the aim is to understand the effect of climate-related shocks on wages in destination counties, through a migration pathway, I keep only data on per capita personal income. This data is further filtered to only consider the years between 2011 and 2021 so as to match the IRS Population Migration Dataset.

3.4 Destination Ambivalence and the Dyadic Settings

I consider specifications involving both a dyadic and a destination ambivalent form of the effect of climate-shocks on emigration and wages. Using the aforementioned data sources, the dyadic data is constructed by joining the 2011-2021 filtered NOAA Storm Events and Tax and Economic Covariates datasets to the 2011-2021 filtered IRS Population Migration dataset detailing domestic emigration. This is done by joint origin-year and joint destination-year. I further construct origin-destination, joint origin-year and joint destination-year indicators to be used in the modified gravity style estimation model detailed in Section 4.2. The result is a working data set which contains 604,703 variables, across 105 climate-shock related and economic variables by origin-joint destination-year and with covariates pertaining to both the joint origin-year and joint destination-year counties.

Destination ambivalent data is constructed by the aggregating dyadic data to the joint origin-year level. Let $n_{i,j,t}$ denote the number of households who migrate such that the household is located in county i at time t and in county j at time $t+1$. Then define $\tilde{n}_{i,t}$

$$\tilde{n}_{i,t} := \sum_{j} n_{i,j,t}$$

to denote the total emigration from county i at time t. Joint origin-year economic and climate-related covariates are consistent across a given joint origin-year county and thus are collapsed to a single observation. The result is working data set, termed destination ambivalent, of 28,736 variables across 40 variables. Aggregation to the origin-year level drops all variation introduced by destination-year and origin-destination variables from the working data set however preserves any origin-year variation.

3.5 Descriptive Statistics

Table 9 provides a visualization of the number of events by type. As noted in Section 3.1, this represents an under counting of the total number of events which occurred between 2011 and 2021 as, for an event to have been included in the NOAA Storm Events Database it must have met at least 1 of the 3 aforementioned conditions. From Table 9, Thunderstorm Wind, Hail and Flash Flood events are most frequently recorded "unusual" events between 2011 and 2021 while Drought, Dense Fog, Excessive Heat and Wildfires only have 1 observation each between 2011 and 2021. Given the construction of Proposition 1, I note that the inclusion of only "unusual" events, as termed by the NOAA Storm Events Database, should not significantly bias the empirically estimated results away from theory. This follows as if only expected events of type m occurred, the individual's updated distribution on λ_m would be similar to their prior. Table 10 provides a summary of the number of origin counties with (1) < 20 events, (2) $20 - 70$ events, (3) $70 - 120$ events and (4) > 120 events for each year between 2011 and 2020. 2021 is not included in the table since the dyadic structure of the data implies that data on 2021 is only included for destination counties. Note that the bulk of counties experience < 20 events in a given year with only a marginal subset of origin counties experiencing > 120 events in any given year.

Table 11 provides summary statistics on the number of emigrants, from all counties, per year. Note that the the 25^{th} percentile, 50^{th} percentile, 75^{th} percentile, mean, standard deviation and total number of emigrating households all follow the same trend with a local minima in the number of emigrants in 2014 and a local maxima in 2016. The 25^{th} percentile and median number of households emigrating domestically decreased in 2020, relative to 2011 levels, while the 75^{th} percentile, mean, standard deviation and total number of households all increased in 2020, relative to 2011 levels. Figure 1 disaggregates emigration by county Rural-Urban Continuum Code and plots the mean emigration by year. Note that while the mean number of emigrants differs by county classification, under the Rural-Urban Continuum Codes, the general trend in emigration reflects the national level averages seen in Table 11. Urban counties appear to track this trend more closely, relative to their rural counter parts, however this can be attributed to number of emigrants in both cases, with the greater number of urban emigrants dominating the trend observed in Table 11. Figure 2 provides a visualization of the distribution of county level wages per capita for even years between 2011 and 2021. Figure 2 demonstrates that for the selected even years between 2011 and 2021, the distribution of wages has remained relatively constant with year on year increases in the mean of the distribution.

4 Model Specification

Given the ability of the data to be aggregated into both a dyadic and a destination ambivalent framework, I estimate the effects of origin climate-related shocks in both settings. In both cases, I note the origin specific climate-related shocks have the capacity to act as push factors, with increases in the number of events leading individuals to emigrate from a given origin county under the framework of Proposition 2. Estimation in the dyadic setting adds a consideration of destination county-year counts of climate-related events which may act as a pull factor when the statement of Equation 3 is met. In contrast estimation in the destination ambivalent framework, while conceptually simpler, does not account for such pull factors. If destination county-year counts of climate-related shocks are not a significant factor in the consideration of a representative migrant, it would then follow that the dyadic estimation should collapse to the destination ambivalent formulation. I begin with a discussion of the destination ambivalent model specifications.

4.1 Destination Ambivalent Specification

Aggregation of the number of emigrant households, $\tilde{n}_{i,t}$, at the level of the origin county, i, allows for a OLS-style estimation of the effect of climate shocks on emigration. Given the differences in how I treat the climate-related shock, these models can broadly be described as "Pooled," "Binned" and "Type Dis-aggregated" and are defined as follows.

A. "Pooled" Destination Ambivalent Model

Equation 6 specifies the pooled model in the destination ambivalent setting.

$$\tilde{n}_{i,t} = \alpha \cdot E_{i,t} + \overrightarrow{\kappa} \cdot \mathbf{X}_{i,t} + \eta_{i,t} + \epsilon_{i,t} \tag{6}$$

$\tilde{n}_{i,t}$ is defined as in Section 3.4 and denotes the total number of emigrants from county i at time t. $E_{i,t}$ denotes the total number of events, of all types, in county i at time t. $\mathbf{X}_{i,t}$ is a vector of origin, i, specific economic covariates at time t. $\eta_{i,t}$ denotes the set of joint origin-year specific fixed effects and $\epsilon_{i,t}$ denotes the error term.

B. "Binned" Destination Ambivalent Model

Equation 7 specifies the binned model in the destination ambivalent setting. Events are binned by decile to average out noise arising from the relatively few number of counties with high counts of climate-related shocks which may bias estimation.

$$\tilde{n}_{i,t} = \overrightarrow{\alpha} \cdot \overrightarrow{F}_{i,t} + \overrightarrow{\kappa} \cdot \mathbf{X}_{i,t} + \eta_{i,t} + \epsilon_{i,t} \tag{7}$$

$\overrightarrow{F}_{i,t}$ denotes a vector of indicators for the decile of events in the origin county, i, at time t. $\tilde{n}_{i,t}$, $E_{i,t}$, $\mathbf{X}_{i,t}$, $\eta_{i,t}$ and $\epsilon_{i,t}$ are defined as in Equation 6.

C. "Type Dis-aggregated" Destination Ambivalent Model

As was the case in the dyadic setting, I note that the estimation of a type dis-aggregated model is motivated by heterogeneities in the effective "treatment" generated by disparate climate-related shock types. Equation 8 specifies the event type dis-aggregated model in the destination ambivalent setting.

$$\tilde{n}_{i,t} = \overrightarrow{\alpha} \cdot \mathbf{Type}_{i,t} + \overrightarrow{\kappa} \cdot \mathbf{X}_{it} + \eta_{i,t} + \epsilon_{i,t} \tag{8}$$

As was the case in the dyadic setting, $\mathbf{Type}_{i,t}$ is a vector such that $\mathbf{Type}_{it} = \left(T^{(1)}_{i,t}, ..., T^{(m)}_{i,t}, ..., T^{(p)}_{i,t}\right)^{\mathrm{T}}$. $T^{(m)}_{i,t}$ is defined to be the type of climate-related shock (i.e. tropical storm, hail, tornado, etc.) and each $1 \leq m \leq p$ denotes a unique type of climate-related shock. $\tilde{n}_{i,t},\ E_{i,t},\ \mathbf{X}_{i,t},\ \eta_{i,t}$ and $\epsilon_{i,t}$ are defined as in Equation 6. This model is estimated in Section 7 towards a consideration of heterogeneous treatment effects in the destination ambivalent setting.

4.2 Dyadic Specification

A. "Pooled" Dyadic Model

Let i denote the origin county and j denote the destination county. The dyadic "gravity-style" migration regression is given by Equation 9.

$$n_{i,j,t} = \alpha \cdot E_{i,t} + \beta \cdot E_{j,t} + \theta \cdot \ln(d_{i,j,t}) + \overrightarrow{\kappa} \cdot \mathbf{X}_{it} + \overrightarrow{\gamma} \cdot \mathbf{Y}_{j,t} + \eta_{i,t} + \delta_{j,t} + \rho_{i,j} + \epsilon_{i,j,t} \tag{9}$$

$n_{i,j,t}$ denotes the number of emigrant households, as measured by the IRS Population Migration data, that were in county i at time t and moved to county j at time $t+1$. As the data is filtered to only consider emigrants, $n_{i,j,t} > 0$ and likewise, $i \neq j$. $E_{i,t}$ and $E_{j,t}$ denote the number of events at time t in the origin, i and destination j respectively. $d_{i,j,t}$ denotes the distance between county i and county j at time t and is taken from the NBER county-distance database. $\mathbf{X}_{i,t}$ is a vector of origin county covariates at time t that describe the mean economic conditions in origin county i which may drive out-migration through push factors. Likewise $\mathbf{Y}_{j,t}$ is a vector of destination county covariates at time t which describe the mean economic conditions in the destination j and may affect emigration from the origin, i, through pull factors. $\eta_{i,t}$ denotes the set of origin county-year fixed effects and $\delta_{j,t}$ denotes the set of destination county-year fixed effects. $\rho_{i,j}$ denotes origin-destination fixed effects. $\epsilon_{i,j,t}$ denotes the error term.

B. "Binned" Dyadic Model

To understand cumulative effects across deciles of events, I re-estimate a binned version of Equation 9. This allows me to average out noise arising from the relatively few number of observations of counties with high numbers of climate-related shocks which may bias the estimation. Equation 10 specifies the "binned" model.

$$n_{i,j,t} = \overrightarrow{\alpha} \cdot \overrightarrow{F}_{i,t} + \overrightarrow{\beta} \cdot \overrightarrow{F}_{j,t} + \theta \cdot \ln(d_{i,j,t}) + \overrightarrow{\kappa} \cdot \mathbf{X}_{it} + \overrightarrow{\gamma} \cdot \mathbf{Y}_{j,t} + \eta_{i,t} + \delta_{j,t} + \rho_{i,j} + \epsilon_{i,j,t} \tag{10}$$

$\overrightarrow{F}_{i,t}$ denotes a vector of indicators for the decile of events in the origin county, i, at time t and $\overrightarrow{F}_{j,t}$ denotes a vector of indicators for the decile of events in the destination county, j, at time t. $d_{i,j,t},\ \mathbf{X}_{i,t},\ \mathbf{Y}_{j,t},\ \eta_{i,t},\ \delta_{j,t},\ \rho_{i,j}$ and $\epsilon_{i,j,t}$ are defined as in Equation 9.

C. Second Stage Model and Reduced Form

Given the dyadic structure, the conceptual effect of migration on wages is then self-evident. For counties experiencing net emigration the labor pool decreases in size so as to shift the labor supply curve inwards. This generates a supply side effect on the market which theoretically raises wages assuming constant labor demand. Likewise counties experiencing net immigration would see their labor pool increase such that the labor supply curve shifts outwards, thus generating an theoretical decrease in wages. In particular, the short-run assumption of constant labor demand is rational given high fixed costs associated with the relocation of firm production from an origin county i to destination county j. Therefore, it is reasonable to suppose that the relative stickiness of firm demand for labor relative to labor mobility between US counties generates an approximately constant demand for labor in counties across the short-run. As I consider year on year changes, the time horizon are sufficiently short enough to suppose that labor has enough time to update beliefs about the place based utility and emigrate while firms do not due to their relatively higher fixed costs. Moreover, a consideration of year on year changes, as opposed to a shorter time horizons, allows for the "averaging-out" of fluctuations in wages due to damages from climate-related shocks. As such, at the time horizon of my data, both hyper-short run fluctuations and long run trends should be minimized providing the best possible environment for an estimation of the causal effect of climate-related shocks on destination county wages. Insofar as climate shocks theoretically generate county-to-county migration, as suggested by Proposition 2, this motivates an estimation in the dyadic setting of the effect of climate-related shocks on destination wages via a migration instrument.

Equation 9 is implemented as the first stage model. Equation 11 then describes the second stage model which seeks to estimate the effect of climate-related shocks on wages to the destination county.

$$w_{j,t+1} = \alpha \cdot \hat{n}_{i,j,t} + \overrightarrow{\gamma} \cdot \mathbf{Y}_{j,t+1} + \delta_{j,t+1} + \nu_{i,j,t+1} \tag{11}$$

$w_{j,t+1}$ denotes the wage in destination county j at time $t+1$, which is after emigration has occurred. $\hat{n}_{i,j,t}$ denotes the fitted number of households migrating from origin county i to destination county j at time t and is estimated as described in Equation 9. $\mathbf{Y}_{j,t+1}$ denotes a vector of destination specific economic covariates at time $t+1$ which is after emigration. $\delta_{j,t+1}$ is a set of destination j fixed effects at time $t+1$. $\nu_{i,j,t+1}$ denotes the error term. In addition, I consider the reduced form expression given by Equation 12.

$$\begin{aligned} w_{j,t+1} = \alpha \cdot E_{i,t} + \beta \cdot E_{j,t} + \theta \cdot \ln(d_{i,j,t}) + \overrightarrow{\kappa} \cdot \mathbf{X}_{it} + \overrightarrow{\gamma}^{(1)} \cdot \mathbf{Y}_{j,t} \\ + \overrightarrow{\gamma}^{(2)} \cdot \mathbf{Y}_{j,t+1} + \eta_{i,t} + \delta^{(1)}_{j,t} + \delta^{(2)}_{j,t+1} + \rho_{i,j} + \epsilon_{i,j,t+1} \end{aligned} \tag{12}$$

Each of the terms in Equation 12 are defined as in a corresponding manner to Equations 9 and 11.

5 Destination Ambivalent Results

I note that amongst the 3,245 counties in the destination ambivalent dataset, 34 never experience a climate-related shock, as defined by NOAA. To determine whether migration outcomes follow county-specific trends, I test for unit specific trends on these never-treated counties. Table 12 provides a summary of the results of these tests. Model (1) performs a simple bivariate regression of the $County \times Year$ interaction on the number of emigrant households. Model (2) adds in county fixed effects. Model (3) adds in a set of origin county economic covariates to the statement of Model (2) and is the main specification of the unit specific trends test. Model (4) adds in fixed effects for Rural-Urban Continuum Codes. I note that while the inclusion of economic covariates and county fixed effects lead coefficients on most $county \times Year$ terms to not be statistically different from 0, there still exist 4 counties which display statistically significant unit specific trends at a 95% confidence level. This is robust to the inclusion of Rural-Urban Continuum Code Fixed effects as in Model (4) of Table 12. The relatively small sample size of 220 observations across 34 counties suggests that the lack of statistical significance amongst on county-year interaction coefficients may be due to the regressions being under-powered. I thus interpret the results, while partially indicative of unit specific trends, as being inconclusive on a whole. In subsequent models, I thus omit the inclusion of any unit-specific trends however include joint origin-year fixed effects, as specified in Equations 6, 7 and 8, to account for any time-variant trends to ensure that subsequent models are unbiased. I turn now to a discussion of the empirical results given by the Pooled and Binned Destination Ambivalent models from Section 4.1. I return to the Type Disaggregated Model in Section 7's consideration of heterogeneity in treatment.

5.1 Pooled Destination Ambivalent Models

Table 1 provides a summary of the results in the case of the pooled model specified by Equation 6 in Section 4.1. Model (1) in Table 1 is a simple bivariate model estimating the relationship between the number of climate-related shocks that a given county experiences and the total number of emigrant households. Model (2) adds in origin state-year fixed effects to the specification of Model (1) to account for time-varying heterogeneities which arise at the state level. County-year fixed effects, while the ideal solution given the ambiguous nature of the unit specific trends problem highlighted earlier, are collinear with the number of climate-related shocks that a county experiences in a given year and thus are not included in any of the specifications in Table 1. Model (3) adds Rural-Urban Continuum Code fixed effects to the specification of Model (2) to account for heterogeneities in the degree to which a county is metropolitan. The inclusion of such fixed effects are motivated by Figure 1 which demonstrates that while, in mean, counties appear to follow a similar trend with regards to emigration, urban counties dominate in terms of the volume of total migrants. Model (4) adds a set of economic covariates to the specification of Model (3). *County Adjusted Gross Income* is added to control for total wealth in a county, at the time of the shock, as higher levels of total wealth in a county may increase amenity values in a location via the increased provision of desirable public or club goods. *Origin Wages* control for per capita mean annual wage in the origin county, at the time of the shock, and are included to control for heterogeneities in initial average wage of an emigrant, by county, as increased wages may decrease the perceived costs of emigration by relaxing budget constraints in the period prior to said emigration. *Business Net Income* is included to control for employment related pull factors in the origin county which may discourage emigration. The *Amount of Taxable Pensions* is included as a proxy for the extent to which a county is composed of retirees in a given year as pensions denote fixed incomes which may

Dependent Variable:	Number of Emigrant Households			
Model:	(1)	(2)	(3)	(4)
Variables				
(Intercept)	492.5***			
	(70.93)			
Number of Events	134.0***	153.4***	121.8***	113.4***
	(6.424)	(7.318)	(6.727)	(7.431)
County Adjusted Gross Income†				-71.29
				(240.9)
Origin Wages				0.1130***
				(0.0118)
Business Net Income†				-113.2**
				(44.97)
Amount of Taxable Pensions†				-0.6134*
				(0.3447)
Number of Farm Returns†				-0.0079
				(0.0078)
Fixed-effects				
State-year		Yes	Yes	Yes
Rural-Urban Continuum Code			Yes	Yes
Fit statistics				
Observations	28,736	28,736	28,736	25,438
R^2	0.06537	0.18701	0.31338	0.33733
Within R^2		0.07898	0.05626	0.08471

Clustered (State-year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 1: Destination ambivalent estimation of the effect of climate-related shocks in the origin county on emigration. Number of emigrant households are measured in unit steps. Note that (†) denotes a variable whose value is rooted to its 2011 value such that it is then expressed as a real-valued multiple of the 2011 value. Standard errors are clustered at the joint origin-year level.

inhibit emigration due to cost of living considerations. Finally, Model (4) includes the *Number of Farm Returns* to control for the existence of farming communities in a county as potential migrants who own farms, and thus file farm returns, face greater fixed costs when emigrating which may then act as a pull factor keeping them in the same county across years. Models (1) - (4) in Table 1 utilize joint origin-year clustered standard errors.

Across all four models in Table 1, I note that the coefficient on the *Number of Events* is positive and statistically significant at a 95% confidence level. Moreover, the magnitude of the coefficient is shown to be robust to the inclusion of joint origin-year fixed effects as well as a set of economic covariates at the joint origin-year level. The results of Model (4) indicate an increase of 113.4 emigrant households for every unit increase in the number of climate-related shocks in a given county. As the climate-related events considered in Models (1)-(4) are denoted as "unusual" storm events, the occurrence these events would lead a Bayesian potential migrant to update their belief about the rates λ_m with which climate related shocks occur thus increasing $\mathcal{N}_{i,t}$. This generates a decrease in $A^c_{i,t+1}$ and under the prediction of Proposition 2, should generate emigration. The results of Model (4) support this hypothesis insofar as emigration is increasing the number of events experienced in the origin county in a given year. Under the specification of Model (4), I note that only *Origin Wages* and *Business Net Income* are statistically significant at a 95% confidence level with increased origin wages yielding a net positive effect and business net income yielding a net negative effect on emigration from the origin county.

5.2 Binned Destination Ambivalent Models

Table 2 provides a summary of the results in the case of binned events model specified by Equation 7 in Section 4.1. Model (5) estimates the relationship between the binned number of climate-related shocks and the number of emigrant households by county with state-year and Rural-Urban Continuum Code fixed effects. As was the case in Table 1, Models (1) - (4), I note that county-year fixed effects are collinear with the binned number of events and thus render it impossible to estimate the effect of said climate-related shocks on emigration if included. I therefore again turn to the second-best solution of applying state-year fixed effects to control for time-variant heterogeneity at the state level which would otherwise bias my estimate of the the number of households emigrating from a county in a given year. Rural-Urban Continuum Code fixed effects are again added to account for heterogeneities in the degree to which a county is metropolitan given the trends shown in Figure 1. Model (6) adds a series of joint origin-year economic covariates to the specification of Model (5). These joint origin-year covariates are identical to those added in the case of Model (4) in Table 1 and likewise are similarly motivated in their addition. Standard errors are clustered by county-year. Note that across Table 2 the binned events appear to generate a positive and statistically significant effect, at a 95% confidence level on the number of emigrant households. The exception to this trend is the coefficient on the bin of 130-140 events. Examining the data, I note that there exist only 28 observations where a county in a given year experience between 130-140 climate-related shocks. Contrasting this to the case where a county experienced < 10, wherein there were 15,962 observations, I conclude that the lack of statistical significance on this coefficient may be due to a lack of statistical power at higher numbers of events per county-year observation. Despite this, I note that, from Model (6), a county which experiences < 10 events expects 832.9 households to emigrate over the course of a year while a county with $50 - 60$ such climate-related events expects $6,685.5$ households to emigrate over the same time period. From Table 2, I thus conclude that the effect of the binned number of climate-related shocks is weakly increasing as the number of binned shocks increases. In aggregate, I thus conclude that models (5) and (6) in Table 2 corroborate the narrative presented in Table 1. More specifically, Models (5) and (6) similarly suggest that increases in the number of binned climate-related shocks generates an increase in the number of emigrants from a county in a given year and that this relationship is robust to the inclusion of state-year and Rural-Urban Continuum Code fixed effects and joint origin-year economic covariates. Of the covariates, I note the only *Origin Wages, Business Net Income* and the *Amount of Taxable Pensions* are statistically significant at a 95% confidence level.

6 Dyadic Results

I now turn to a discussion of the results in the dyadic case. Throughout the models discussed in this section, I apply joint origin-year fixed effects to control for heterogeneities arising from time-variant trends in the origin county and similarly apply joint destination-year fixed effects to control for heterogeneities arising from time-variant trends in the destination county. Origin-destination pair fixed effects are also frequently implemented in the models in this section to control for heterogeneities that arise between pairs of origin and destination counties. In addition to an examination of the models specified by Equations 9 and 10, I here also exploit the dyadic structure of the data to consider the effect of climate-related shocks on destination county wages through an migration instrument, as specified in Equations 11 and 12.

6.1 Pooled Dyadic Models

Table 3 provides a summary of the results in the case of the Pooled Dyadic model specified by Equation 9. Model (7) applies the framework of a trade gravity-style model to the case of emigration and regresses the number of origin climate-related shocks, the number of destination climate-related shocks and the distance between the origin and destination counties on the number of emigrant households. Moreover Model (7) applies destination state-year and origin state-year fixed effects to control for time-variant heterogeneities in the origin and destination counties. Similarly, origin Rural-Urban Continuum Code and destination Rural-Urban Continuum Code fixed effects are included to control for heterogeneities in the degree to which the origin and destination county is metropolitan, given the trends demonstrated in Figure 1. Model (8) adds origin and destination economic covariates to Model (7). These take the form the form of covariates describing *Origin County Adjusted Gross Income, Destination County Adjusted Gross Income, Origin Wages, Destination Wages, Origin Number of Farm Returns* and *Destination Number of Farm Returns*. Model (9) adds origin and destination economic covariates, in the form of *Origin Taxable Pension Amount* and *Destination Taxable Pension Amount*, to the specification of Model (8). The motivations between the inclusion of the adjusted gross income, wages, number of farm returns and taxable pension amount remain unchanged, in the dyadic setting, from the discussion in Section 5.1 and thus is not repeated here. Model (10) adds covariates for *Origin Unemployment Compensation* and *Destination Unemployment Compensation* to Model (9). The inclusion of origin and destination controls for unemployment are used to proxy for unemployment rates by county and are motivated the fact that, like a pension, unemployment can introduce liquidity constraints which inhibits emigration due to cost of living considerations and a fixed budget constraint. Model (11) adds origin-destination pair fixed effects to the estimation of Model (10). Model (12) re-estimates Model (11) using a Poisson Pseudo-Maximum Likelihood (PPML) estimator. The use of a PPML estimator in this setting is motivated by the discrete count structure of the number of emigrant households as the outcome variable wherein $n_{i,j,t} > 0$ as defined in Equation 9. Moreover, the missing-ness of migration data in the case where the number of emigrant households from a given origin to a given destination county is less than 20 is akin to the missing data problem faced by the trade literature and which has been partially resolved by the use of such a PPML estimator.

I note that across Models (7) - (10), in Table 3, the effect of both the number of origin climate-related shocks and number of destination-climate related shocks is positive and statistically significant at the 95% confidence level. Moreover, I note that the coefficients on both origin and destination counts of climate-related shocks is robust to the inclusion of economic covariates, in both the origin and destination counties, and to the use of the aforementioned set of fixed effects. Models (7) - (10) notably lend further evidence in support of Proposition 2 as the number of emigrant households is shown to be increasing in the number of origin events in Table 3 with 0.9401 households emigrating for every unit increase in climate-related shocks in the origin county in Model (10). However, the effect on the number of emigrant households also appears to be increasing in the number of destination events as 0.8848 additional households emigrate from the origin to the destination for every unit increase in the number of destination climate-related events. In this way, the number of events in the destination appears to act counter to a hypothetical pull factor. Table 3 does not provide any information on the relative number of events in the origin and destination and thus it is impossible to determine whether the positive coefficient on the number of destination climate-related shocks is due to relative differences between the origin and destination numbers or not. Of the origin and destination economic covariates included in Model (10), I note that wages, and unemployment compensation, in both the origin and destination, are shown to be statistically significant at the 95% confidence level. Specifically, the number of emigrant households is seen to be increasing in both origin and destination wages and decreasing in origin and destination unemployment compensation. Model (10) further demonstrates that the number of emigrant households is decreasing in the taxable pension amount in the destination and is increasing in destination adjusted gross income. Put together, these results suggest that beyond the effect of the climate-related shock, increases overall income facilitates emigration while fixed budget constraints impede it.

As Model (11) re-estimates Model (10) with the inclusion of origin-destination pair fixed effects, I note that the addition of origin-destination pair fixed effects renders it impossible to identify a statistically significant effect of the number of origin and destination climate-related shocks on the number of emigrant households. Given the similarities of Model (11) to the gravity model, I thus re-estimate Model (11) using a PPML estimator to determine whether the loss of significance is a consequence of the structure of OLS model or whether it reflects an actual sensitivity to controlling for heterogeneities between origin and destination county pairs. Using a PPML estimator, Model (12) demonstrates that the loss of statistical significance on the coefficients of the number of origin and destination climate-related shocks is most likely a consequence of the inherent structure of an OLS estimator relative to the PPML estimator. with Model (12), I show that the number of emigrant households is indeed increasing in the number of origin and destination climate-related shocks. Moreover, Table 3 demonstrates that this relationship is statistically significant at a 95% confidence interval and is robust to the inclusion of origin-destination pair fixed effects when a Poisson Pseudo-Maximum Likelihood Estimator is used. More concretely, I note that from Model (12) a unit increase in the number of origin climate related shocks generates a 0.03% increase in the number of emigrant households to the destination county. Given the coefficient on the number of destination county climate-related shocks, I similarly conclude that a unit increase in the number of destination climate-related shocks generates a 0.03% increase in the number of emigrant households to the destination county.

6.2 Binned Dyadic Models

Table 4 summarises the results in the case of the binned dyadic models given by Equation 10. Model (13) estimates the effect of replaces the a series of indicators for the binned deciles of climate related shocks in the origin county and a series of indicators for the binned deciles of climate related shocks in the destination county on the number of emigrant households from the origin county. Model (13) also includes a covariate to control for the distance between the origin and destination counties. Unlike the case of the binned destination ambivalent model in Section 5.2, I note that only the coefficients on binned origin indicators for > 150 events, $100 - 110$ events, $130 - 140$ events, $50 - 60$ events and $70 - 80$ events are statistically significant at a 95% confidence level. Moreover only the coefficient on a binned destination indicator for $140 - 150$ events is statistically significant at a 95% level. I further note that wherein the parity of coefficients was continuously positive in the prior specifications in Tables 1, 2 and 3, Model (13) illustrates significant variability in the parity of the coefficients on binned origin and destination decile indicators of climate-related shocks. Model (14) adds in the same set of origin and destination economic covariates for wages, adjusted gross income, number of farm returns, taxable pension amount and unemployment compensation to the specification of Model (13). I note that the statistically significant coefficients estimated in Model (13) are robust to the inclusion of such economic origin and destination county covariates. Despite this, only the coefficient on the indicator for $60 - 70$ binned events in the origin shifts from not being statistically significant in Model (13) to being statistically significant in Model (14). Moreover, of the covariates included, on origin wages, destination unemployment compensation and adjusted gross income and taxable pension amount, in both the origin and destination counties, are statistically significant at a 95% confidence interval. To test whether the lack of statistical significant throughout Models (13) and (14) is a consequence of the specification of an OLS style estimator, I use Model (15) to re-estimate the specification of Model (14) using a PPML estimator. Doing so increases the number of origin and destination coefficients that are found to have a statistically significant effect on the number of emigrant households. I note that changing the specification of the estimator does not alter the observed variation in parity thus suggesting that such observed variation in coefficient parity is an accurate descriptor of the effect of each decile in the Model as it is shown to be robust to fixed effects, origin and destination economic covariates and changes to the functional form of the estimator. I note that amongst the origin decile indicators in Model (14), a county which experiences > 150 climate-related shocks will then experience a 3.987% increase in the number of emigrant households while a county which experiences $130 - 140$ climate related-shocks will experience a -6.452% increase in the number of emigrant households.

6.3 Climate Shock Effects on Wages

Having estimated the empirical forms of Equations 6, 7, 9 and 10, I now turn to a consideration of the effect of climate-related shocks on destination county wages via a migration instrument. As discussed in Section 4.2, I note that the year on year time horizon of my data allows me to minimize both hyper-short run fluctuations arising, from the recovery process from a climate-related shock, and the long run trends, which generate multiple causal pathways for climate events to affect wages. As such the year on year time horizon of my data enables me to study the effect of an origin climate shock on destination wages given the theoretical justification presented in Section 4.2. Table 5 provides a summary of the 2 stage least squares regression as well as the reduced form model.

Model (16), in Table 5 denotes the reduced form expression given by Equation 12. In particular, Model (16) estimates the effect of climate-related shocks, in the origin and destination county, on the number of emigrant households given controls for origin and descrination economic covariates; joint origin-year and joint destination-year fixed effects and origin-destination fixed effects. Note that Model (16) indicates a -0.01% decrease in mean annual destination income per capita per unit increase in the number of origin climate-related shocks and a 0.03% increase in mean annual income per capita per unit increase in the number of destination climate-related shocks. Both results are statistically significant at a 95% confidence level. As such, the reduced form expression of Model (16) would suggest that an increase in the number of climate related shocks to an origin county does indeed have a negative effect on wages to the destination county. This lends support to the notion put forwards in Section 4.2 that the re-allocation of labor, due to migration induced from climate-related shocks, should have a negative affect on wages in destination counties are labor flows in and thus expands the potential labor pool relative to an approximately constant demand. Note that the $t+1$ destination county economic covariate given by the Adjusted Gross Income is dropped due to collinearity. Consequently mean annual income in the destination county is recorded at time $t+1$ while all remaining covariates in Model (16) are recorded at time t. At a 95% confidence level, I note that Model (16) also demonstrates that the mean annual wage in the destination county is increasing in the distance between the origin and destination counties, the adjusted gross income in the destination county in the prior year, the number of destination county farm returns and in destination county unemployment compensation. Conversely, Model (16) indicates that at a 95% confidence level, mean annual income in the destination county is decreasing in taxable pension amounts in the origin and destination counties.

In Table 5, Model (17) utilizes the 2 stage least squares approach specified in Equations 9 and 11 to estimate the effect of climate-related shocks on mean annual income in the destination county, via an emigration instrument. Note that all covariates and fixed effects applied in Model (16) are applied in Model (17). Model (18) re-estimates Model (17) with the addition of origin-destination pair fixed effects. As the coefficient on the *Fitted Number of Emigrant Households* is statistically significant at a 95% confidence level in Model (17) but fails to be statistically different from 0, at the same level in Model (18), I note that the fitted effect of climate-related shocks on mean annual wages in the destination county is not robust to the inclusion of origin-destination pair fixed effects. As noted in the comparison of Models (11) and (12) in Table 3, this may be due to the method of OLS estimation as Model (12) indicates that when the number of emigrant households is estimated using a PPML estimator, the results are robust to the inclusion of origin-destination pair fixed effects. Given the results of Table 5, I thus conclude that while there may exist a statistically significant and robust second stage relationship between the fitted number of emigrant households and mean annual wages in the destination county, I do not find strong evidence in favor of such a relationship here.

7 Robustness

Thus far I treat disparate types of climate shocks as having the same effect on the potential migrant. Given the differential frequency and perceived potential harms of disparate shocks, this may fail to hold and thus may bias results. I thus test for heterogeneous treatment effects, in the Destination Ambivalent setting, using Equation 8. Moreover, given the high number of simultaneous hypothesis tests run in the type disaggregated and binned models, I apply a False Detection Rate correction.

7.1 Heterogeneity in Treatment

Table 6 provides a summary of the results in the case of the Type Disaggregated model specified by Equation 8. Model (19) re-estimates Model (3) in the setting where the "treatment" effect of disparate climate-related shocks cannot be considered to be identical. Model (20) adds in a series of origin county economic covariates to Model (19). Note that this practically is a re-estimation of Model (4) under the assumption that disparate event types cannot be treated identically due to heterogeneous treatment effects. Model (19) demonstrates that of the 16 types of climate-related shocks, for which a coefficient can be estimated, a statistically significant relationship, at the 95% confidence level, exists for only *Flash Flood, Flood, Hail, Thunderstorm Wind, Funnel Cloud, Heavy Rain, Dust Devil, Lightning, Fog* and *Strong Wind* events. Moreover, I note that disaggregation by type reveals heterogeneities in the parity of the effect on the number of emigrant households as, of the events with statistically significant coefficients, only *Heavy Rain* is shown to have a negative effect on the number of emigrant households. Model (20) indicates that the coefficient estimates for the effects of *Heavy Rain* and *Fog* on the number of emigrant households are not robust to the inclusion of origin county-year economic covariates. In particular, Model (20) indicates that of each of the 16 types of climate-related shocks, for which a coefficient can be estimated, Dust Devils appear to have the greatest effect on the number of emigrant households with a predicted 3,025.6 households emigrating for a unit increase in the number of Dust Devils. Putting the results of Models (19) and (20) together, I thus conclude that, at a minimum, *Flash Floods, Floods, Hail, Thunderstorm Wind, Funnel Clouds, Dust Devils* and *Lightning* have a positive and statistically significant effect on the number of emigrant households from a given origin county.

7.2 False Detection Rate Corrections

Considering the results of Tables 2 and Table 6, I note the large number of simultaneous hypothesis tests being performed. To correct for the expected number of false positives amongst parameter estimates in these models, I employ a false detection rate correction for Models (6) and (20). Table 7 provides the resultant p-values, by coefficient in Model (6), with and without the false detection rate correction. From Table 7, I note that the application of the false detection rate correction does not alter the statistical significance of any climate-shock related coefficients at a 95% confidence interval thus implying that even with a correction for potential type I error, the conclusions reached in Section 5.2 remain valid. Table 8 summarises the p-values, with and without the false detection rate correction, for the type disaggregated results given in Table 6 Model (20). As was the case with Table 7, I note that applying a correction for the false discovery rate did not alter the statistical significance, at the 95% level, of terms which were already significant at or above the 95% confidence level. Thus, I conclude that the conclusions drawn in Section 7.1, from Model (20) remain valid.

8 Conclusions

Dealing with the increased risk of climate shocks in the modern world is a foregone conclusion given anthropogenic climate change. In a destination ambivalent framework, I show that when such shocks are measured by the total of number "usual" events in a given county, the effect of such a shock is to increase the number emigrant households from the origin county by 113.4 per unit increase in the number of shocks. Using a decile binned setting, I further show that the effect on the number of emigrant households is weakly increasing in accordance with the number of shocks. Moreover, when shocks are disaggregated by type, I demonstrate that only Flash Floods, Floods, Hail, Thunderstorm Wind, Funnel Clouds, Dust Devils and Lightning have a statistically significant impact on the number of emigrant households. I note that these results may be a feature of the data and that the lack of significance on events such as observations of "unusual" Wildfires, Tornadoes and Tropical Storms, which generate greater damages, are limited.

In a dyadic setting, I show that that the effect on the number of emigrant households, of an increase in the number of origin climate-related shocks, is positive and statistically significant. In an OLS setting without origin-destination fixed effects, I estimate that for every addition climate related shock in the origin county, 0.9401 households emigrate from the origin to the destination. Surprisingly, I find a positive effect that arises when the number of shocks in the destination county increase. In the same OLS setting without origin-destination fixed effects, I estimate that every addition climate related shock in the destination county increases emigration from the origin to the destination by 0.8843 households. I show that given the similarity to gravity-style models a Poisson Pseudo-Maximum Likelihood estimator is well suited to estimate the effect on the number of emigrant households when origin-destination pair fixed effects are included. I estimate an increase 0.03%, with respect to the number of emigrant households, when for a unit increase in the number of events in either the origin county or the destination county. I find similar results in a binned setting however note that relative. I finally show that while there is a positive and statistically significant, post emigration, increase of 0.00807% in mean annual destination county wages per unit increase in origin county climate shock, this result is not robust to the inclusion of origin-destination fixed effects. In contrast I demonstrate that my migration regressions, in the destination ambivalent setting, are robust to the correction for heterogeneous treatments arising from disparate event types and are similarly robust to a false detection rate correction.

In aggregate, my results are broadly supportive of the major theoretical implications of Proposition 2 and indicate that climate shocks to the origin county, measured not by temperature fluctuations or changes in precipitation but rather by event counts, do indeed have a statistically significant and positive relationship with migration. In effect, such shocks appear to act as push factors, potentially altering the place based utility for the hypothetical migrant and leading them to relocate to a new county within the US. Future work might look at why destination county climate shocks appear to have a pull effect that is positive, as opposed to negative, such that emigration to the destination county is increasing in the number of shocks to the destination. Moreover, future work might look more closely at the robustness of 2 stage least squares relationship between climate shocks in the origin county and wages, post emigration, in the destination given that the results here were statistically significant under certain specifications. Future work might also look at the effects of international migration using event type data, as opposed to the more tradition temperature and precipitation fluctuation data, as the treatment. Finally, future work may approach the case where r_m and s_m are not fixed but rather have an underlying distribution of their own such that the aggregate population is constructed from a distribution of individuals with differential priors on the distribution of the rate of climate shocks of type m in location i. Such an approach would shift away from a representative migrant to instead consider how distributional changes in priors impact migration outcomes and is particularly relevant given the existing literature which focuses on individual learning given climate risks.

References

[1] Bureau of Economic Analysis. 2024. "Regional Economic Accounts: CAINC30: Economic Profile by County." Apps.bea.gov. 2024. https://apps.bea.gov/regional/downloadzip.cfm.

[2] Cattaneo, Cristina, and Giovanni Peri. 2015. "The Migration Response to Increasing Temperatures." NBER, October. https://doi.org/10.3386/w21622.

[3] Chen, Shuai, Paulina Oliva, and Peng Zhang. 2017. "The Effect of Air Pollution on Migration: Evidence from China." Social Science Research Network, November. https://doi.org/10.3386/w24036.

[4] Fan, Qin, Karen Fisher-Vanden, and H. Allen Klaiber. 2018. "Climate Change, Migration, and Regional Economic Impacts in the United States." Journal of the Association of Environmental and Resource Economists 5 (3): 643–71. https://doi.org/10.1086/697168.

[5] Internal Revenue Service. 2024a. "SOI Tax Stats - US Population Migration Data ." www.irs.gov. February 7, 2024. https://www.irs.gov/statistics/soi-tax-stats-migration-data.

[6] Internal Revenue Service. 2024b. "SOI Tax Stats - County, Metropolitan and Micropolitan Data." www.irs.gov. February 29, 2024. https://www.irs.gov/statistics/soi-tax-stats-county-data.

[7] Knutson, Thomas, Maya Chung, Gabriel Vecchi, Jingru Sun, Tsung-Lin Hsieh, and Adam Smith. 2021. "Climate Change Is Probably Increasing the Intensity of Tropical Cyclones." https://sciencebrief.org/uploads/reviews/ScienceBrief_Review_CYCLONES_Mar2021.pdf.

[8] Leah Platt Boustan, Price V Fishback, and Shawn Kantor. 2007. "The Effect of Internal Migration on Local Labor Markets: American Cities during the Great Depression." Library Union Catalog of Bavaria, Berlin and Brandenburg (B3Kat Repository), July. https://doi.org/10.3386/w13276.

[9] Mitterwallner, Veronika, Manuel Steinbauer, Gregor Mathes, and Anna Walentowitz. 2024. "Global Reduction of Snow Cover in Ski Areas under Climate Change." PLOS ONE 19 (3): e0299735–35. https://doi.org/10.1371/journal.pone.0299735.

[10] Mullins, Jamie T, and Prashant Bharadwaj. 2021. "Weather, Climate, and Migration in the United States." NBER, March. https://doi.org/10.3386/w28614.

[11] National Bureau of Economic Research. n.d. "County Distance Database." NBER. Accessed March 20, 2024. https://www.nber.org/research/data/county-distance-database.

[12] National Oceanic and Atmospheric Administration NCEI. 2019. "Storm Events Database | National Centers for Environmental Information." NOAA.gov. 2019. https://www.ncdc.noaa.gov/stormevents/.

[13] Turco, Marco, John T Abatzoglou, Sixto Herrera, Yizhou Zhuang, Sonia Jerez, Donald D Lucas, Amir AghaKouchak, and Ivana Cvijanovic. 2023. "Anthropogenic Climate Change Impacts

Exacerbate Summer Forest Fires in California." EARTH, ATMOSPHERIC, and PLANETARY SCIENCES 120 (25). https://doi.org/10.1073/pnas.2213815120.

[14] USDA Economic Research Service. 2024. "USDA Economic Research Service Rural-Urban Continuum Codes." www.ers.usda.gov. January 22, 2024. https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/.

[15] Westerling, A.L. 2016. Increasing western U.S. forest wildfire activity: Sensitivity to changes in the timing of spring. Phil. Trans. R. Soc. B. 371:20150178.

[16] Xi, Dazhi, Ning Lin, and Avantika Gori. 2023. "Increasing Sequential Tropical Cyclone Hazards along the US East and Gulf Coasts." Nature Climate Change, February, 1–8. https://doi.org/10.1038/s41558-023-01595-7.

A Proofs

A.1 Proof of Proposition 1

Given that

$$\lambda_m \sim Gamma(s_m, r_m) \Rightarrow f_0(\lambda_m) = \frac{r_m^{s_m}}{\Gamma(s_m)} \lambda_m^{s_m - 1} e^{-r_m \lambda_m}$$

for $\lambda_m > 0$ and with hyperparameters s_m, r_m that describe the shape of the distribution. Then over the interval $[t, t+1)$ with n disjoint subintervals and observations of $\{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\}$ shocks in each corresponding sub-interval, l's posterior distribution on λ_m at time $t+1$ is then given by

$$f_1(\lambda_m | \{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\}) \propto g(\{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\} | \lambda_m) f_0(\lambda_m)$$

where $g(\{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\} | \lambda_m)$ denotes the likelihood function such that $y_{m,k} | \lambda_m \sim Poisson(\lambda_m)\ \forall\ 1 \leq k \leq n$. Since on the interval $[t, t+1)$, l has not yet updated the distribution f_0, consider each $y_{m,k}$ to be an i.i.d draw such that

$$g(\{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\} | \lambda_m) = \prod_{k=1}^{n} \frac{e^{-\lambda_m} \lambda_m^{y_{m,k}}}{(y_{m,k}!)}$$

Then, note that

$$\Rightarrow f_1(\lambda_m | \{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\}) \propto \frac{e^{-n\lambda_m} \lambda_m^{\sum y_{m,k}}}{\prod_{k=1}^{n} (y_{m,k}!)} \cdot \frac{r_m^{s_m}}{\Gamma(s_m)} \lambda_m^{s_m - 1} e^{-r_m \lambda_m}$$

where the $y_{m,k}$'s are assumed to be independent. Considering only terms in r, s and noting that $\lambda_m \sim Gamma(s_m, r_m)$ is the conjugate prior to $Poisson(\lambda_m)$,

$$f_1(\lambda_m | \{y_{m,1}, y_{m,2}, \ldots, y_{m,n}\}) \propto e^{-(r_m + n)\lambda_m} \lambda_m^{-1 + s_m + \sum y_{m,k}}$$

Thus i's belief about the distribution of the frequency of shocks of type m at time $t+1$ is then given by

$$\lambda_m \sim Gamma\left(s_m + \sum_{k=1}^{n} y_{m,k}, r_m + n\right)$$

where n denotes the number of time steps between $[t, t+1)$.

Now take $N_m(t)$ to define a Poisson Process with intensity λ_m. During $[t, t+1)$, individual l has not updated their belief about the distribution of λ_m and thus believes that in expectation, for $t' \in [t, t+1)$ and for some fixed $\tau > t+1$,

$$\begin{aligned} \mathbb{E}_0((N_m(\tau) - N_m(t'))) &= \mathbb{E}_0(\mathbb{E}_0(N_m(\tau) - N_m(t') | \lambda_m = k)) \\ &= \int_0^\infty \mathbb{E}_0(N_m(\tau) - N_m(t') | \lambda_m = k) f_0(k) dk \end{aligned}$$

where $\mathbb{E}_0((N_m(\tau) - N_m(t')))$ defines the expectation during $[t, t+1)$. By definition of a Poisson process, $\mathbb{E}_0(N_m(\tau) - N_m(t)|\lambda_m = k) = k(\tau - t')$ and thus

$$\mathbb{E}_0((N_m(\tau) - N_m(t'))) = \int_0^\infty k(\tau - t')f_0(k)dk = (\tau - t')\mathbb{E}_0(\lambda_m)$$

$$\Rightarrow \mathbb{E}_0((N_m(\tau) - N_m(t'))) = (\tau - t') \cdot \frac{s_m}{r_m}$$

Similarly, for $t + 1 \leq t'' < \tau$, note that

$$\mathbb{E}_1((N_m(\tau) - N_m(t''))) = (\tau - t'') \cdot \frac{s_m + \sum_{k=1}^n y_{m,k}}{r_m + n}$$

Then for fixed τ

$$\frac{s_m + \sum_{k=1}^n y_{m,k}}{r_m + n} > \frac{s_m}{r_m} \Rightarrow \lim_{t' \to (t+1)^-} E_0((N_m(\tau) - N_m(t'))) < \lim_{t'' \to (t+1)^+} E_1((N_m(\tau) - N_m(t'')))$$

thus implying that the expectation changes discontinuously at $t+1$. Then, $E_1((N_m(\tau) - N_m(t+1)))$ yields the number of events of type m that individual l expects between $[t + 1, \tau]$ at location i.

Recognizing that m is arbitrary, note that this holds for all m and thus the individual l updates the distribution of each λ_m over the interval $[t, t + 1)$ such that for each m after updating,

$$\lambda_m \sim Gamma\left(s_m + \sum_{k=1}^n y_{m,k}, r_m + n\right)$$

holds. Superposition of Poisson Processes then implies that if $N_1(t), N_2(t), ..., N_m(t), ..., N_p(t)$ are Poisson Processes describing the number of climate shocks of types $1, 2, ..., m, ..., p$, then $N(t) = N_1(t) + ... + N_m(t) + ... + N_p(t)$ also denotes a Poisson Process describing the total number of shocks such that $\forall\, t' \in [t + 1, \tau]$. Then

$$\mathbb{E}_1(N(\tau) - N(t')) = \mathbb{E}_1((N_1(\tau) - N_1(t')) + ... + (N_p(\tau) - N_p(t')))$$

$$= \mathbb{E}_1(N_1(\tau) - N_1(t')) + ... + \mathbb{E}_1(N_p(\tau) - N_p(t'))$$

$$= \mathbb{E}_1(\mathbb{E}_1(N_1(\tau) - N_m(t')|\lambda_1 = k)) + ... + \mathbb{E}_1(\mathbb{E}_1(N_p(\tau) - N_p(t')|\lambda_p = k))$$

so the number of total climate shocks that individual l expects in location i is given by

$$\mathcal{N}_{i,t+1} := \mathbb{E}_1(N(\tau) - N(t + 1)) = (\tau - t') \sum_{m=1}^p \left(\frac{s_m + \sum_{k=1}^n y_{m,k}}{r_m + n}\right)$$

The above formulation of $\mathcal{N}_{i,t+1}$ then allows for changes in the expectation of individual climate shock types to alter the individual's expectation of the total number of shocks that will be experience in location i.

A.2 Proof of Proposition 2

Without a loss of generality, I demonstrate the claim for migration at time $t+1$. All other cases are similar since $k \in \mathbb{Z}^+$. At time $t+1$, individual l updates their belief about $\mathcal{N}_{i,t}$, following proposition 1, and thus updates $u_t(i)$ to be $u_{t+1}(i)$. For each $j \in L$, the decision for individual l to migrate is rational if, for a cost, $C(i,j)$, of migration from location i to location j at time t, the following inequality is satisfied

$$u_t(i) \leq u_t(j) - C(i,j)$$

Rearranging terms,

$$u_{t+1}(i) - u_t(i) + u_t(i) \leq u_{t+1}(j) - C(i,j)$$

$$\Rightarrow \Delta u_t(i) + u_t(i) \leq u_{t+1}(j) - C(i,j)$$

By substitution,

$$\Delta u_t(i) = A_{t+1}(i)c_{t+1}(i)z_{t+1}(i) - A_t(i)c_t(i)z_t(i)$$

By assumption of the proposition, $A^{nc}_{t+1} = A^{nc}_t(i)$, $c_{t+1} = c_t$ and $z_{t+1} = z_t$, or equivalently that non-climate amenity values, consumption and wages remain constant across short run time steps t and $t+1$,

$$\Delta u_t(i) = \left(A^c_t(i) \cdot h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) - A^c_t(i)\right) c_t z_t$$

$$\Rightarrow \left(A^c_t(i) \cdot h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) - A^c_t(i)\right) c_t z_t + u_t(i) \leq u_{t+1} - C(i,j)$$

For completeness, note that by construction of the function h,

$$\frac{\partial h}{\partial \mathcal{N}_{i,t}} < 0 \ \wedge \ h(\mathcal{N}_{i,t} = k, \mathcal{N}_{i,t-1} = k) = 1 \Rightarrow \begin{cases} h(\mathcal{N}_{i,t} < \mathcal{N}_{i,t-1}) > 1 \\ h(\mathcal{N}_{i,t} > \mathcal{N}_{i,t-1}) < 1 \end{cases}$$

and thus $\mathcal{N}_{i,t} < \mathcal{N}_{i,t-1} \Rightarrow A^c_t(i) < A^c_{t+1}(i)$ so that a decrease in the expected number of climate shocks over equal length time intervals leads to an increase in the amenity value of a given location, i. Now let $\mathcal{F}_{i,t}$ denote the set of locations to which it was rational for individual l to migrate to at time t such that

$$\mathcal{F}_{i,t} := \{j \in L : \Delta u_t(i) + u_t(i) + C(i,j) \leq u_{t+1}(j)\}$$

Similarly define $\mathcal{F}_{i,t+1}$ to denote the set of locations for which it is rational for individual l to migrate to at time $t+1$. Note that since, $A_t, c_t, z_t > 0$, the definition of the function $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1})$ implies that

$$(h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1 \Leftrightarrow \Delta u_t(i) > 0) \Rightarrow |\mathcal{F}_{i,t}| \geq |\mathcal{F}_{i,t+1}|$$

$$(h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1 \Leftrightarrow \Delta u_t(i) < 0) \Rightarrow |\mathcal{F}_{i,t}| \leq |\mathcal{F}_{i,t+1}|$$

Thus the number of locations $j \in L$, for which it is rational for individual l to migrate from i, is increasing between t and $t+1$ if $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1$ and is decreasing if $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1$.

A.3 Proof of Proposition 3

I begin by demonstrating the first claim. For the base case, note

$$h(\mathcal{N}_{i,1}, \mathcal{N}_{i,0}) > 1 \Rightarrow \mathbb{P}_i(\mathcal{F}_{i,1}) < \mathbb{P}_i(\mathcal{F}_{i,0})$$

Since $M_{i,0}$ is exogenously given and $M_{i,0} \geq M_{i,1}$ by assumption, it immediately follows that $M_{i,0}\mathbb{P}_i(\mathcal{F}_{i,0}) \geq M_{i,1}\mathbb{P}_i(\mathcal{F}_{i,1})$ and thus there are a greater number of expected emigrants at time $t = 0$ relative to time $t = 1$. Now suppose that $M_{i,0} \geq M_{i,1} \geq ... \geq M_{i,t}$ and $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1$ for some arbitrary $t > 1$. By the inductive hypothesis, $M_{i,t-1} \geq M_{i,t}$ and $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) > 1 \Rightarrow \mathbb{P}_i(\mathcal{F}_{i,t-1}) < \mathbb{P}_i(\mathcal{F}_{i,t})$. Thus, $M_{i,t-1}\mathbb{P}_i(\mathcal{F}_{i,t-1}) \geq M_{i,t}\mathbb{P}_i(\mathcal{F}_{i,t})$ and by the claim

$$M_{i,0}\mathbb{P}_i(\mathcal{F}_{i,0}) \geq M_{i,1}\mathbb{P}_i(\mathcal{F}_{i,1}) \geq ... \geq M_{i,t-1}\mathbb{P}_i(\mathcal{F}_{i,t-1}) \geq M_{i,t}\mathbb{P}_i(\mathcal{F}_{i,t})$$

Since for arbitrary t, $M_{i,t}\mathbb{P}_i(\mathcal{F}_{i,t})$ denotes the expected number of emigrants, it follows that the expected number of emigrants is weakly decreasing over time.

A proof of the second claim is similar. As the base case follows identical logic, I consider it obvious and show only the inductive step. Suppose that $M_{i,0} \leq M_{i,1} \leq ... \leq M_{i,t}$ and $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1$ for some arbitrary $t > 1$. Then $M_{i,t-1} \leq M_{i,t}$ and $h(\mathcal{N}_{i,t}, \mathcal{N}_{i,t-1}) < 1 \Rightarrow \mathbb{P}_i(\mathcal{F}_{i,t-1}) > \mathbb{P}_i(\mathcal{F}_{i,t})$. Thus, $M_{i,t-1}\mathbb{P}_i(\mathcal{F}_{i,t-1}) \leq M_{i,t}\mathbb{P}_i(\mathcal{F}_{i,t})$ and by the claim

$$M_{i,0}\mathbb{P}_i(\mathcal{F}_{i,0}) \leq M_{i,1}\mathbb{P}_i(\mathcal{F}_{i,1}) \leq ... \leq M_{i,t-1}\mathbb{P}_i(\mathcal{F}_{i,t-1}) \leq M_{i,t}\mathbb{P}_i(\mathcal{F}_{i,t})$$

Since for arbitrary t, $M_{i,t}\mathbb{P}_i(\mathcal{F}_{i,t})$ denotes the expected number of emigrants, it follows that the expected number of emigrants is weakly increasing over time.

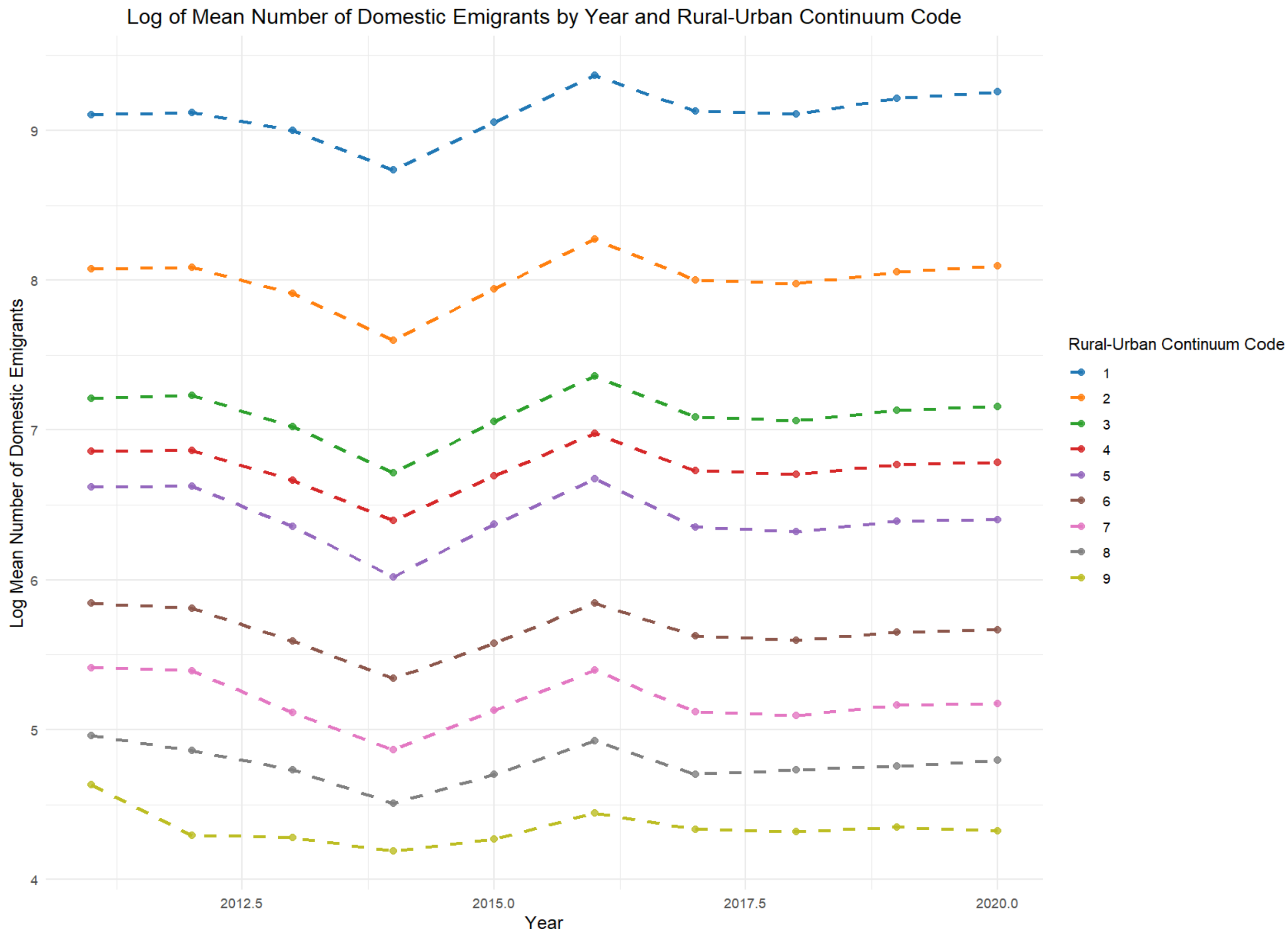

Figure 1: Logarithm of the mean number of emigrants (2011 - 2020) disaggregated by rural-urban continuum code. Note that a continuum code of 1 denotes the most metropolitan counties while a code of 9 denotes the most rural counties.

B Additional Tables and Figures

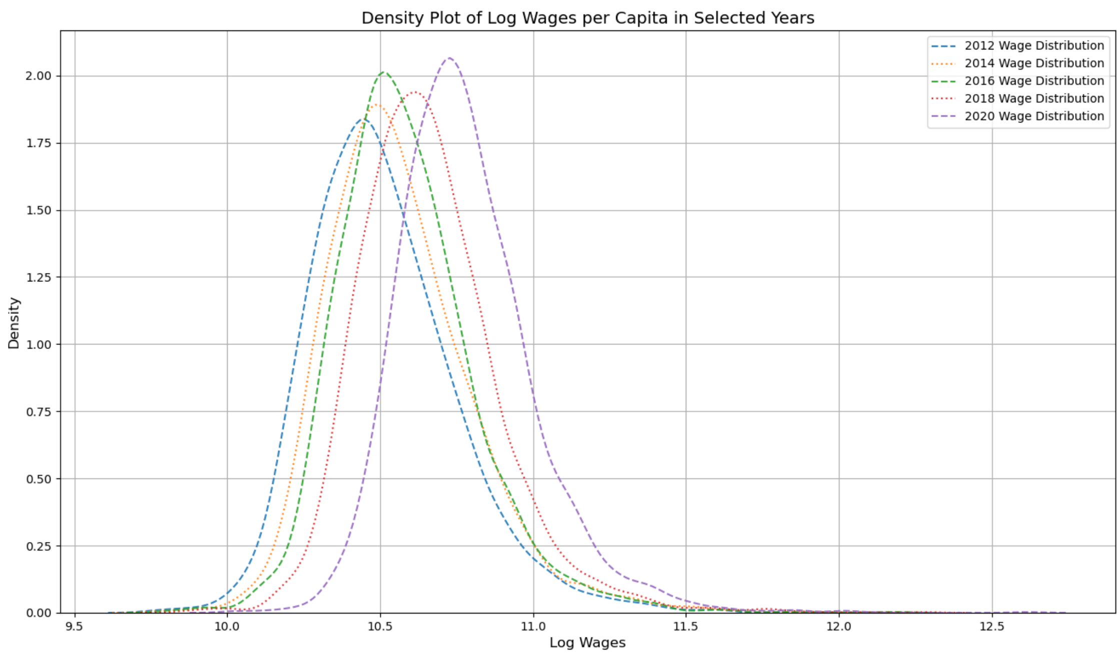

Figure 2: Density plot of log wages per capita, at the county-level, for 2012, 2014, 2016, 2018 and 2020

Dependent Variable:	Number of Emigrant Households	
Model:	(5)	(6)
Binned Events		
10-20	940.6***	832.9***
	(82.83)	(87.79)
20-30	1,765.5***	1,574.1***
	(138.5)	(144.4)
30-40	2,941.4***	2,702.1***
	(273.0)	(283.4)
40-50	4,654.3***	4,059.5***
	(502.4)	(520.2)
50-60	7,157.5***	6,685.5***
	(782.1)	(842.9)
60-70	7,564.6***	6,843.6***
	(1,143.4)	(1,263.6)
70-80	7,928.6***	8,153.2***
	(1,616.6)	(1,742.1)
80-90	13,735.9***	14,809.9***
	(3,171.3)	(3,477.4)
90-100	10,099.4***	8,515.5***
	(1,674.9)	(1,588.0)
100-110	10,765.1***	11,618.7***
	(2,556.6)	(2,951.3)
110-120	11,332.6***	7,616.0***
	(2,502.3)	(1,594.3)
120-130	8,758.7***	8,754.5***
	(2,984.2)	(3,086.7)
130-140	3,519.2	-17.61
	(3,224.7)	(4,000.5)
140-150	40,119.0**	39,393.7**
	(19,128.3)	(18,952.9)
>150	16,482.1***	11,028.1***
	(4,833.2)	(4,221.9)
County Adjusted Gross Income†		-95.42
		(241.1)
Origin Wages		0.1141***
		(0.0118)
Business Net Income†		-103.4**
		(42.23)
Amount of Taxable Pensions†		-0.7999**
		(0.3372)
Number of Farm Returns†		-0.0084
		(0.0080)
Fixed-effects		
State-Year	Yes	Yes
Rural-Urban Continuum Code	Yes	Yes
Fit statistics		
Observations	28,736	25,438
R^2	0.31373	0.33966
Within R^2	0.05673	0.08793

Clustered (County-year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2: Destination ambivalent estimation of the effect of the binned number of climate-related shocks in the origin county on emigration. Number of emigrant households are measured in unit steps. Note that (†) denotes a variable whose value is rooted to its 2011 value such that it is then expressed as a real-valued multiple of the 2011 value. Standard errors are clustered at the joint origin-year level.

Dependent Variable:	Number of Emigrant Households					
Model:	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	OLS	OLS	OLS	OLS	Poisson
Variables						
Number of Origin Events	0.9682***	0.9161***	0.9400***	0.9401***	-0.0168	0.0003***
	(0.0828)	(0.0847)	(0.0860)	(0.0860)	(0.0149)	(6.8×10^{-5})
Number of Destination Events	0.9124***	0.8757***	0.8843***	0.8848***	0.0051	0.0003***
	(0.0828)	(0.0865)	(0.0880)	(0.0880)	(0.0138)	(6.3×10^{-5})
ln(distance)	-125.9***	-128.7***	-125.2***	-125.2***	-80.99	-0.8138
	(4.422)	(4.591)	(4.459)	(4.460)	(1,486,910.9)	(31,256.8)
Origin Wages		0.0013***	0.0014***	0.0014***	0.0011***	4.24×10^{-6}***
		(0.0003)	(0.0003)	(0.0003)	(0.0001)	(6.67×10^{-7})
Destination Wages		0.0012***	0.0011***	0.0011***	-0.0002	-3.84×10^{-6}***
		(0.0002)	(0.0002)	(0.0002)	(0.0001)	(4.82×10^{-7})
Origin County Adjusted Gross Income†		8.423	9.334	9.272	27.26***	0.2455***
		(12.89)	(13.11)	(13.11)	(4.393)	(0.0255)
Destination County Adjusted Gross Income†		18.15*	43.85***	43.85***	15.04***	0.1230***
		(10.29)	(11.31)	(11.31)	(2.857)	(0.0179)
Origin Number of Farm Returns†		-0.0006	-0.0005	-0.0005	2.05×10^{-5}	-1.33×10^{-7}
		(0.0007)	(0.0007)	(0.0007)	(5.14×10^{-5})	(2.31×10^{-7})
Destination Number of Farm Returns†		-0.0003	-0.0003	-0.0003	1.4×10^{-5}	7.35×10^{-9}
		(0.0010)	(0.0010)	(0.0010)	(7.69×10^{-5})	(1.97×10^{-7})
Origin Taxable Pension Amount†			-0.0748	0.0129	-15.74***	-0.0235
			(0.0602)	(0.0630)	(5.162)	(0.0259)
Destination Taxable Pension Amount†			-85.66***	-85.74***	43.37***	0.3731***
			(22.10)	(22.10)	(5.311)	(0.0253)
Origin Unemployment Compensation†				-0.0882***	0.0039	-0.0002
				(0.0194)	(0.0033)	(0.0001)
Destination Unemployment Compensation†				-0.0004***	5.87×10^{-6}**	1.25×10^{-7}***
				(2.02×10^{-5})	(2.52×10^{-6})	(1.93×10^{-8})
Fixed-effects						
Destination State-Year	Yes	Yes	Yes	Yes	Yes	Yes
Origin State-Year	Yes	Yes	Yes	Yes	Yes	Yes
Origin Rural-Urban Continuum Code	Yes	Yes	Yes	Yes	Yes	Yes
Destination Rural-Urban Continuum Code	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Destination Pair					Yes	Yes
Fit statistics						
Observations	572,620	504,728	410,372	410,372	410,372	410,372
Squared Correlation	0.10751	0.11065	0.10875	0.10876	0.96600	0.99143
Pseudo R^2	0.00759	0.00783	0.00771	0.00772	0.22660	0.97330
BIC	8,530,830.0	7,509,813.7	6,086,230.2	6,086,250.9	6,096,751.9	4,452,506.2

Clustered (Origin-Destination Pair) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3: Dyadic estimation of the effect of the number of origin climate-related shocks on emigration. Number of emigrant households are measured in unit steps. Note that (†) denotes a variable whose value is rooted to its 2011 value such that it is then expressed as a real-valued multiple of the 2011 value. Standard errors are clustered at the origin-destination pair level. Note that Model (12) is a restimation of Model (11) using a Poisson Pseudo-Maximum Likelihood estimator.

Dependent Variable:	Number of Emigrant Households		
Model:	(13)	(14)	(15)
	OLS	OLS	Poisson
Origin			
>150	-4.109**	-6.015**	0.0391***
	(1.842)	(2.570)	(0.0107)
10-20	0.3788	-0.2467	0.0030
	(0.4844)	(0.5678)	(0.0021)
100-110	3.799**	4.801**	0.0609***
	(1.595)	(2.336)	(0.0093)
110-120	-0.4197	-2.091	0.0304***
	(1.551)	(2.141)	(0.0109)
120-130	-0.0358	-1.401	0.0070
	(2.819)	(4.357)	(0.0286)
130-140	-6.445***	-15.39***	-0.0667***
	(2.243)	(2.561)	(0.0214)
140-150	3.499*	-11.60**	-0.0076
	(1.838)	(4.949)	(0.0151)
20-30	0.8010	-1.069	-0.0052*
	(0.5018)	(0.6649)	(0.0030)
30-40	-0.9543	-2.678***	-0.0097**
	(0.6950)	(0.9062)	(0.0046)
40-50	1.803**	0.1099	0.0112**
	(0.8788)	(1.063)	(0.0045)
50-60	2.717**	3.499**	0.0339***
	(1.109)	(1.656)	(0.0078)
60-70	2.027*	3.477**	0.0352***
	(1.138)	(1.545)	(0.0070)
70-80	-4.870***	-9.562***	-0.0228***
	(1.786)	(2.468)	(0.0066)
80-90	1.670	2.086	0.0301***
	(2.077)	(2.198)	(0.0085)
90-100	-2.229	-3.263	0.0328***
	(2.105)	(3.140)	(0.0104)
Destination			
>150	0.3763	-3.495	0.0420***
	(1.878)	(2.459)	(0.0116)
10-20	-0.0491	-0.0900	-0.0004
	(0.3834)	(0.5208)	(0.0026)
100-110	1.293	0.7326	0.0129
	(1.393)	(1.820)	(0.0089)
110-120	0.6702	-3.372*	0.0150
	(1.655)	(1.744)	(0.0109)
120-130	-2.749	-6.443*	-0.0353**
	(2.405)	(3.324)	(0.0153)
130-140	-1.893	-3.785	-0.0268
	(2.720)	(3.789)	(0.0314)
140-150	-4.327**	6.637***	0.0500***
	(2.004)	(2.391)	(0.0120)
20-30	-0.2693	-1.150*	-0.0083**
	(0.5154)	(0.6746)	(0.0039)

30-40	-0.0571	-0.6700	0.0019
	(0.7439)	(0.9320)	(0.0050)
40-50	-0.7120	-1.984*	-0.0079
	(0.8212)	(1.022)	(0.0049)
50-60	1.596	3.681**	0.0287***
	(1.008)	(1.584)	(0.0063)
60-70	0.5606	-0.0405	-0.0014
	(1.086)	(1.311)	(0.0056)
70-80	0.0087	-0.4784	0.0226***
	(1.635)	(2.202)	(0.0087)
80-90	2.878*	2.423	0.0245***
	(1.734)	(1.905)	(0.0075)
90-100	-4.611	-3.095	0.0330***
	(3.035)	(3.993)	(0.0102)
ln(distance)	-39.94	-57.47	0.1564
	(1,475,511.0)	(1,495,225.4)	(30,569.3)
Origin Wages		0.0011***	4.4×10^{-6}***
		(0.0002)	(6.71×10^{-7})
Destination Wages		-0.0002*	-3.82×10^{-6}***
		(0.0001)	(4.82×10^{-7})
Origin County Adjusted Gross Income†		27.39***	0.2495***
		(4.387)	(0.0256)
Destination County Adjusted Gross Income†		15.38***	0.1242***
		(2.906)	(0.0183)
Origin Number of Farm Returns†		1.85×10^{-5}	-1.5×10^{-7}
		(5.15×10^{-5})	(2.36×10^{-7})
Destination Number of Farm Returns†		1.45×10^{-5}	1.21×10^{-8}
		(7.69×10^{-5})	(1.97×10^{-7})
Origin Taxable Pension Amount†		-16.26***	-0.0234
		(5.128)	(0.0253)
Destination Taxable Pension Amount†		42.55***	0.3688***
		(5.305)	(0.0257)
Origin Unemployment Compensation†		0.0030	-0.0002*
		(0.0027)	(0.0001)
Destination Unemployment Compensation†		2.07×10^{-5}***	2.03×10^{-7}***
		(3.86×10^{-6})	(2.27×10^{-8})
Fixed-effects			
Destination State-Year	Yes	Yes	Yes
Origin State-Year	Yes	Yes	Yes
Origin Rural-Urban Continuum Code	Yes	Yes	Yes
Origin Rural-Urban Continuum Code	Yes	Yes	Yes
Origin-Destination Pair	Yes	Yes	Yes
Fit statistics			
Observations	572,620	410,372	410,372
Squared Correlation	0.97193	0.96602	0.99155
Pseudo R^2	0.23841	0.22663	0.97333
BIC	7,986,951.4	6,096,939.4	4,448,897.1

Clustered (Origin-Destination Pair) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4: Dyadic estimation of the effect of the binned number of origin climate-related shocks on emigration. Number of emigrant households are measured in unit steps. Note that (†) denotes a variable whose value is rooted to its 2011 value such that it is then expressed as a real-valued multiple of the 2011 value. Standard errors are clustered at the origin-destination pair level. Note that Model (12) is a restimation of Model (11) using a Poisson Pseudo-Maximum Likelihood estimator.

Dependent Variable:	Log Mean Annual Destination Income per capita		
Model:	(16)	(17)	(18)
Variables			
Number of Origin Events	-0.0001***		
	(2.98×10^{-5})		
Number of Destination Events	0.0003***		
	(2.85×10^{-5})		
ln(distance)	0.0217***	0.0315***	0.6044
	(0.0008)	(0.0027)	(2,378.1)
Origin County Adjusted Gross Income†	0.0147*	0.0117	0.0715
	(0.0084)	(0.0085)	(0.0546)
Destination County Adjusted Gross Income†	0.5693***	0.5621***	0.1731***
	(0.0388)	(0.0385)	(0.0185)
Origin Number of Farm Returns†	8.09×10^{-8}	1.29×10^{-7}	4.48×10^{-8}
	(3.02×10^{-7})	(3.09×10^{-7})	(8.7×10^{-8})
Destination Number of Farm Returns†	$1.28 \times 10^{-6***}$	$1.3 \times 10^{-6***}$	2.24×10^{-8}
	(1.63×10^{-7})	(1.85×10^{-7})	(1.13×10^{-7})
Origin Taxable Pension Amount†	$-7.82 \times 10^{-5**}$	$-8.14 \times 10^{-5***}$	-0.0422
	(3.06×10^{-5})	(3.1×10^{-5})	(0.0293)
Destination Taxable Pension Amount†	-0.2951***	-0.2902***	0.0132
	(0.0241)	(0.0242)	(0.0557)
Origin Unemployment Compensation†	4.56×10^{-5}	$5.47 \times 10^{-5*}$	$-2.41 \times 10^{-5*}$
	(3.05×10^{-5})	(3.03×10^{-5})	(1.28×10^{-5})
Destination Unemployment Compensation†	$3.97 \times 10^{-7***}$	$4.15 \times 10^{-7***}$	$2.09 \times 10^{-7***}$
	(3.63×10^{-8})	(3.55×10^{-8})	(3.3×10^{-8})
Fitted Number of Emigrant Households		$8.07 \times 10^{-5***}$	-0.0014
		(2.3×10^{-5})	(0.0012)
Fixed-effects			
Destination State-Year	Yes	Yes	Yes
Origin State-Year	Yes	Yes	Yes
Origin Rural-Urban Continuum Code	Yes	Yes	Yes
Destination Rural-Urban Continuum Code	Yes	Yes	Yes
Origin-Destination Pair			Yes
Fit statistics			
Observations	415,656	415,656	415,656
R^2	0.52198	0.51376	0.82688
Within R^2	0.10396	0.08854	-19.846

Clustered (pair) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 5: 2 Stage least squares and reduced form estimation of the effect of the number of origin and destination climate-related shocks on wages in the destination county. Number of emigrant households are measured in unit steps. Note that (†) denotes a variable whose value is rooted to its 2011 value such that it is then expressed as a real-valued multiple of the 2011 value. Standard errors are clustered at the origin-destination pair level.

Dependent Variable:	Number of Emigrant Households	
Model:	(19)	(20)
Variables		
Flash Flood	312.4***	308.1***
	(28.10)	(31.08)
Flood	109.2***	97.46***
	(26.70)	(32.40)
Hail	93.62***	93.89***
	(11.67)	(12.56)
Thunderstorm Wind	60.94***	49.12***
	(11.90)	(12.79)
Tornado	56.09	56.90
	(47.48)	(51.45)
Funnel Cloud	540.2***	486.7***
	(127.3)	(125.0)
Heavy Rain	-79.11***	-43.72
	(25.38)	(27.58)
Dust Devil	2,823.9***	3,025.6***
	(695.8)	(736.9)
Debris	79.77	-134.9
	(213.9)	(189.8)
Lightning	1,385.8***	1,265.8***
	(148.3)	(158.0)
High Wind	-426.2	-401.6
	(1,308.4)	(1,353.4)
Fog	2,591.1***	
	(335.2)	
Heat	-641.9	-633.5
	(613.5)	(513.1)
Dust Storm	250.5	635.2***
	(183.2)	(193.9)
Tropical Storm	996.5	1,373.3
	(6,003.6)	(6,010.2)
Strong Wind	1,159.5**	71.39
	(550.1)	(573.3)
County Adjusted Gross Income†		-79.27
		(253.5)
Origin Wages		0.1030***
		(0.0115)
Business Net Income†		-106.3***
		(39.12)
Amount of Taxable Pensions†		-0.6123*
		(0.3474)
Number of Farm Returns†		-0.0047
		(0.0082)
Fixed-effects		
state-year	Yes	Yes
Rural-Urban Continuum Code	Yes	Yes
Fit statistics		
Observations	27,167	24,117
R^2	0.34931	0.36805
Within R^2	0.08925	0.11113

Clustered (county-year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6: Destination ambivalent estimation of the effect of disparate types of climate-related shocks in the origin county on emigration. Number of emigrant households are measured in unit steps. Note that (†) denotes a variable whose value is rooted to its 2011 value such that it is then expressed as a real-valued multiple of the 2011 value. Standard errors are clustered at the origin county-year level.

	p-values	FDR Correction	Significance
>150	0.00900	0.01286	$*$
10-20	0.00000	0.00000	$*$
100-110	0.00008	0.00014	$*$
110-120	0.00000	0.00000	$*$
120-130	0.00457	0.00703	$*$
130-140	0.99649	0.99649	
140-150	0.03767	0.04432	$*$
20-30	0.00000	0.00000	$*$
30-40	0.00000	0.00000	$*$
40-50	0.00000	0.00000	$*$
50-60	0.00000	0.00000	$*$
60-70	0.00000	0.00000	$*$
70-80	0.00000	0.00001	$*$
80-90	0.00002	0.00004	$*$
90-100	0.00000	0.00000	$*$
County Adjusted Gross Income	0.69232	0.72875	
Origin Wages	0.00000	0.00000	$*$
Business Net Income	0.01433	0.01911	$*$
Amount of Taxable Pensions	0.01770	0.02213	$*$
Number of Farm Returns	0.29638	0.32931	

Table 7: p-values with and without the false detection rate correction corresponding to Model (6). Note that those coefficients which are significant at a 95% level following the false detection rate correction are denoted with a $*$ in the significance column.

	p-values	FDR Correction	Significance
Flash Flood	0.00000	0.00000	$*$
Flood	0.00263	0.00584	$*$
Hail	0.00000	0.00000	$*$
Thunderstorm Wind	0.00012	0.00035	$*$
Tornado	0.26881	0.38401	
Funnel Cloud	0.00010	0.00033	$*$
Heavy Rain	0.11302	0.18837	
Dust Devil	0.00004	0.00016	$*$
Debris	0.47716	0.63621	
Lightning	0.00000	0.00000	$*$
High Wind	0.76665	0.85183	
Heat	0.21691	0.33371	
Dust Storm	0.00106	0.00264	$*$
Tropical Storm	0.81927	0.86239	
Strong Wind	0.90090	0.90090	
County Adjusted Gross Income	0.75448	0.85183	
Origin Wages	0.00000	0.00000	$*$
Business Net Income	0.00660	0.01320	$*$
Amount of Taxable Pensions	0.07797	0.14177	
Number of Farm Returns	0.56540	0.70675	

Table 8: p-values with and without the false detection rate correction corresponding to Model (20). Note that those coefficients which are significant at a 95% level following the false detection rate correction are denoted with a $*$ in the significance column.

Event Type	Number of Occurrences
Flash Flood	53099
Flood	37161
Hail	139901
Thunderstorm Wind	234542
Tornado	19327
Lightning	6111
Funnel Cloud	4484
Heavy Rain	19018
Dust Devil	123
Debris Flow	1405
High Wind	7
Drought	1
Dense Fog	1
Heat	5
Dust Storm	3
Tropical Storm	2
Excessive Heat	1
Strong Wind	3
Heavy Snow	5
Wildfire	1

Table 9: Number of events by type between 2011 and 2021 in US counties. Data is aggregated from the NOAA storm events database. Note that this does not include all weather events in counties in the US between 2011 and 2021 but rather only includes those which are defined as (1) having sufficient intensity to cause significant damage or other harms; (2) being rare in a given location; or (3) representing a local extrema in the type of event according to NOAA.

Year	Group	Number of Counties
2011	20-70	876
2011	70-120	31
2011	<20	2137
2011	>120	2
2012	20-70	496
2012	70-120	14
2012	<20	2529
2013	20-70	429
2013	70-120	14
2013	<20	2403
2014	20-70	365
2014	70-120	15
2014	<20	2379
2015	20-70	425
2015	70-120	20
2015	<20	2391
2016	20-70	412
2016	70-120	7
2016	<20	2467
2016	>120	2
2017	20-70	484
2017	70-120	16
2017	<20	2350
2018	20-70	418
2018	70-120	21
2018	<20	2381
2018	>120	2
2019	20-70	528
2019	70-120	30
2019	<20	2272
2019	>120	5
2020	20-70	431
2020	70-120	23
2020	<20	2356
2020	>120	2

Table 10: Number of climate-related shocks by year in origin counties. Note that group refers to the binning of county level observations of events into 1 of the following 3 groups: (1) < 20 events, (2) $20 - 70$ events, (3) $70 - 120$ events and (4) > 120 events

Year	Number of Households	P_{25}	P_{50}	P_{75}	Standard Deviation	Mean
2011	6213920.00	152.00	384.00	1142.75	6518.47	2040.03
2012	6278252.00	148.00	381.00	1157.00	6537.99	2065.89
2013,	5397206.00	128.00	344.50	1078.50	6040.31	1896.42
2014	4077939.00	107.00	281.00	861.50	4832.16	1478.05
2015	5635951.00	128.00	346.00	1103.25	6410.49	1987.29
2016	7720963.00	160.00	440.00	1444.75	8643.61	2673.46
2017	6028155.00	131.00	354.00	1143.75	6909.30	2115.14
2018	5903777.00	127.00	348.50	1130.75	6805.09	2092.05
2019	6462009.00	140.00	369.00	1199.50	7671.38	2279.37
2020	6731320.00	142.00	372.00	1216.50	8145.44	2391.23

Table 11: Summary statistics of the number of households emigrating within the US (2011 - 2020). Note that while the data covers 2021, it is not included in this table as only destination specific terms can be determined for 2021 variables. Thus summary statistics on emigration in 2021 are not supported by this dataset.

Dependent Variable:	Number of Emigrant Households			
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	22,182.4***			
	(6,370.8)			
02013 × Year	-11.00***	-4.792***	6.947	6.947
	(3.164)	(2.65×10^{-15})	(10.22)	(10.40)
02016 × Year	-10.98***	-6.212***	-2.473	-2.473
	(3.161)	(7.24×10^{-16})	(6.642)	(6.759)
02050 × Year	-10.93***	-5.479***	-3.175	-3.175
	(3.161)	(3.73×10^{-16})	(6.257)	(6.367)
02060 × Year	-11.02***	-1.000***	64.62	64.62
	(3.168)	(2.41×10^{-16})	(46.03)	(46.83)
02068 × Year	-10.99***	-0.5471***	13.27	13.27
	(3.162)	(7.02×10^{-17})	(10.12)	(10.29)
02070 × Year	-10.99***	-1.630***	1.900	1.900
	(3.161)	(2.33×10^{-16})	(6.559)	(6.674)
02090 × Year	-10.08***	-66.53***	-62.49***	-62.49***
	(3.161)	(5.6×10^{-15})	(6.685)	(6.802)
02100 × Year	-11.02***	1.000***	20.01	20.01
	(3.167)	($NaN \times 10^{-Inf}$)	(13.34)	(13.57)
02105 × Year	-11.00***	0.1456***	2.034	2.034
	(3.163)	(1.17×10^{-17})	(6.702)	(6.819)
02110 × Year	-10.87***	-26.81***	-22.46***	-22.46***
	(3.161)	(1.87×10^{-15})	(6.937)	(7.058)
02122 × Year	-10.73***	-33.50***	-31.10***	-31.10***
	(3.161)	(2.24×10^{-15})	(6.877)	(6.998)
02130 × Year	-10.97***	-8.988***	-6.005	-6.005
	(3.161)	(1.77×10^{-15})	(6.731)	(6.849)
02150 × Year	-10.94***	-11.00***	-7.280	-7.280
	(3.161)	(1.68×10^{-15})	(6.189)	(6.297)
02164 × Year	-11.00***	0.7165***	3.346	3.346
	(3.162)	($NaN \times 10^{-Inf}$)	(4.880)	(4.965)
02170 × Year	-10.41***	-28.96***	-26.97***	-26.97***
	(3.161)	(6.72×10^{-15})	(6.872)	(6.992)
02180 × Year	-10.97***	-3.382***	0.7411	0.7411
	(3.161)	(1.87×10^{-16})	(6.202)	(6.311)
02185 × Year	-10.97***	-4.521***	-12.35	-12.35
	(3.161)	(3.73×10^{-16})	(9.096)	(9.255)
02188 × Year	-10.97***	-2.697***	1.274	1.274
	(3.161)	(1.12×10^{-15})	(5.789)	(5.890)
02195 × Year	-11.01***	-6.214***	0.3817	0.3817
	(3.166)	(1.03×10^{-15})	(14.05)	(14.29)
02198 × Year	-11.00***	-0.4059***	5.016	5.016
	(3.161)	(1.53×10^{-16})	(6.935)	(7.056)
02240 × Year	-10.97***	-6.103***	-1.555	-1.555
	(3.161)	(6.54×10^{-16})	(6.133)	(6.240)
02261 × Year	-10.94***	-11.19***	-9.294	-9.294
	(3.163)	(1.47×10^{-15})	(7.261)	(7.388)
02270 × Year	-10.99***	-15.50***		
	(3.167)	(3.85×10^{-15})		
02275 × Year	-11.02***			
	(3.169)			
02282 × Year	-11.03***			
	(3.169)			
02290 × Year	-10.96***	-2.406***	2.172	2.172
	(3.161)	(1.87×10^{-16})	(5.506)	(5.602)
06003 × Year	-11.02***			
	(3.167)			
08065 × Year	-10.97***	0.2441***	5.594	5.594
	(3.161)	(8.67×10^{-16})	(7.780)	(7.916)
16025 × Year	-11.02***			
	(3.167)			
46113 × Year	-10.98***	-2.000***		
	(3.167)	(4.81×10^{-16})		
48211 × Year	-11.01***			
	(3.169)			
49009 × Year	-11.02***			
	(3.167)			
49033 × Year	-11.02***			
	(3.167)			
51735 × Year	-10.91***	3.006***		
	(3.161)	(7.47×10^{-16})		
County Adusted Gross Income†			13.86	13.86
			(42.66)	(43.40)
Number of Farm Returns†			-0.0003	-0.0003
			(0.0003)	(0.0003)
Wages			-0.0041	-0.0041
			(0.0029)	(0.0029)
Fixed-effects				
County		Yes	Yes	Yes
Rural-Urban Continuum Code				Yes
Fit statistics				
Observations	220	220	204	204
R^2	0.93383	0.94364	0.94395	0.94395
Within R^2		0.19679	0.20256	0.20256

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 12: Test for unit specific trends amongst never shocked counties. Note that (†) denotes that a variable is rooted to its 2011 value and thus is expressed as a real-valued multiple of said 2011 value. Numbers $(XXXXX) \times Year$ denote the county FIPS code.