Escape the Heat: The Dynamics of Migration as Adaptation to Climate Change

Robert Baluja*

This Version: January 8, 2025

Most Recent Version

Earth's climate is changing, which is widely expected to drive net reductions to human welfare. In this paper, I study how effectively migration will reduce experienced climate damages. To provide answers to my research questions, I develop and estimate a dynamic lifecycle model of migration within Mexico. I combine this with a non-stationary and spatially varying model of the climate, in which I allow for both fully informed and naive expectations of the future progression of climate change. Estimation of the climate model uses daily-level historical weather data and output from state-of-the-art climate simulations. Estimation of the lifecycle model uses a sample of life histories, covering the years 1950–2019, and follows a nested full solution pseudo-maximum likelihood routine. I find that climate damages from business-as-usual warming would be 28% higher if domestic migration within Mexico was no longer available as a tool of adaptation to climate change. Moreover, the fraction of the population that I estimate as forming naive expectations of the climate system would experience an average of 2% less lifetime climate damages from becoming fully informed on the climate transition. Given that most of the increased damages this population faces come from a reduced propensity to migrate, one way to reduce these losses is to subsidize migration. I find that subsidizing migration at the average level of the internality reduces their welfare losses by 8–19%. The exact value of this reduction depends on whether the policy forces people to use the subsidy in a particular period. Policies that allow individuals to choose *when* to use them are over twice as valuable to the affected population because they do not overly incentivize dynamically suboptimal moves. This sort of dynamically-available policy is common; examples include provisions from the recent Inflation Reduction Act and first-time homeowners tax credits.

JEL Codes: D15, D84, J61, Q54, Q56

Keywords: climate change, dynamics, migration, information, adaptation, lifecycle

---

*University of Arizona, Department of Economics. Email: RobertBaluja@gmail.com. I am deeply indebted to my advisor, Ashley Langer, for continuous support and guidance. I thank Fiona Burlig, Christian Cox, Jonathan Dingel, Price Fishback, Gautam Gowrisankaran, Konan Hara, Hidehiko Ichimura, Sarah Jacobson, Derek Lemoine, Alvin Murphy, Juan Pantano, Giovanni Peri, Stanley Reynolds, Ivan Rudik, Pietro Tebaldi, Christopher Timmins, Matthijs Wildenbeest, the attendees of the Arizona ENREE workshop and the 2024 AERE Meetings, and the faculty and attendees of the 2nd Summer School on the Economics of Migration, for insightful comments. This work used High Performance Computing (HPC) resources supported by the University of Arizona TRIF, UITS, and Research, Innovation, and Impact (RII) and maintained by the UArizona Research Technologies department. All errors are my own.

1 Introduction

Earth's climate is changing in ways that modern humanity has never experienced. We are already observing large-scale changes to the physical systems governing our oceans and the atmosphere, resulting in changes to the frequency and severity of extreme heat and precipitation events throughout the planet (IPCC, 2022). Although we have yet to understand the full extent of these changes on the habitability of the many regions of the world (Sherwood and Huber, 2010; Mora et al., 2017), history tells us that humans will find a way to adapt to our new environment. One of the most prominent ways that we have ancestrally adapted to environmental change, and specifically to climate change, is through migration (Willis et al., 2004; Carto et al., 2009; Timmermann and Friedrich, 2016).

In this paper, I ask how valuable migration will be as a tool to mitigate the damages associated with *future* changes to the climate. To answer this question, one must first understand how the decision to migrate will be impacted by the changing climate. I develop and estimate, using a full solution routine, a microfounded dynamic model of migration within Mexico in response to climate change. Individuals are modeled as making a series of sequential decisions of where to live within Mexico. Each time they make this decision, they understand that they will live in their chosen destination until they have the opportunity to move again. The dynamic nature of the decision of where to live requires that individuals form expectations about the future climate in each of their potential destinations. Given the complexity of the climate system, I allow individuals to form either fully-informed or naive expectations about a climate system that is changing at different rates throughout the country and over time.

Migration is both costly and a difficult decision to reverse (Kennan and Walker, 2011). Moreover, past work has highlighted that a sizeable fraction of the population would experience prohibitively high psychological costs from moving (Koşar et al., 2022). Completely accounting for these costs is critical when seeking to understand the effectiveness of migration as an adapting strategy to the damages of climate change. As such, the model allows individuals to have such high psychological frictions from migration that they never consider it as an option. Those that do consider the trade-offs of migration make their decision of whether and where to move based on a comparison of the expected benefit of moving to a particular location, accrued across the remainder of their lifetime, with the psychological and monetary cost associated with making such a move today.

I find that domestic migration within Mexico is an important mechanism for limiting the damages of climate change. As of the 2020 Census, the average 17–19-year-old stands to face 28% less climate damages, from business-as-usual warming, over their lifetime, because of their ability to migrate. The average reduction in damages masks important variation across the population. For those living in the warmest regions of the country, the ability to migrate is projected to reduce their expected lifetime welfare losses from business-as-usual warming by 35%, as compared to the 3%

reduction faced by those already living in the temperate regions. Those employed in agricultural sectors and those with the lowest levels of education stand to face reductions of 32%, compared to the 28% reduction faced by those in other sectors and those with at least 12 years of education. As time, and as a result, warming, progresses, our ability to migrate in response to climate change will become increasingly important. Those who were less than 2-years-old in 2020 are projected to experience an average of 33% less damages over their lifetime from business-as-usual warming.

Our ability to migrate in response, or more broadly dynamically adapt, to climate change rests on our ability to correctly anticipate future changes to the climate. This has the potential to be a cause for concern, as recent international surveys have highlighted that although a large fraction of the population understands that climate change is happening, they also report a lack of understanding of it (Leiserowitz et al., 2022). Using the estimated model, I find that the projected lifetime climate damages, on average, would be 2% less for the climate-naive if they were to become fully-informed.

The majority of the welfare losses to the climate-naive come from a reduced propensity to move, rather than from the decision to move to suboptimal locations. Indeed, relative to the decisions one would make with full information, I find that, for the climate-naive, mistakenly choosing to remain in one's location leads to an average loss to lifetime welfare of $9,007. This is much larger than the average welfare losses from making a mistake during one's move: $3,796. Moreover, the climate-naive are more than three-times as likely to mistakenly remain in their location as they are to mistakenly move. Taken together, this implies that one way to reduce the welfare losses to this population is to subsidize domestic migration.[1] I find that such policy has the potential to be quite effective: subsidizing migration at the value to the climate-naive from becoming fully-informed reduces the lifetime welfare losses borne by the climate-naive by between 8% and 19%.

The differences in the effectiveness of policies which affect moving costs are driven by the dynamic structure of the different policy options. More specifically, dynamic policies create value to the population by giving individuals the flexibility to choose to move when it is most valuable for them to do so. This additional flexibility causes dynamic policies to positively select from the portion of the population marginal to a similarly-priced static policy. The selection effects of dynamic policies are of first-order concern: differences in *who* decides to use the different policies causes dynamically-minded reductions to moving costs to be over twice as valuable to movers as similar static policies. Moreover, the value placed on the ability to move in the future drives a similar increase in the value of the policy to the climate-naive.

To answer my research questions, I estimate the lifecycle model using rich microdata from the Mexican Migration Project (MMP). These data contain information on the life histories of more than 11,000 individuals living and working in Mexico from 1950 to 2019. Given the large changes

[1] Such policies are commonplace: Canada, the United Kingdom, and the United States have tax codes which historically allow for tax-deductible moving expenses.

to the climate in Mexico over this time period, and similar to the long-difference strategy employed in the reduced-form literature (Burke and Emerick, 2016), I exploit true change in the climate over time to identify the relevant primitives of the model. For example, I identify the proportion of the population forming naive and fully-informed expectations about the climate system using data on observed individual decisions along with variation in the rates of warming across both space and time.

Estimation of the model primitives follows a two-step procedure. In the first stage, I use high-quality gridded weather data to estimate a non-stationary model of the climate system that varies throughout the country. I then estimate heterogeneous mappings from the relevant climate variables (average daily temperature and agricultural-season precipitation) to various statistics of weather that are designed to differentially capture the impact of extreme heat on crop yields and human well-being, along with the risk of drought and flood. I supplement the MMP data with historical surveys on wages to estimate a wage equation that depends on individual demographics, sector of employment, weather, and the location of residence. I take these estimated mappings to the second stage, where I estimate the model primitives. I employ a nested full-solution pseudo maximum likelihood routine (Rust, 1987), taken over a finite mixture (Heckman and Singer, 1984), to estimate the model primitives: the preference parameters and the mixing probabilities for the unobserved moving and belief types.

My estimation framework allows for an internally consistent model of climate-induced migration that avoids the need to calibrate anything other than the discount factor. I combine the estimated model with data from the 2020 full-count Census and high-quality future climate projections under a business-as-usual scenario (Thrasher et al., 2022) to simulate the choices for each Census-enumerated male living in Mexico in 2020, across the remainder of his lifetime. This framework allows me to make changes to the environment, through changes to the climate system and the moving costs, to recover welfare metrics of interest. For example, once the model is estimated, to estimate the value of migration one need only calculate average welfare under both the status-quo and in a counterfactual environment where moving is completely restricted

Relation to the Literature

This article builds on three main streams of the literature. First, I contribute to the literature seeking to understand the damages of climate change.[2] This body of work can be broken into two primary branches: a reduced form and a general equilibrium literature. The reduced form work in this sphere has made use of various econometric tools to identify how economic outcomes of interest will be affected by climate change. For example, Mendelsohn et al. (1994) exploits cross-sectional variation

[2] Carleton et al. (2024) provides an insightful and in-depth analysis of the climate damages and adaptation literature.

in the climate across space to understand how the value of U.S. farmland will evolve with the climate. Later work highlighted that such spatial comparisons are prone to large biases, which fueled the use of panel variation in weather to identify the effect of unanticipated deviations in weather from historical averages on economic outcomes (Schlenker and Roberts, 2009; Feng et al., 2010; Missirian and Schlenker, 2017). While panel variation credibly identifies the relationship between weather shocks and economic outcomes, in many settings it does not capture ex-ante adapting behaviors, which will serve to reduce actual climate damages (Dell et al., 2014; Mendelsohn, 2019; Lemoine, 2024). To account for this, a relatively recent literature studies how the response to weather shocks vary across different climates to identify variation in equilibrium ex-ante adapting behaviors (Dell et al., 2012; Rode et al., 2021; Carleton et al., 2022).

The discussed work specializes to static forms of adaptation (Carleton et al., 2024); however, many potentially valuable types of adapting behaviors are instead driven by long-term expectations over a non-stationary climate system, with these expectations formed heterogeneously by the population of interest. I contribute to this literature by providing novel estimates of the value of dynamic forms of adaptation to climate change, by jointly identifying the cost and benefit of migration as adaptation to climate change.[3] Moreover, I provide an empirical framework that future work can use to similarly study the effectiveness of dynamic forms of adaptation to climate change.

Another branch of this literature has applied general equilibrium tools, which allow them to study the role of adaptation for limiting the damages of climate change through counterfactual analyses (Desmet and Rossi-Hansberg, 2015; Rudik et al., 2022; Cruz and Rossi-Hansberg, 2023). In applying modern tools from the spatial economics literature, this work imposes climate change as a series of steady state to steady state changes, as in the reduced form literature. In contrast, I develop a rich, partial equilibrium framework[4] upon which I am able to exploit microdata[5] to explicitly study the frictions that will inherently be present in between any two steady states. I contribute to this literature in two primary ways. First, I provide a framework which is able to cleanly identify the dynamic drivers behind climate migration, including the large moving costs associated with migration, which are allowed to vary across the population. Second, I highlight the role of expectations about the future climate system on current adapting decisions, by both identifying heterogeneity in these beliefs and by showing how such biases will lead to a reduction in the value of dynamic adaptation to climate change.

[3]Much of the literature has worked towards identifying the benefit of adaptation to climate change. Identifying the costs associated with these actions has remained elusive (Carleton et al., 2024).

[4]As discussed in Section 6, I design a stylized framework to check how general equilibrium forces, through congestion impacts on wages, may affect my primary results. I find that these impacts do not seem to be of first-order concern in my context. Much of this is likely due to the relatively short time horizon I analyze.

[5]Makridis and Ransom (2018) and Mathes (2024) also use microdata to estimate models of migration in response to climate change. Both use data from the United States and assume a stationary climate with no uncertainty, from the purview of the decision makers. I extend their work by allowing for a non-stationary and uncertain climate.

Second, I contribute to the literature which has emphasized the existence of an "adaptation gap" with respect to climate change (Carleton and Hsiang, 2016). Zappalà (2024) is the first, and to my knowledge only, article to highlight that a portion of this gap is likely driven by inconsistent beliefs about the climate. They use cross-sectional data on reported farmer beliefs surrounding recent changes in the frequency of drought to estimate how such beliefs lead to differential irrigation responses to short-term dry spells in Bangladesh. I complement their work by instead using data on individual decisions to recover beliefs on the climate system, similar to Kala (2017). As in Zappalà (2024), I find that naive expectations of the climate system will lead to lower-than-optimal rates of adaptation to climate change. I highlight the generality of this finding, and extend the discussed literature by arguing for the broad role that policy, by subsidizing adapting behaviors, can play in helping to close the "belief gap."

Third, I contribute to the literature seeking to uncover the value of migration, irrespective of climate change. One branch of this literature has used reduced form techniques, generally in the form of quasi-experiments. More specifically, they study how arguably exogenous shifts in the propensity to migrate change lifetime outcomes (Deryugina et al., 2018; Nakamura et al., 2021; Sarvimäki et al., 2022). This literature has found that migration can be quite valuable, however almost all of this work has found that this value varies strongly throughout the population.

Another branch of the literature studying the value of migration has used dynamic discrete choice methods to directly model the moving decision to account for the fact that migration is an endogenous choice (Bishop, 2008; Thom, 2010; Kennan and Walker, 2011; Lessem, 2018). I contribute to this literature by highlighting how the value of migration can change, across the population, over time, in a non-stationary environment, in this case, from climate change. Moreover, similar to a budding literature in industrial organization (e.g., Aguirregabiria and Jeon, 2020), I relax the assumption of full-rationality that is oftentimes present in this general body of work by allowing for different levels of sophistication surrounding expectations of the climate system.

The remainder of this paper is structured as follows: Section 2 describes the data and various reduced-form empirics and summary statistics that motivate the dynamic model. Section 3 presents the dynamic model. Section 4 provides intuition on the estimation strategy and sources for identifying variation in the preference parameters. Section 5 presents the estimated parameters and details how well the model fits reality. Section 6 provides the setup used in counterfactual simulations and the results of these exercises. Section 7 discusses and contextualizes the results of the paper, concludes, and provides avenues for future research on this topic.

2 Data & Empirical Motivation

I use data from the Mexican Migration Project (MMP) surveys to estimate the model described in Section 3. I supplement these data with historical climate data from Livneh et al. (2015) and Thornton et al. (2022); climate projection data from Thrasher et al. (2022); and wage data from the ENIGH and ENE Mexican surveys. Furthermore, in counterfactual simulations, I use data on demographic distributions at the municipality level from the 2020 Mexican Census count.

Data

Mexican Migration Project

I use the life-history files from the Mexican Migration Project surveys to estimate the preference parameters governing the structural model. These surveys contain annual information on the location of employment at the municipality level for a sample of over 28,000 individuals[6] living in Mexico for each year of their life. I also use both demographic and personal characteristics contained in these files, such as their birth location, level of education, and the year of birth for each of their children.

Most climate-induced migration to date has been within a country's border (IPCC, 2022). Therefore, to study only internal migration, I restrict the life histories to individuals who reside within Mexico for their entire recorded history. The publicly available life history data censor many of the municipalities of employment and birth to maintain the confidentiality of the respondents; however, I received the censoring algorithm and used it to fill in these missing observations where possible. I therefore further restrict the data to only those individuals who either have a known municipality of birth or whose municipality of birth was successfully decensored. Similarly, I remove all individuals who, after employing the decensoring and an imputation algorithm, have any observations with missing employment municipalities. After these restrictions, I am left with 11,194 unique individuals, covering the years 1950 through 2019.[7] After aggregating the data from annual to triennial observations,[8] I am left with 90,578 person-year observations for estimation.

Wages

A limitation of the MMP data is that although information on annual location and sector of employment is provided, the survey did not collect information on income. I therefore use data

[6]The MMP data primarily contain the life histories for the head of the household. In Mexico, during my sample period (1950–2019), this is an overwhelmingly male population. For this reason, I focus my analysis on males.

[7]As discussed in Section 3, I begin modeling each individual's choices when they turn 17. As such, I model decisions from 1967 to 2064. The end point of this range represents the year of the final decision for a 17-year-old in 2018.

[8]The aggregation procedure is discussed in Appendix A.1.

from the following two census surveys to estimate a series of wage equations that I then use when estimating the model: La Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) and La Encuesta Nacional de Ocupación y Empleo (ENOE). I use the ENIGH surveys from 1984 to 2006, and the quarterly ENOE surveys from 2007 to 2019. I subset the above survey data to match the selected MMP sample: only males between the ages of 17 and 60 who are heads of household and not self-employed, who reside in one of the municipalities lived in by an individual in the MMP data, who have a positive monthly income, and whose sector is not listed as missing. This leaves me with 5,841,519 observations in the repeated cross-section. I discuss these data in detail in Appendix A.2

Historical Climate

I use historical weather data from two sources to understand the relationship between extreme weather outcomes, migration, and wages: Livneh et al. (2015) (Livneh) and Daymet (Thornton et al., 2022). Both sources provide gridded data based on daily data from weather stations.[9] These data are used to obtain measures of the temperature and total level of precipitation across each municipality in Mexico. I use Livneh for the early years of the panel (1930–1979) and Daymet for the later years (1980–2018). I match the datasets on their 1980 values to ensure that the combined panel is corrected for differences in levels between the two.

I use data on the standardized precipitation-evapotranspiration index (SPEI) to calculate both drought and flooding at the municipality-year level (Vicente-Serrano et al., 2010). SPEI is calculated by applying a normalizing transformation to the difference between precipitation and potential evapotranspiration at a given timescale and point in space. Following the literature (e.g., McKee et al., 1993), drought is defined as occurring when the calculated SPEI is below -1.5. To focus on drought most relevant to agriculture, I define a location as being in a drought when the minimum daily SPEI is below the drought cutoff during the summer agricultural growing season, defined as May through October (Jessoe et al., 2018). Given that the SPEI can be interpreted as the approximate number of standard deviations that the climatic water balance is above or below the mean, I define flooding analogously to drought — that is, a maximum daily SPEI above 1.5.

Climate Projections

I use the output from 25 high-resolution, bias-corrected climate simulation models, provided by NASA Earth Exchange (NEX) (Thrasher et al., 2022). These data provide information on daily projected temperatures and precipitation levels on a 0.25×0.25 degree grid, at a global scale. Even after bias-correcting algorithms are used, individual climate models are generally biased in that their predictions are model-specific (Auffhammer et al., 2013). To partially remedy this issue, I

[9] Livneh is gridded at at 1/16 $^\circ$ resolution ($\sim$ 6km at the equator) and Daymet at 1 km.

perform two additional debiasing exercises. First, I average the daily gridded values over many of the individual models (Burke et al., 2015; Rudik et al., 2022). Second, I match the resulting data with the 2020 values from the historical weather data to correct for biases in levels from the climate model output. I use a moderate warming scenario to represent climate change: the Shared Socioeconomic Pathway 2-4.5 (SSP2-4.5). This climate pathway is one of five primary scenarios used in the Intergovernmental Panel on Climate Change's sixth assessment report and represents a business-as-usual scenario following historical patterns (Riahi et al., 2017). Hsiang and Kopp, 2018 provides a high-level overview of climate change science for economists and discusses the different SSPs.

Daily Measures of Weather

Each of the sources of weather data is produced at the daily level. Given the annual nature of the MMP data, these daily weather outcomes must be aggregated up to annual statistics. To do this, I construct three different measures of extreme temperature outcomes, differentiated by their importance for agricultural productivity and amenity values.

Following the literature (e.g., Feng et al., 2010; Liu et al., 2023), I model the impact of temperature on wages primarily through an agricultural channel, and more specifically, through its impact on maize production based on growing degree days.[10] A growing degree day (GDD) is a piece-wise linear function, with a crop-specific upper and lower discontinuity. In the case of maize, the functional form is

$$GDD(\text{temp}) = \begin{cases} 0 & \text{if temp} < 8 \\ \text{temp} - 8 & \text{if } 8 \leq \text{temp} \leq 32, \\ 24 & \text{if temp} \geq 32. \end{cases}$$

This mapping from daily heat is used throughout the agronomy literature (Herrero and Johnson, 1980; Wilson and Barnett, 1983; Bassetti and Westgate, 1993) and captures the fact that plants are able to absorb additional energy from heat, up until some threshold. I calculate growing degree days at the municipality level as the sum of all daily levels across the summer growing season (May through October).

Growing degree days are specifically designed with crops in mind. To capture amenities over extreme heat, or lack thereof, I also make use of degree days above and below a given threshold x.

[10] Maize is the most sown crop in Mexico (SIAP, 2022) and is grown during the summer months in every state in the country. The impact of weather on wages is modeled as entering through maize yields. This directly impacts the wages of agricultural employees, which spills over into non-agricultural sectors through general equilibrium channels. One of these is through decreases in local demand for non-agricultural products (Liu et al., 2023). As will be made apparent in Section 4, I capture such effects through a richly specified first stage wage equation.

Degree days above x is a recentered and truncated linear function in daily temperature, which is zero until x then increases linearly without bound when the temperature is above x. Similarly, degree days below x is zero when the daily temperature is above x and increases linearly in the absolute difference between x and the daily temperature when the daily temperature is below x.[11] Such statistics are computed using the average daily temperature. To bring these statistics to the annual level, I sum the calculated daily-level degree days over all days in a given year. Mathematically, daily-level degree days above x can be written as:

$$DD_x(\text{temp}) = \begin{cases} 0 & \text{if temp} \leq x \\ \text{temp} - x & \text{if temp} \geq x, \end{cases}$$

and daily-level degree days below x as:

$$DD^x(\text{temp}) = \begin{cases} x - \text{temp} & \text{if temp} \leq x \\ 0 & \text{if temp} \geq x. \end{cases}$$

I use degree days above 26°C, which I label DD26, and degree days below 14°C, which I label DD14. I justify these thresholds in the next section.

Empirical Motivation

Migration is a relatively rare event in one's life. This fact is readily apparent in the MMP data: just over 2.3% of person-year observations include an inter-municipality move from 1967 through 2018. Moreover, since many individuals never decide to move during their lifetime, the gains from such moves are concentrated in a small subsample of the overall population. The MMP data record that just under 25% of the observed population ever moved in their recorded life-history. Panel (a) of Table 1 looks at how the subpopulations of movers and nonmovers differ along a few important dimensions. In particular, migrants, on average, have higher levels of education, are less likely to be employed in agriculture, and are somewhat older when they have their first child.

Panel (b) of Table 1 studies the characteristics of the moves that migrants make. Most moves occur when an individual is relatively young, with the average age of a migrant on their initial move being just over 22 years old, and the average age for all moves being a bit older: 25 years old. Almost 70% of moves are to rural locations. To put this into perspective, 70% of moves originating in a rural location are to another rural location, while 65% of moves originating in an urban location are to a rural location. Among individuals born in urban locations, this proportion is much smaller

[11] Cooling and heating degree days are generally used by the U.S. Energy Information Administration to predict energy requirements for cooling and heating. They use a cutoff of 65°F for these calculations.

Table 1: Descriptive Statistics

	Non-Migrants		Migrants		2020 Census	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Panel (a)						
Education	7.66	4.25	8.76	4.96	9.03	3.63
Agricultural Worker	0.32	0.47	0.22	0.41	0.31	0.46
Children	0.95	0.21	0.93	0.26	-	-
Age at Birth of First Child	23.83	4.90	24.37	4.90	-	-
Panel (b)						
Number of Moves	-	-	2.16	1.84	-	-
Age at First Move	-	-	22.69	6.30	-	-
Move to Urban Location	-	-	0.32	0.46	-	-
Stay > 1 Year	-	-	0.74	0.44	-	-
Panel (c)						
Born in Dry Climate	0.28	-	0.26	-	0.24	-
Born in Temperate Climate	0.41	-	0.48	-	0.59	-
Born in Warm Climate	0.32	-	0.26	-	0.17	-
Born in Urban Location	0.25	-	0.24	-	0.41	-
Number of Individuals	8406		2788			

Note: Education and age are measured in years, agricultural worker is one if an individual works in agriculture at least 25% of their observed life history, and zero otherwise. Children is one if an individual is a parent at the time of the survey, and zero otherwise.

— about 38% for both moves originating in rural and urban locations. Figure A1 plots the empirical density of the number of moves made by migrants. One can readily see the high value associated with the consideration of future migration: the majority of migrants (65%) move more than once, with a large fraction (21%) moving more than twice. Moreover, the average number of moves is just over two, and almost three-quarters of observed moves last over one year, with the average length of stay being 7.4 years (median of 4 years). Given the long duration of most moves, a model that does not allow for dynamic considerations will mistakenly attribute the value of individual moves, across time, to a single year. Moreover, dynamics are necessary to rationalize both the high number of moves made by migrants and the relative infrequency of overall migration.

Panel (c) of Table 1 shows the distribution of birth locations across both the three primary Köppen climate zones in Mexico and urban locations. There is a good amount of variation along these dimensions; however, more importantly, the proportions of births within each of these groups matches somewhat closely to the locations of individuals in the 2020 decennial census. Specifically,

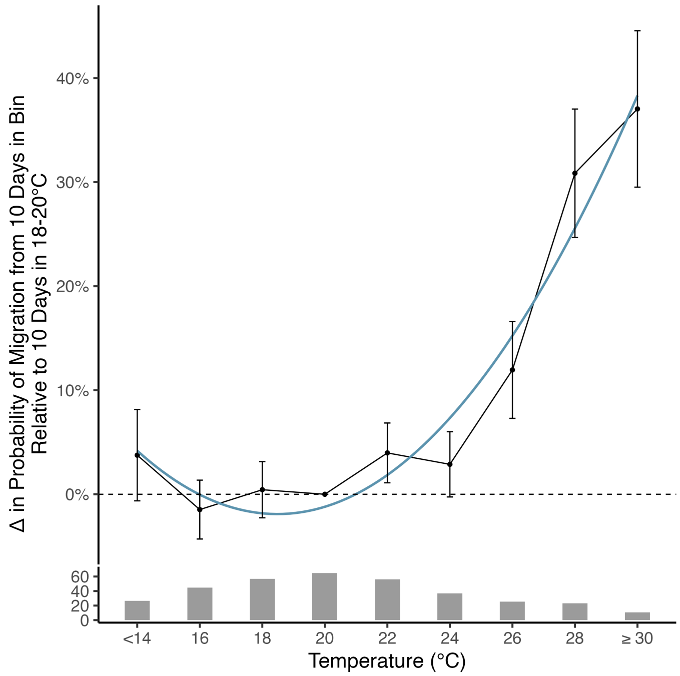


Figure 1: Nonlinear effect of daily temperatures on the probability of migration, conditional on expected wage, age, and person and state-year fixed effects. Linear probability model using MMP data for migrants. The histogram plots the average number of days in each bin, taken across the estimation sample. The 95% confidence bands, corresponding to standard errors clustered at the person and year-by-municipality level, are plotted around the point estimates.

the 2020 census places 24% of the population in dry climates (versus 28%), 59% in temperate climates (versus 42%), 17% in warm climates (versus 30%), and 41% in urban locations (versus 25%).

Figure 1 plots the results from a linear probability model of the decision to migrate on the number of days in the previous year that the average daily temperature falls within each 2° C bin, conditional on age, expected wage, and individual and state-by-year fixed effects. Estimation of the marginal probabilities induced by this model makes comparisons only among migrants, comparing the number of days in each temperature bin in years that they do move to years that they do not, conditional on any macroeconomic trends that occur at the state-year level. The x-axis plots the rightward limit of the bin,[12] with < 14 containing all days below 14° C, and ≥ 30 containing all days with an average temperature greater than 30° C. The plotted effect shows the increase in the average propensity to migrate if I replace 10 days in the 18–20° C bin with 10 days in the given bin. One can see a clear nonlinear relationship between daily temperatures and migration decisions. In particular, days above 26° C are more than twice as important in predicting migration decisions as the very next bin containing days between 24 and 26° C. Although Figure 1 makes clear that

[12] For example, "28" corresponds to the bin between 26 and 28° C.

Table 2: Daily Extreme Weather Increases the Probability of Migration

	Dependent Variable: $\mathbb{1}\{\text{Migrate}\}$			
	(1)	(2)	(3)	(4)
Average Temperature	0.00143***	0.00142***	0.0045***	0.0106***
	(0.00032)	(0.00033)	(0.0013)	(0.0016)
Degree Days above 26°C	0.0129***	0.0130***	0.0059***	0.0149***
	(0.0021)	(0.0021)	(0.0013)	(0.0026)
Degree Days below 14°C	0.00303***	0.00318***	0.0075***	0.0083***
	(0.00055)	(0.00057)	(0.0014)	(0.0017)
Agricultural Worker	-0.0052***			
	(0.0010)			
Education	0.00084***			
	(0.00015)			
Children	-0.0136***			
	(0.0011)			
Num.Obs.	260472	260472	260472	260472
R2	0.052	0.050	0.168	0.201
FE: Year-State	✓	✓		✓
FE: Person			✓	✓

Note: DD26 (DD14) represents 100 annual degree days above (below) 26°C (14°C). Education is measured in years of schooling, agricultural worker and children are indicator variables representing whether an individual is employed in the agricultural sector and has children, respectively. Standard errors are two-way clustered at the person and year-by-municipality level. *p < 0.1; **p < 0.05; ***p < 0.01

heat is a much stronger correlate with the decision to migrate than cool temperatures, there is still a significant impact of cool days below 14° C. For this reason, I focus on degree days above 26° C and below 14° C.

Table 2 presents the results from a series of linear probability models of the decision to migrate with different levels of fixed effects and controls. All four columns show that the primary weather variables of interest — average temperature, degree days above 26°C, and degree days below 14°C — are important correlates with the decision to migrate out of an area. This remains true even when making comparisons only among migrants (column 3), and even when additionally controlling for any local state-by-year-level macroeconomic shocks (column 4). The first column shows that the primary demographic variables used in the model strongly correlate with the decision to migrate. Agricultural workers are 23% less likely to move than those employed outside this sector; parents are 59% less likely to move; and each additional year of education, on average, correlates with a 4% increase in the likelihood of a move.

Although the above analysis shows that individual migration decisions are correlated with the

Table 3: Reduced Form Mechanisms

	Dependent Variable: $\mathbb{1}$\{Migrate\}			
	(1)	(2)	(3)	(4)
Degree Days above 26°C	0.0113***	0.0048***	0.0180***	0.0043**
	(0.0018)	(0.0010)	(0.0024)	(0.0017)
Degree Days below 14°C	0.00189***	0.0028***	-0.000027	-0.0012
	(0.00039)	(0.0011)	(0.001459)	(0.0013)
DD26: Away From Birth Loc	0.0071***	0.017***	0.0123***	
	(0.0025)	(0.004)	(0.0038)	
DD14: Away From Birth Loc	-0.00249***	-0.0024	-0.00029	
	(0.00078)	(0.0024)	(0.00233)	
Num.Obs.	260472	260472	260472	260472
R2	0.064	0.176	0.205	0.341
FE: Year-State	✓		✓	✓
FE: Person		✓	✓	
FE: Person-Location				✓

Note: Degree days 26 (14) represents degree days above (below) 26 (14)° C. "Location" refers to the elements of the choice set, defined in Section 3 of the text. Standard errors are two-way clustered at the person and year-by-municipality level. *p < 0.1; **p < 0.05; ***p < 0.01

heat and cold experienced, it may be the case that the MMP sampling locations happen to have a different climate than the destinations of most moves. To address this possibility, I conduct two empirical robustness exercises and report the results in Table 3. The first three columns test the differential impact of experienced degree days above 26°C and below 14°C on the migration propensity when an individual is and is not currently residing in their location of birth. Column 3 includes both person and state-year fixed effects and is thus the preferred specification. The estimates of the non-interacted degree day effects come from comparisons of the number of degree days when an individual lives in their birth location in the years that they do and do not choose to migrate. I find that a 100 unit increase in degree days above 26°C increases their probability of migrating by 78%. On the other hand, the estimate of the impact of degree days above 26°C when an individual is away from their location of birth compares the number of degree days above 26°C in all years that they are away from their birth location to the number of degree days in the years that they decide to move and are living somewhere other than their birth location. I find that an increase of 100 degree days above 26°C increases the probability of migration by an additional 41%. Most importantly, this exercise shows that individuals are responsive to the heat both when they are living in and away from their location of birth.

The final column of Table 3 studies the impact of degree days above 26°C and degree days below 14°C on the probability of migration, controlling for person-location and state-year fixed effects.

Table 4: Model-Free Evidence of Forward Looking Beliefs over the Weather

	Dependent Variable: $\mathbb{1}$\{Migrate\}		
	(1)	(2)	(3)
Current Temperature	0.0097***	0.0052***	0.0011
	(0.0013)	(0.0014)	(0.0016)
Last Year's Temperature		0.0050***	0.0044***
		(0.0015)	(0.0015)
Next Year's Temperature			0.0052***
			(0.0014)
Num.Obs.	260472	260472	260472
R2	0.200	0.200	0.200
FE: Year-State	✓	✓	✓
FE: Person	✓	✓	✓

Note: Temperature is measured in degrees Celsius. Standard errors are two-way clustered at the individual and year-by-municipality level. *p < 0.1; **p < 0.05; ***p < 0.01

Locations are defined as the elements of the choice set, as given in Section 3, and are areas with similar climates and financial opportunity. The estimation of this specification compiles the average number of degree days each year that an individual living in a particular location experiences and compares this number in years that they do not decide to move to the years when they do decide to move, holding fixed any macroeconomic shocks at the state-year level. I find that individuals are more likely to move in years when there happens to be higher-than-normal levels of degree days above 26°C, where "normal" is defined against the average number of degree days that this individual experienced in that particular location.

The evidence in Table 3 verifies a strong correlation between the weather a person has experienced and the likelihood of moving. I now examine whether individuals are forward-looking with respect to weather and how they form these beliefs about future outcomes. The first column of Table 4 shows that, conditional on person and year-state fixed effects, the current year's municipality-level experienced temperature is a strong correlate of the decision to migrate. When I add the previous year's temperature to the linear probability model in Column 2, the estimated coefficient drops by nearly half, with the current and previous years' temperatures each receiving half of the importance in predicting migration decisions. One explanation for the importance placed on the previous temperature, conditional on the current temperature, is that individuals use this information in predicting next year's temperature, which is what is likely to drive this year's migration decisions. To provide evidence of this mechanism, in the third column, I add next year's temperature to the model. One can now see that both the future and previous year's temperature, conditional on the current temperature, are the important drivers behind the decision to migrate.

Table 5: Static Logit Parameters

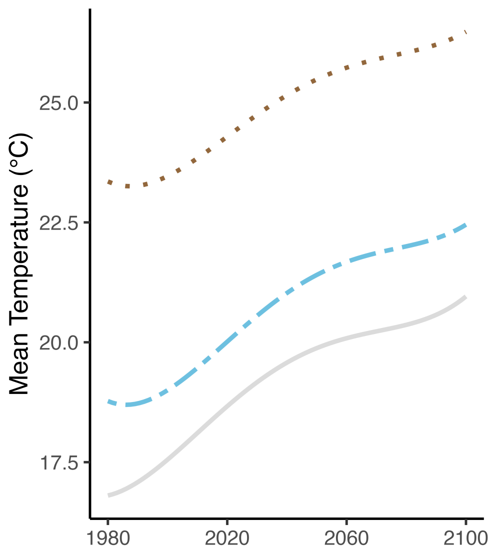
Flow Utility		Moving Costs	
Income	0.059	Moving Intercept	2.891
	(0.006)		(0.073)
Degree Days above 26°C	0.067	Distance	0.139
	(0.058)		(0.011)
Degree Days above 26°C^2	-0.074	Children	0.129
	(0.021)		(0.042)
Degree Days below 14°C	0.134	Age	0.209
	(0.011)		(0.009)
Degree Days below 14°C^2	-0.022		
	(0.002)		
Urban Location	-0.342		
	(0.050)		
Birth Location	2.452		
	(0.024)		
Log-Likelihood: -25255.83			

Note: Income is measured in annual hourly 2010 pesos, degree days are divided by 100, distance is measured in log kilometers. Standard errors are computed with the inverse Hessian.

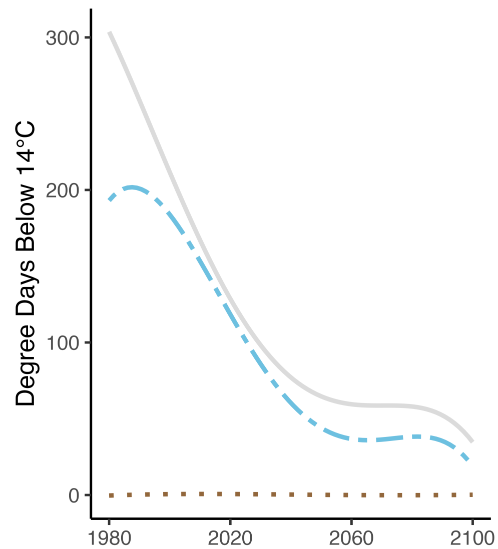
I now explore whether individuals, when deciding both whether and where to migrate, take into account not just the weather in their current location, but also that of their potential destinations. To do so, I estimate a static discrete choice model of migration within Mexico, which is similar to the dynamic model discussed in the following section. Specifically, I use a conditional logit framework to estimate the preference parameters corresponding to the utility of an individual living in location ℓ and then moving to location ℓ':

$$\begin{aligned} u(\ell,\ell') = {} & \alpha_1 \text{income} + \alpha_2 \mathbb{1}\{\ell' \in \mathcal{U}\} + \alpha_3 \mathbb{1}\{\ell' = \nu^\ell\} \\ & + \alpha_4 DD26 + \alpha_5 DD26^2 + \alpha_6 DD14 + \alpha_7 DD14^2 \\ & - \mathbb{1}\{\ell' \neq \ell\} \times [\alpha_8 + \alpha_9 d(\ell,\ell') + \alpha_{10} \mathbb{1}\{\text{children}\} + \alpha_{11} \text{age}]. \end{aligned} \tag{1}$$

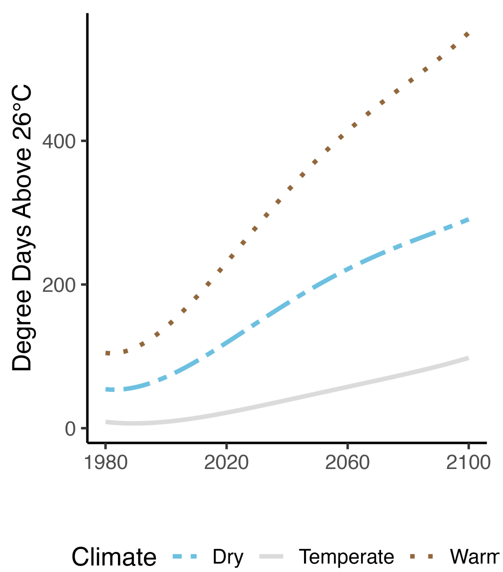
Here, $DD26$ represents last year's number of degree days above 26°C; $DD14$ last year's number of degree days below 14°C; $\mathcal{U}$ is the set of urban locations; ν^ℓ is an individual's birth location; and d is the Euclidean distance metric. Table 5 presents the estimated parameters, all of which are as expected. Individuals view higher incomes and living in the same area where they were born as positive attributes in their location of residence. Moving is costly, and this cost increases



(a)



(b)



(c)

Figure 2: Climate transitions in Mexico under the Coupled Model Intercomparison Project Phase 6 (CMIP6) Shared Socioeconomic Pathway 2-4.5 (SSP2-4.5) through 2100 and smoothed using a fourth-order polynomial in the year

with distance, parenthood, and age. Moreover, the estimated parameters on the weather show that individuals enjoy moderate amounts of both heat and cold, but dislike extreme levels. To put this finding in perspective, note that the change in utility from a change in the number of degree days above 26°C is:

$$\frac{\partial u}{\partial DD26} = \alpha_4 + 2\alpha_5 DD26.$$

This quantity is positive when $DD26 < -\frac{\alpha_4}{2\alpha_5} = 0.48$, or when a location experiences fewer than 48 degree days above 26°C. The marginal impact of an additional degree day above 26°C below the 68th percentile of the observed data has a positive impact on welfare. This relation is even stronger for degree days below 14°C: an additional degree day below the 77th percentile of the observed data has a positive impact on welfare.

To provide further evidence of a forward-looking dynamic in expectations about future weather on the decision to migrate, I modify the above static logit framework to additionally allow for preferences about future temperature. I include an additional preference parameter for the average of the daily temperature distribution over the next decade in the location chosen today. I present the estimated parameters in Appendix D.2, Table D2. The results show a strongly significant preference for future expected temperatures, suggesting that individuals form forward-looking expectations about the climate when deciding where to live.

Given that heat and cold are enjoyed in moderation but disliked in abundance, it is important to understand how the degree days of interest are projected to evolve in the future. Figure 2 plots changes in the relevant weather outcomes across Mexico's three primary Köppen climate zones and time under a projected moderate warming scenario.[13] Of particular note is the smooth increase in the average temperature across time, albeit at a reduced rate in the climate projections than under the historical observation. This increase in annual average temperatures maps strongly to increases in tail-events. Figure 2c highlights that an increase of less than 2°C in average temperatures from 2020 through 2100 maps to a more than doubling in the number of annual degree days above 26°C. Similarly, Figure 2b shows that this increase maps to a sharp decrease in the number of degree days below 14°C in the dry and temperate climates in Mexico. The expected increase in the number of degree days above 26°C and decrease in the number of degree days below 14°C work together to show the potential welfare losses via amenities from climate change.

This section has documented a correlation between experienced, future, and past temperatures with migration decisions. In the next section, I present the model framework, which allows me to study the role of dynamics in shaping the value of migration in response to climate change.

[13] Technically, there are four first-level climates in Mexico, with a handful of areas considered to have a polar climate. As there are no municipalities in Mexico with a majority of land area contained in a polar climate, these areas are subsumed into nearby climates.

3 Model

Overview

Individuals are modeled as making a sequence of 15 decisions about where to live. Each time they make this choice, they take into account that the next time they will have the opportunity to move, they will be a bit older, and that they will be making that choice from an origin of today's choice. This decision of where to live takes place once every three years, beginning at age 17 and ending at 61.[14] These decisions are drawn from a choice set containing 27 distinct options within Mexico, where each option is defined as a cluster of municipalities with similar climates and employment opportunities

Individuals are modeled as varying in unobserved ways across two dimensions: migration costs and expectations about the climate. First, some individuals never truly consider the option of moving for a variety of reasons, including deep-rooted familial ties to their birth location, a strong fear of the unknown, or any other highly heterogeneous reason that makes a person absolutely desire to remain in the same location for their life. Other individuals actively consider the trade-offs of living in a different location—be it higher incomes or improved amenities and the like—versus the monetary and psychological costs of moving.

Second, individuals form beliefs about the climate in each of their potential choices in different unobservable ways. In the spirit of the literature studying the sophisticated ways that consumers form long-term expectations (e.g., Busse et al., 2013), I allow for individuals to form rich expectations about a spatially heterogeneous climate that is changing heterogeneously throughout time. Other work in economics has criticized the assumption that individuals are able to form correct expectations over complicated dynamics (e.g., Simon, 1955). For this reason, I additionally allow individuals to form naive expectations about the future climate. I assume that these individuals see the weather in the current period and predict, with certainty, that all future weather will be the same. I use the model framework to estimate the unobserved portion of the population associated with each of the belief processes and moving types, at the same time as the other model primitives.

After observing the current year's weather and forming expectations of the implied future climate, but prior to making the decision of where to move during this period, an idiosyncratic, transitory, and location-specific preference shock affecting a combination of incomes and moving costs is realized. This individual then compiles the available information and chooses to live in the location that maximizes their lifetime welfare, as defined against their expectation process. If they decide to move to a new location, they then pay an upfront cost that depends on their demographic

[14]The vast majority of observed moves occur well before an individual is 61 years old — only a single individual moves (one time) after this age in the MMP data used for estimation.

characteristics as well as the characteristics of their origin and destination locations. After making this decision, agricultural season drought and flooding shocks affecting their income are realized.[15] They then obtain three years' worth of discounted flow utilities associated with their destination, and they begin the next period with an updated set of state variables in their current location.

A Dynamic Model of Where to Live

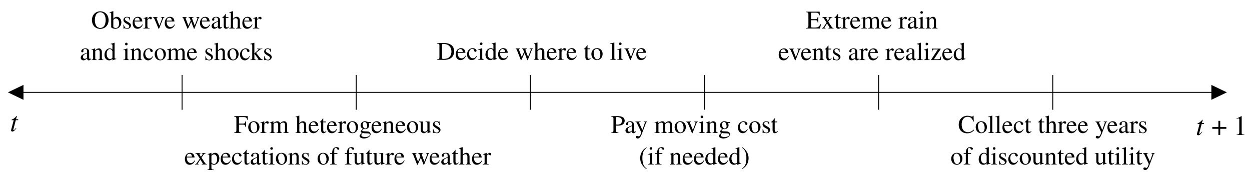


Figure 3: Within Period Timing

More formally, take some individual i in some time period $t \in \{1, \ldots, 15\}$. At the start of period t, i begins in their previous-period choice location, ℓ, and observes their vector of state variables, ω, along with a vector of location-specific utility shocks, ε. Given this framework, an individual receives a flow utility associated with the decision to move to any particular location, ℓ', as given by $u(\ell, \ell', \omega, r) + \varepsilon_{\ell'}$, where r is a two-dimensional vector containing the drought and flooding realizations. Individual i is modeled as sequentially choosing locations ℓ'_t, consistent with maximizing the following discounted sum of future flow utilities, given their time t decision[16]:

$$\max_{k \in \mathscr{C}^{|T-t|}} \left\{ \sum_{s=0}^{T-t} \delta^{3s} \mathbb{E}\left[u_s(k_s, k_{s+1}, \omega_s, r; \theta, \tau) + \varepsilon_{k,s} | \omega, \tau \right] \right\}.$$

When making this decision, i is assumed to be fully aware of their current state, as given by ω_τ, and the current-period vector of idiosyncratic shocks. However, future realizations of these values are unknown from a time t perspective, as are the within-period rainfall shocks. Therefore, i uses their knowledge of the distribution over which these variables are realized when deciding where to live today. Mild regularity conditions on this joint distribution, along with an assumption of the separability of the deterministic flow utility values from the stochastic utility shocks, allow for a translation from the above sequential representation of the modeled decision to that of a Bellman

[15] I model extreme rainfall events as occurring after the moving decision for two primary reasons. First, doing so better accounts for higher levels of uncertainty surrounding localized precipitation than localized temperatures. Second, conditional on the realized weather states, it is assumed to be uncorrelated across time. (This is a computational requirement in that these events are therefore not elements of the state space).

[16] I calibrate the annual discount rate, δ, to 0.95.

equation. Specifically, I can reformulate the time $t < T$ decision as:[17]

$$V_t(\ell, \omega; \theta, \tau) = \max_k \left\{ \mathbb{E}_R[u(\ell, k, \omega, r; \theta, \tau)] + \varepsilon_{tk} + \delta^3 \mathbb{E}_{\Omega,\varepsilon|\tau}[V_{t+1}(k, \omega'; \theta, \tau)|\omega] \right\}.$$

The time T value function, representing the sum of expected current flow utility and a discounted retirement value can be written as:

$$V_T(\ell, \omega; \theta, \tau) = \max_k \left\{ \mathbb{E}_R[u(\ell, k, \omega, r; \theta, \tau)] + \varepsilon_{tk} + \frac{\delta^3}{1-\delta} \mathbb{E}_{R,\Omega|\tau}[u(k, k, \omega', r'; \theta, \tau)|\omega] \right\},$$

where the final term of the above equation represents a retirement value associated with living indefinitely in location k, without the ability of further moves. Assuming that the state no longer evolves after an individual's final decision allows me to use a recursive strategy to fully solve the model, as in Gowrisankaran et al. (2024).

Formulating the individual decision in terms of a Bellman equation highlights the dynamic interplays present in such a model. Of particular note, each period an individual faces a high number of temporal trade-offs between current and future utilities, including an individual's awareness of future shocks which may unexpectedly change the future flow utilities associated with any decision made today. These shocks enter through both the idiosyncratic draws and future weather realizations throughout Mexico.

Components of Flow Utilities

Flow utilities associated with living in any particular location ℓ include three components: income, temperature as an amenity, and level shifters from both living in one's location of birth and living in an urban location. Specifically, flow utility of an individual deciding to live in ℓ' is modeled as:

$$\begin{aligned} \bar{u}(\ell', \omega, r; \theta) = {} & \theta_1 \text{inc}(\ell', \omega, r) \\ & + \theta_2 DD26(\ell', \omega) + \theta_3 DD26(\ell', \omega)^2 + \theta_4 DD14(\ell', \omega) + \theta_5 DD14(\ell', \omega)^2 \\ & + \theta_6 \mathbb{1}\{\ell' = \nu^\ell\} + \theta_7 \mathbb{1}\{\ell' \in \mathcal{U}\}. \end{aligned} \tag{2}$$

This parameterization is given for a single year. However, since each model period is three years long, an individual's period flow utility is instead given by the present discounted sum of the annual flow utility over three years. Specifically, in each period, an individual receives a flow payoff of $\sum_{s=0}^{2} \delta^s \bar{u}(\ell', \omega)$.

[17] $\mathbb{E}_R$ is an expectation about extreme rainfall events, while $\mathbb{E}_{\Omega,\varepsilon}$ is an expectation about the joint state variable and utility shock distribution. $\mathbb{E}$, with no subscript, is an expectation about the joint distribution of all three sources of stochasticity.

Income

Income is parameterized as a function of one's location, individual-specific state variables, and the weather, specifically, a flexible function of growing degree days, drought, and flooding.[18]

Specifically,

$$\text{inc}(\ell, \omega, r) = f(\ell, \omega, r),$$

where f is assumed to be a known mapping to be estimated, and r contains indicators for drought and inundation. Within-period uncertainty over wages comes from uncertainty about extreme rainfall outcomes, and future uncertainty comes from the transitory shocks, weather shocks, and extreme rainfall outcomes.

Temperature as an Amenity

Weather is allowed to affect flow utility both indirectly through wages and directly as an amenity. Amenity values are flexibly modeled as entering through extreme degree days, that is, through degree days below 14° C and degree days above 26° C. These values were chosen for two primary reasons. First, almost all the locations in the choice set have historically experienced nonzero levels of these statistics.[19] Second, as shown in Figure 1, at these thresholds, a nonlinearity in the reduced-form probability of migration are present. The specification presented in Equation 2 remains agnostic towards the sign and magnitude of the linear and quadratic terms for both extreme heat and cold. In particular, it allows for a bliss-level of both where utility is increasing until this point and then decreasing after it.

Flow Utility Shifters

To appropriately account for the high proportion of the population who either never leaves their location of birth or who leaves and subsequently returns, I allow for a level shifter of period utility from living in an individual's location of birth (Kennan and Walker, 2011). Further, to account for differences in the populations living in urban and rural locations that are not explainable by income differentials, I also allow for a utility shifter from living in an urban location.

Moving Costs

Given the relative infrequency of moves seen in data, it is important to accurately capture the costs associated with a given move. I model moving costs as depending on both observed and unobserved

[18] This parameterization is discussed in Section 4.

[19] 22 of 27 locations have experienced nonzero degree days above 26°C, while 26 have experienced nonzero degree days below 14°C.

sources of individual, origin, and destination characteristics. This cost is paid upfront in the first year of a model period that an individual decided to move. I parameterize it as follows:

$$\bar{c}(\ell, \ell', \omega; \theta, \tau) = \gamma_1(\tau) + \gamma_2 d(\ell, \ell') + \gamma_3 \mathbb{1}\{k_\omega \geq 1\} + \gamma_4 \text{age}_\omega, \tag{3}$$

where one source of unobserved heterogeneity, fixed within individual, enters through $\gamma_1(\tau)$. d represents the Euclidean distance metric, k_ω the number of children, and age_ω the current period time-step.

I allow for two unobserved permanent moving-cost types of individuals. The first type comprises individuals who consider the trade-offs associated with moving each period. These individuals are used to estimate the magnitude of γ_1. The other type is composed of individuals who do not actively consider the option to move in any period; their γ_1 is prohibitively high. Allowing for this source of unobserved heterogeneity is standard in the literature (Kennan and Walker, 2011; Lessem, 2018; Oswald, 2019), particularly, because past work has generally found that large masses of the population have very high psychological costs associated with moving (Koşar et al., 2022).

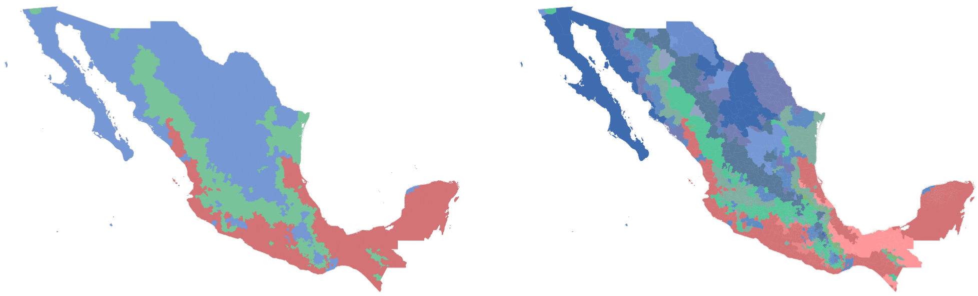
Observed individual heterogeneity enters through level-shifters in the cost function. Specifically, the moving cost can differ for individuals as they progress through their lifetime by allowing this cost to depend on their age. The moving cost also is allowed to be different once an individual has their first child. Heterogeneity in costs across space are modeled through different moving costs based on the distance from an individual's current location to their destination.

A Model of Heterogenous Climates in Mexico

The climate is generally defined as a high-dimensional distribution linking physical systems, such as oceans and the atmosphere (e.g., Hsiang and Kopp, 2018). Weather, itself a high-dimensional object, is just a realization of this distribution at a particular time and place. Computational feasibility requires that I take a summary statistic approach to modeling climate dynamics. I model the mean of the daily temperature distribution and the level of agricultural-season precipitation as following a multivariate Gaussian distribution, with the mean and variance estimated using data from the previous 30 years, from the perspective of year y. That is,

$$\begin{pmatrix} \text{temp} \\ \text{prec} \end{pmatrix} \sim \mathcal{N}(\mu^y, \Sigma^y), \tag{4}$$

where temp and prec are both three-dimensional vectors containing the relevant realizations across the three level-one Köppen climates in Mexico. Figure 4a plots the level-one Köppen climate zones in Mexico, with blue representing the "dry," teal the "temperate," and red the "warm" climates.



(a) Level-One Köppen Climate Zones (b) Level-Two Köppen Climate Zones

Figure 4: Köppen Climates in Mexico

As discussed in Section 4, individuals, in part, choose among locations defined against level-two climates, which are nested within the level-one climates. Figure 4b plots these throughout space. Appendix A.3 provides a detailed discussion of this classification system and how it is used in this paper.

The model allows for two different belief formation processes, and as discussed in Section 4, the mass of the population associated with each is estimated along with the other model primitives. The first process is one of naivety. Individuals aligned with this framework observe the weather each period and assume that all future weather will be drawn from a degenerate distribution at this point. In particular, I only assume that these individuals are aware of the realized weather in the current period for all locations in their choice set.

The second process is one of fully-informed expectations. More specifically, individuals who form expectations under this process understand the functional form of μ^y and Σ^y, as given in 4, where

$$\mu_w^y = \alpha_w^y + \alpha_{w1}^y \mu_w^{y-1} + \alpha_{w2}^y y,$$

and w represents the level-one average temperature or agricultural-season precipitation. I model fully-informed individuals as forming their understanding of α^y and Σ^y based on realized weather over the previous 30 years, as is generally accepted as the length of a "climate."

Although computational tractability requires that I only track summary statistics of the climate system, past work has found that local weather impacts outcomes in a nonlinear fashion (Schlenker and Roberts, 2009; Aragón et al., 2021; Miller et al., 2021). To appropriately account for this fact in the dynamic model, conceptually, I allow for heterogeneous climate distributions for each element of the choice set, where the particular climate distribution is determined by the relevant level-one weather realizations.[20] I assume that, on average, the individuals who form fully-informed expectations about the climate system are aware of these distributions and that they use the relevant realized level-one climate temperature to pin down the daily temperature distribution for a particular location. They use the expected daily distribution to integrate out any necessary degree day calculation. Operationally, I assume that they use location-specific mappings from the level-one climate mean temperature into the expected number of degree days in a location. Mathematically, they understand that the expected number of a particular type of degree day (DD) is given by

$$DD_\ell = \mathbb{1}\{\lambda_\ell^D + \lambda_{1\ell}^D \text{temp}_{C(\ell)} > 0\} \times \left[\lambda_\ell^D + \lambda_{1\ell}^D \text{temp}_{C(\ell)}\right], \tag{5}$$

where $C(\ell)$ represents the level-one climate for location ℓ, and temp is the realized mean of the location's daily temperature distribution.

Precipitation levels are generally of first-order concern in agriculture (Deschênes and Greenstone, 2007; Seo and Mendelsohn, 2008; Dell et al., 2014); however, their impact on outcomes has proven elusive to capture empirically. This difficulty is especially true when precipitation is aggregated to large spatial areas. To account for this issue, I instead model that individuals use a mapping from level-one climate agricultural season precipitation levels to location-specific probabilities of drought and flooding. Specifically, they form expectations based on the following:

$$\text{prob}(R_\ell) = \Phi\left(\eta_\ell^R + \eta_1^R \text{P}_{C(\ell)} + \eta_2^R \text{temp}_{C(\ell)}\right), \tag{6}$$

where R is either drought or flooding, and again, $C(\ell)$ represents the level-one climate for location ℓ, P the agricultural-season precipitation level, and temp the mean of the daily temperature distribution. Appendix B.1 discusses these mappings and their estimation in greater detail.

Figure 5 provides a graphical representation of the climate model used by the individuals forming fully-informed expectations. Starting at the top, these individuals understand that the weather is drawn from a Gaussian distribution characterized by the mean and variance described earlier in this section. A draw from this distribution gives the mean daily temperature and level of agricultural-season precipitation for each of the three primary Köppen climates: dry, temperate, and warm. These individuals then use these values to form expectations about the number of degree days and the probability of drought and flooding in each location. They form these expectations based on Equations 5 and 6, respectively. These equations are estimated in the first stage, which allows for a rich set of different weather variables to influence migration decisions. In particular, I use six different weather variables to characterize an individual's state. After combining these variables

[20]For example, both "dry warm dry" and "dry warm temperate" are different level-two climates (and therefore define different locations in the choice set) that share a level-one climate of "dry".

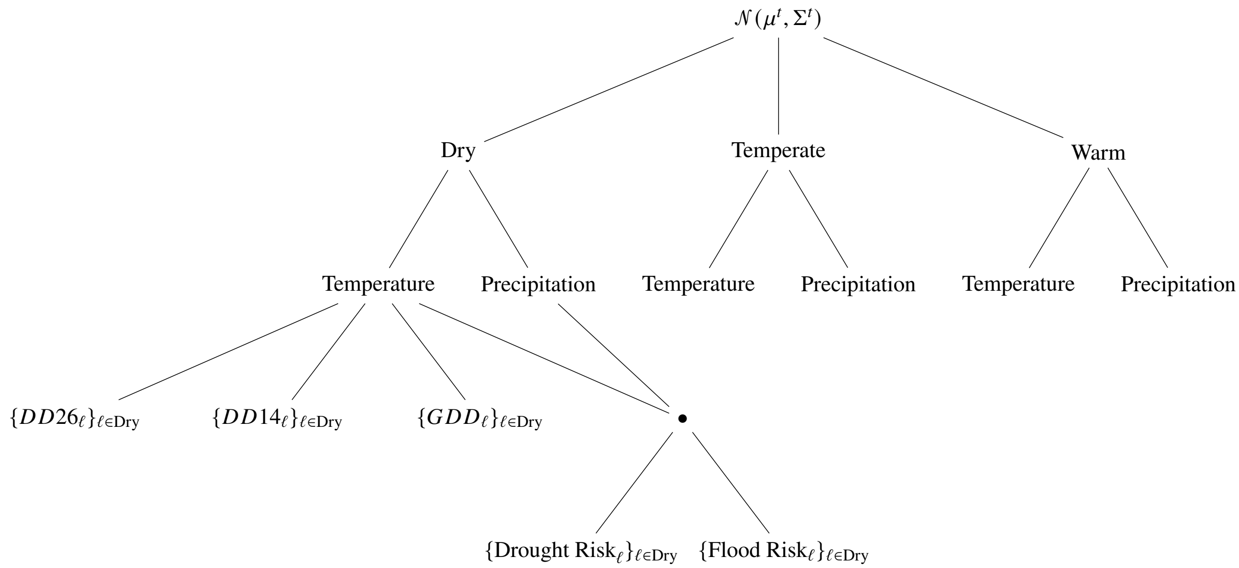


Figure 5: Graphical Representation of the Climate Model

with the estimated mappings, I am able to use variation from 141 different weather variables in the dynamic model.

4 Identification & Estimation

Identification

I briefly discuss the intuition behind the identifying variation for each of the model primitives; a more formal argument is presented in Appendix B.3. Simply put, the income and temperature flow utility parameters are identified from two sources of variation in the life history data. Conditional on making a move, the first source of identifying variation compares differences between the given quality in the observed choice location and that in all other options. For example, part of the income parameter comes directly from comparisons of the expected income in an individual's choice location to that across the menu of choices an individual faced in the year of their move. The second source of variation identifying these parameters instead comes from comparisons in average rates of migration across differences in these qualities, conditional on all other state variables. More specifically, this source of identification holds constant all state variables (and other preference parameters) except that which one is seeking to identify and compares how changes in this variable influence the expected migration rate within each state variable grouping.

The variation identifying the preference for living in one's home location also comes from two facets of the data. The first of these compares the aggregate migration rate away from an individual's home location to the migration rate away from all other locations. The second compares the duration of stay when an individual is and is not in their birth location. Similar comparisons are made when identifying the preference for living in urban locations.

Identification of the mixing probability for the different moving types comes from a comparison of the migration rate for individuals on their second or later move to those who would be making their first move. Intuitively, the first-time migration rate contains a mixture of information about both moving types, while the migration rate for later moves only contains information about the movers. The estimation routine exploits this difference, conditional on all other state variables, to back out the mass of the population associated with either type. The moving cost intercept is identified from the migration rate across the population, conditional on all state variables and other parameters, most importantly, on the mixing probabilities. Thus, the intercept's value is identified from comparisons of migration rates within first and later moves. The distance parameter is identified from comparisons of the distance traveled for each move in the data, the different cost borne by parents comes from migration rate comparisons across individuals with and without children, and the age parameter comes from migration rate comparisons across the age distribution.

Finally, one part of the differential mass of each climate belief type is identified from the variation, caused by climate change, in the future relative values of living in the different locations in an individual's choice set. To build intuition, imagine that an individual can only choose between two locations, a and b. Further, suppose that today a is preferable to b, but that a's future desirability is decreasing at a faster rate than that of b. A fully-informed individual would choose b with a higher probability than a naive individual. Another source of identifying variation of the belief masses comes instead from comparisons of the distribution of migration decisions over time. In particular, as presented in Figure 2, the rate of warming across the three climate zones in Mexico has varied quite strongly throughout the sample period. Individuals deciding where to live in the early years of the sample, therefore, faced a climate which was changing at a much slower rate than the climate faced by individuals in the latter years of the sample. The differences in the rate of warming across time cause otherwise identical individuals to have different choice probabilities, but only if they are fully-informed on the climate system. The estimation framework, discussed later in this section, makes comparisons such as the two discussed here, across the model framework and the migration decisions observed in the MMP data, to identify the mass associated with either belief process.

As a final point of emphasis, much of the model's ability to identify point estimates associated with the preference parameters rests on normalizations imposed on the stochastic shocks. In particular, I normalize the scale of this distribution to one, which allows for separate identification for each preference parameter.[21]

[21] An alternative strategy would be to normalize the marginal flow utility of income to one and instead estimate the variance of the stochastic shocks. This identification would follow from comparisons of migration outcomes across

Choice Set

The choice set is defined as the product of the level-two Köppen climate zones in locations where individuals from the MMP data were seen to reside and an indicator for urban status, all defined at the municipality level. Figure 4 plots the location of each level- one and -two Köppen climate group, with the level-one climates shaded salmon for the "dry" climates, green for the "temperate" climates, and blue for the "warm" climates.

The second element of the choice set is the classification as either rural or urban. I use the Mexican government's metro zone classification to make this distinction. Specifically, a municipality is considered "urban" if there is a metro zone within its boundaries, and "rural" otherwise. Each element of the choice set is therefore a cluster of municipalities, defined such that each contains all municipalities of the same level-two climate grouping and "urban" status.[22]

Transition Probabilities

I estimate the state variable transition distribution in a first stage. The estimated distribution is used in the second stage to integrate over the value of future states within each of the available choices. For numerical tractability, I assume that occupation (in or out of agriculture), education level, and one's birth location are constant within individual. Appendix B provides evidence in support of these restrictions. Except for one's moving-belief type, the remaining state variables transition across an individual's lifetime.

Given the triennial structure imposed on the model timing, an individual's age increases by 3 years each time step with probability one. If an individual has had a child before or during the current period, then they do not transition out of the parenthood state; however, if they have not had a child, then they may become a parent with probability $p_c(a)$. This probability depends on a second-order polynomial of their age and is estimated using a probit regression of the MMP data in the first stage.

I compress each year into its decade, which transitions to the next element with probability 0.3. The decade is used to track the relevant climate distribution, as described in Section 3. The model's solution procedure, discussed later in this section, requires that the model be solved through the year 2064. For all years beyond 2019, the parameters are estimated using data from the CMIP6 business-as-usual climate scenario. Estimates of all first-stage transitions are provided in Appendix B.2.

individuals with identical states who make different decisions.

[22] One climate group has no metro zones within any municipality in its borders; therefore, there are 27 different locations in the choice set.

Wages

Expected wages are estimated in a first step, using external census survey data, as discussed in Section 2. I assume that expected wages are of the following form:

$$
\begin{aligned}
\mathrm{inc}(\ell, \omega, r) = \sum_a \sum_e \beta_{a,e} + \beta_{\mathrm{ag}} ag(\omega) \\
+ \sum_{c=1}^{3} \left[\beta_c^{GDD,C} GDD(\omega) + \beta_c^{GDD2,C} GDD(\omega)^2 + \beta_c^{D,C} D(r) + \beta_c^{I,C} I(r)\right] \\
+ \sum_{ag=0}^{1} \left[\beta_{ag}^{GDD,Ag} GDD(\omega) + \beta_{ag}^{GDD2,Ag} GDD(\omega)^2 + \beta_{ag}^{D,Ag} D(r) + \beta_{ag}^{I,Ag} I(r)\right] \\
+ \xi_\ell + \eta_y(\omega), \qquad (7)
\end{aligned}
$$

where a indexes age and e education level, and ag is an indicator for whether an individual is employed in agriculture. This functional form offers a large degree of flexibility, while still allowing for a structural interpretation of the resulting parameter estimates. Each location has its own intercept, ξ_ℓ. The age-education intercepts allow for a flexible return to experience and education, which, given the year fixed effects $\eta_y(\omega)$, accounts for potential biases arising from differences across birth-cohorts. The remaining terms represent the differential impact across level-one climates and agricultural employment of different measures of weather on wages. $GDD(\omega)$, representing agricultural-season growing degree days, enter into wages quadratically. D is an indicator for drought and I for inundation (flooding), both measured during the agricultural growing season. Appendix A.2 provides details on the estimated coefficients from Equation 7.

Estimation Framework

I estimate the model primitives using a nested full-solution routine within a pseudo-maximum likelihood framework (Rust, 1987). I briefly discuss the format of this procedure here, and provide more details in Appendix B. The likelihood is a mixture over the unobserved moving-belief types. Specifically, for a given vector of the mass of unobserved types π_τ and a vector of structural parameters θ, the log-likelihood is the sum of the individual log-likelihoods, where each individual likelihood integrates over the distribution of unobserved heterogeneity: the moving and belief types, represented by τ. Mathematically,

$$\begin{aligned}\Lambda(\theta, \pi) &= \sum_i \log\left(\mathbb{E}_\tau[\mathscr{L}_i(\theta; \tau)]\right) \\ &= \sum_i \log\left(\sum_\tau \pi_\tau \cdot \mathscr{L}_i(\theta; \tau)\right). \end{aligned} \tag{8}$$

Given θ and π_τ, Λ can be computed directly from the data and interpreted as being the log-likelihood of observing the choices made in the MMP data, given the model structure. The assumption that ε_ℓ follows a type-one extreme value distribution provides a closed form of $\mathscr{L}_i(.)$. To see this, I can first define the alternative-specific value function ($\overline{V}_t(\ell, k, \omega; \theta, \tau)$) as the value afforded to an individual in state ω who currently lives in ℓ from choosing option k, net of the current period stochastic shock. Mathematically,

$$\overline{V}_t(\ell, k, \omega; \theta, \tau) = \mathbb{E}_R[u(\ell, k, \omega, r; \theta, \tau)] + \delta^3 \mathbb{E}_{\Omega,\varepsilon|\tau}[V_{t+1}(k, \omega'; \theta, \tau)|\omega]. \tag{9}$$

Conditional on the structural parameters and the unobserved type, the individual's choice problem can then be rewritten as

$$\max_k \left\{\overline{V}_t(\ell, k, \omega; \theta, \tau) + \varepsilon_k\right\}.$$

Following McFadden (1973), rewriting the choice problem in this way allows one to write the model-generated choice probabilities as

$$\mathscr{L}_{it}(\theta; \tau) = \frac{\exp(\overline{V}_t(\ell_{it}, \ell_{it+1}, \omega; \theta, \tau))}{\sum_{j \in \mathscr{C}} \exp(\overline{V}_t(\ell_{it}, j, \omega; \theta, \tau))}.$$

This formulation of the choice probabilities makes clear that one need only obtain the set of alternative-specific value functions at each time step and location in the state space.

The distributional assumption on the preference shocks additionally assists with this endeavor. In particular, Equation 9 can be rewritten as

$$\begin{aligned}\overline{V}_t(\ell, k, \omega; \theta, \tau) &= \mathbb{E}_R[u(\ell, k, \omega, r; \theta, \tau)] + \delta^3 \mathbb{E}_{\Omega,\varepsilon|\tau}[V_{t+1}(k, \omega'; \theta, \tau)|\omega] \\ &= \mathbb{E}_R[u(\ell, k, \omega, r; \theta, \tau)] + \delta^3 \int \max_j \{\overline{V}_{t+1}(k, j, \omega'; \theta, \tau) + \varepsilon_j\} dF_{\omega', \varepsilon|\omega,\tau} \\ &= \mathbb{E}_R[u(\ell, k, \omega, r; \theta, \tau)] + \delta^3 \int \log\left(\sum_{j \in \mathscr{C}} \exp(\overline{V}_{t+1}(k, j, \omega'; \theta, \tau))\right) dF_{\omega'|\omega,\tau},\end{aligned}$$

where the second line uses the definition of the alternative-specific value function to replace the value of beginning the next period in location k with state ω' with the maximum alternative-specific value, inclusive of each preference shock, across the various choices available.[23] The third line exploits the iid nature of the type-one extreme value preference draws to break apart the joint expectation into a sequential expectation over the state variable transitions and the preference shocks. It then uses the properties of the type-one extreme value distribution to rewrite the expectation of this maximum in a closed form.[24]

To obtain the set of alternative-specific value functions given a vector of preference values, I employ a backward-recursion strategy (Wolpin, 1984). Briefly, given some vector of preference parameters, one can start at the terminal period and calculate the set of alternative-specific values associated with making any decision from the perspective of an individual at every spot in the state space. The time T values can then be used to calculate the time T-1 alternative-specific value functions. One can work backward in this fashion until the full set of alternative-specific value functions has been recovered. With these in hand, one can directly calculate Equation 8 and maximize its value over the structural parameters and mixing probabilities to obtain consistent estimates of their values (Rust, 1988).

Substitution Patterns

As a minor point of emphasis, this framework additionally relaxes the independence of irrelevant alternatives (IIA) assumption that is ubiquitous in much of the spatial literature to allow for a more accurate representation of the substitution patterns that a potential migrant in Mexico faces when they make the decision of where and whether to move (Chipman, 1960; de Dios Ortuzar, 1983; Brownstone and Train, 1998). To see why making this change is important in the context of migration and climate change, I provide the following stylized example. Suppose that individuals choose from three locations in the country: dry, temperate, and warm. Further, suppose that in relative terms, the temperate location is becoming more desirable with time, while the dry and warm locations are becoming less desirable with time. Moreover, suppose that the desirability of the warm location is declining faster than that of the dry location. Notice that the increase in the value of living in the temperate location would cause a higher proportion of the population to decide to live there. IIA substitution patterns would require that this increase is proportionately driven by the warm and dry locations. In reality, we may expect that relatively more people would move to the temperate location from the warm location than the dry location. The dynamic framework that I

[23] Conditional on one's unobserved type, the integral on the second line is thirty-five dimensional. This comes from twenty-seven location shocks, and the state transition for: one's parental status, the year, and the six weather state variables. I don't include one's birth location, their education, and their sector of employment in this calculation, as computational tractability requires that those are assumed constant within individual.

[24] The resulting integral is now eight dimensional.

employ allows for such substitution patterns to occur.

5 Results

Model Estimates

Table 6: Structural Parameter Estimates

Description	Parameter	Coefficient	Standard Error
Flow utility			
Income	θ_1	0.006	(0.0007)
100 Degree days above 26°C	θ_2	0.034	(0.005)
100 Degree days above 26°C squared	θ_3	-0.015	(0.002)
100 Degree days below 14°C	θ_4	0.020	(0.002)
100 Degree days below 14°C squared	θ_5	-0.003	(0.0003)
Living in an urban location	θ_6	-0.041	(0.003)
Living in location of birth	θ_7	0.183	(0.002)
Moving costs			
Fixed migration cost, for movers	γ_1	2.599	(0.073)
Distance of move	γ_2	0.140	(0.011)
Cost shifter: parenthood	γ_3	0.072	(0.022)
Cost shifter: age	γ_4	0.172	(0.006)
Unobserved heterogeneity			
Mass of movers	π_m	0.628	(0.013)
Mass of fully-informed expectations	π_b	0.673	(0.206)

Note: Income is measured in annual hourly 2010 pesos. Distance is measured in log kilometers. The likelihood contains 90,578 individual-year observations from 11,194 individuals. Asymptotic standard errors are calculated using the score of the likelihood.

The estimated preference parameters are provided in Table 6. Similar to past work using U.S. data (Bishop, 2008; Kennan and Walker, 2011), I estimate a strong preference for living in one's birthplace. The estimate implies that individuals are indifferent between spending one year away from their hometown with an income roughly twice that received in their birth location. Even with such strong attachments to this location, I still find that a large portion of the population actively considers moving each period — 63%. Similarly, I estimate a negative preference for living in an urban area. This preference is reminiscent of a compensating differential associated with living and working in an urban center (Rosen, 1986). Specifically, the average individual requires a roughly

24% higher income to be indifferent between living in an urban versus rural area.

I estimate that the behavior of more than half of the population (67%) is more consistent with a model of full-information than one of naivety.[25] I am unable, however, to reject the null hypothesis of a unit-mass of fully-informed individuals.[26] I find that individuals dislike extreme amounts of both heat and cold, as measured by degree days above 26°C and below 14°C. To put these preference estimates into perspective, one can study how flow utilities change with respect to changes in the number of degree days. In particular, the marginal flow utility of an additional degree day above 26° C is above zero when the number of degree days is below $\sim$120. Similarly, the marginal impact on flow utilities of an additional degree day below 14° C is positive when the number of degree days is below $\sim$260. The quadratic preference specification over degree days allows for a "bliss-point" interpretation, with this value defined as the peak of the relevant quadratic.

To put these values into perspective, 17% of the individual histories used in estimation experience annual degree days above 26°C which are above their bliss-point, and 28% above that for degree days below 14°C. If these same individuals experienced the 2050 business-as-usual climate, 49% of them would face a number of annual degree days above 26°C to the right of their bliss point and less than 1% would face annual degree days below 14°C to the right of the same. I explore the consequences of these outcomes in the next section.

Similar to past work in developed countries (Kennan and Walker, 2011; Oswald, 2019; Ransom, 2022), I estimate a high migration cost that is increasing in age and in the distance between the destination and origin. I also find that these costs increase once one becomes a parent. One way to understand the magnitude of the moving cost is by comparing the marginal utility of income to the intercept of the moving cost equation to translate from utils to dollars.[27] This exercise reveals a fixed migration cost of $99,000, or roughly 10 years of income.[28] This calculation is the full cost, inclusive of all psychological and monetary costs associated with a forced move in an arbitrary period to an arbitrary location.[29] Since migration is an endogenous decision, individuals often decide to move because they received a high-preference shock (represented by ε) in their

[25] I assume a homogenous discount rate across the population; which allows me to identify heterogeneity in climate-beliefs. That said, the estimate of the mass of climate-naive is locally not sensitive to changes in the discount rate.

[26] I have also estimated two different versions of the model where I assume that the climate-naive expect that future weather draws are drawn from a either a stationary or non-stationary degenerate distribution at the average weather over the last thirty years. The estimated preference parameters and the mass of movers are statistically indistinguishable from those presented here. I estimate a smaller mass of the climate-naive (8% and 12% vs 33%), but with much more precision (a standard error of 0.019 and 0.014).

[27] Appendix B.7 discusses the mapping I use to translate from utils to dollars.

[28] Here, I define the fixed cost simply by the moving intercept. The level shifters, children and age, will, in practice, increase the magnitude of this fixed cost.

[29] This value is still informative for the discussion of migration as adaptation to climate change: this is the cost of forcing an individual to adapt via migration without giving them the ability to choose the timing nor the location of said adaptation.

chosen destination. I incorporate this fact into the moving cost calculation by then subtracting off the expected maximum of $N-1$ iid type-one extreme value draws.[30] This step reduces the fixed moving cost to a moving *subsidy* of \$50,000.[31] Kennan and Walker (2011) provide another way to contextualize the moving costs by defining them as the difference in expected flow utilities between the origin and the destination, inclusive of moving costs, net of the expected difference in preference shocks, given the decision made. This definition incorporates the idea that when an individual is willing to leave a relatively desirable location for one that is less so, their unobserved preference shock in the destination must have been sufficiently high, compared to that in their origin. Calculating moving costs in this way provides for an average *subsidy* of \$50,000. Given the strong preference for living in one's birth location, the calculated moving cost back home is much higher — a cost of \$29,000. These calculations highlight the endogenous nature of migration, and the high value associated with allowing individuals to choose the timing and destination of their migration decisions in the context of dynamic adaptation to climate change. I discuss these calculations in more detail in Appendix B.4.

The above calculations highlight the highly heterogeneous nature of migration. In particular, model observables are not, by themselves, able to fully capture the incentives associated with moving in any given period to any given location. However, combining these observables with the underlying model structure, including the distributional assumptions on the preference shocks, allows one to obtain a more accurate picture of the importance placed on each component of the migration decision process. The model uses the variation described in Section 4, in conjunction with the normalized preference shocks, to provide insights into the relevance of climate on migration decisions, through both incomes and amenities. In the next section, I provide evidence that the estimated model, along with the imposed structure, does a good job of predicting migration decisions within Mexico.

Model Fit

To understand how well my estimated model fits the observed data, I use it to simulate each of the decisions of the individuals used in its estimation. Table 7 shows the average migration rates across a series of demographics in the data and across 100 simulations of the model. The model generally slightly overestimates migration rates across the population. This occurs primarily because of the retrospective nature of the MMP data. In particular, I only reproduce exactly those decisions in the data, and as such, I do not predict decisions made at points in life later than those observed (the

[30] N, here, represents the number of elements in the choice set.

[31] I am defining a moving subsidy as a negative moving cost. Given that the stochastic shocks are modeled as entering through incomes, another way to view this would be through the lens of a transitory (3-year-long) positive shock to incomes in their destination.

Table 7: Model Fit

	Migration Rate	
	Model	Data
Overall	5.4%	4.7%
Children		
Yes	5.2%	3.4%
No	5.4%	7.2%
Agricultural Worker		
Yes	4.5%	3.5%
No	5.8%	5.3%
Years of Education		
0-5	5.0%	4.2%
6-11	5.3%	3.8%
12+	6.1%	7.5%

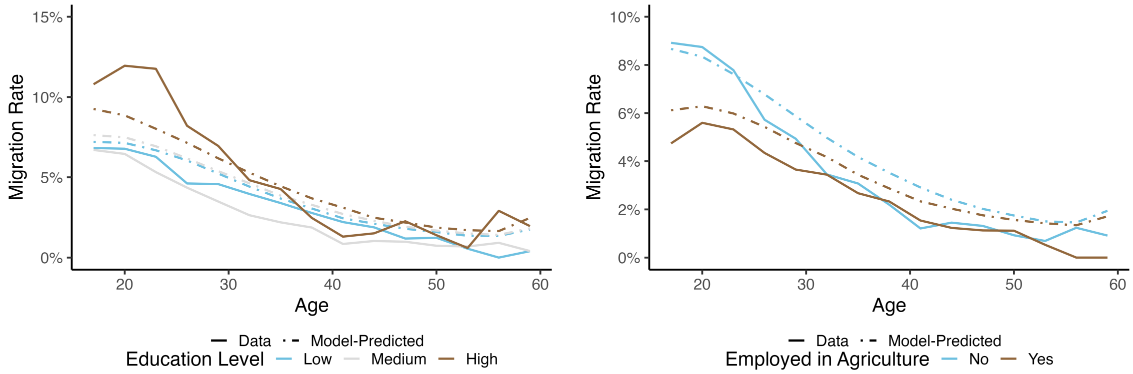
Note: For each category, I calculate the average migration predicted by 100 model simulations of the life trajectories of the individual-year observations used in estimation. The data migration rates are calculated as the average migration rate for each category.

age at which an individual is surveyed). This leads to an overprediction of migration from a higher mass of young individuals than the estimation routine accounts for. Appendix D.1 replicates Table 7 when instead simulating the entire life history of each individual. This alternative specification leads to a much closer matching of the demographic moments.

Figure 6a plots model predicted versus observed migration rates, broken apart by age and education levels. The dashed curve represents the model-predicted migration rate across ages and education levels; the solid curve represents that from the estimation data. The model does a good job of capturing trends over time and education levels. Perhaps reassuringly, the model only recognizes differences across levels of education through differences in the wage process, and even still, it is able to correctly capture the initial gap in migration rates by young individuals with high levels of education compared to those with lower levels.[32] It also correctly predicts both the convergence of migration rates across education levels as individuals age and the eventual uptick in migration rates at the end of the life cycle.

Figure 6b tells a similar story as Figure 6a, this time looking at agricultural employment. As in the setup for education, the model only incorporates an individual's sector of employment through

[32]Some of these individuals are likely moving to college. The model captures this via a higher wage opportunity in all locations.



(a) (b)

Figure 6: (a) plots the average migration rates in the data against those predicted by the model, broken down by education levels. Low, mid, and high education refers to less than 6 years, 6–11 years, and 12 or more years of education, respectively. (b) plots the average migration rates in the data against those predicted by the model, broken down by agricultural employment.

the wage process. It is still able to correctly match many of the differential life-cycle trends across these groups. In particular, the model matches the initial uptick in migration at the beginning of the life cycle for agricultural workers, but not for non-agricultural workers. It also correctly predicts the eventual convergence in rates across the two groups and the increase in rates for non-agricultural workers at the end of the life cycle. Appendix D.1 plots the same figures but from simulating decisions across the entire life cycle for all individuals, rather than just those decisions made and observed in the data.

6 Counterfactuals

I use the estimated model to simulate forward the life histories of all Census-enumerated Mexican males through 2080 under a variety of counterfactual scenarios. To do so, I combine data from the 2020 Census count and the MMP life history data. In particular, I use Census data on the number of individuals in each age bin residing in a municipality assigned to each location in 2020, along with the location-specific education and agricultural employment distributions to create an individual-level cross-section. Through 2038, I incorporate new cohorts of young men into the simulation, where the number and locations of new entrants comes from the age distribution for children in the 2020 census. Appendix F discusses this approach, and the data used, in more detail.

I use the counterfactual simulations to understand the effectiveness of migration in limiting future climate damages. In particular, I design a series of counterfactual exercises which modify the climate system, moving costs, and the ability to migrate. I use these to understand both how the value of the ability to migrate varies with the degree of warming as well as how migration directly affects expected climate damages.

The baseline scenario, a business-as-usual climate change projection, is modeled directly through an extension to the framework used in estimation. More specifically, the full solution procedure used in estimation requires solving the model through 2064, to account for the terminal period of those who turn 17 in 2018. Throughout the estimation framework, I assume that the individuals who form fully-informed expectations of the future climate use the previous thirty years of weather data to form their expectations. I use historical data through 2020, and use daily projected weather data from the SSP2-4.5 climate projections for the remaining years. This helps maintain an internally-consistent framework upon which to study counterfactual scenarios.

To understand the difference between migration decisions under business-as-usual and a world without climate change, I use historical weather data from 1950–1979 to construct a stationary climate system with limited climate change. For this scenario, I model that the climate system is governed by a Gaussian distribution with a constant mean and variance, both of which are set to the observed values from 1950–1979. As in estimation, in both the business-as-usual and stationary climate scenarios, I assume that there are both movers and non-movers and those who form fully-informed expectations and those who form naive expectations about future weather draws, with the mass of each of these groups set to the estimated proportions from the MMP data.

Business-as-Usual Warming Increases the Value of Migration

The first set of simulations shed light on how climate change will affect the value that different individuals place on the ability to choose whether or not to migrate throughout their lifetime. To study this question, I use the estimated model to simulate the choices of each Census-enumerated male for the remainder of their lives. When doing so, I track their model-predicted welfare, when they can and cannot choose to move domestically as well as when the climate is characterized by the business-as-usual scenario and the counterfactual scenario of limited warming. Mathematically, the average value of migration under scenario s, in the base year (2020), is given by:

$$\frac{1}{n(y)} \sum_{i(y)} \left\{ \tilde{V}(\ell_i, \omega_i | s) - \tilde{V}(\ell_i, \omega_i | s, \gamma_1 = \infty) \right\},$$

where $\tilde{V}(\ell_i, \omega_i|s)$ is the observed value of individual i's lifetime value under scenario s, appropriately scaled to 2024 dollars, and y indexes the year.

Figure 7 plots the results. I find that migration is valuable: the average 17–19-year-old in 2020 values the ability to move at approximately 2 years of annual income. The entering cohorts consistently have values of migration that are greater than those of the last. Panel (a) shows that by 2038, this quantity increases by 70% relative to its 2020 value. Much of both the initial level and

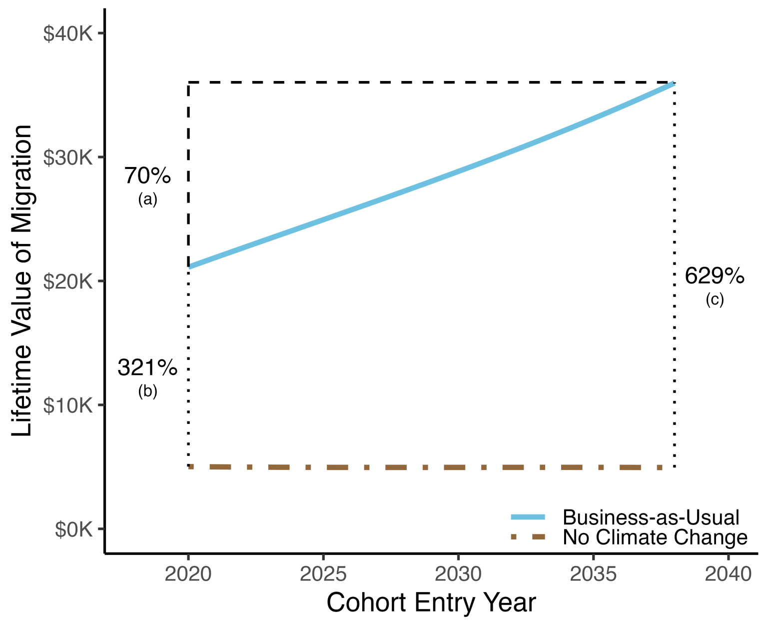


Figure 7: Value of Migration over Time

Note: Plotted is the average value of migration for 17–19-year-olds over time, with the percentage change in this value from 2020 to 2038. (a) represents the percentage change in this value under the business-as-usual climate scenario from 2020 through 2038. (b) represents the percent difference in this value between the business-as-usual and stationary climate scenarios in 2020. (c) represents the same as (b), but in 2038. All curves are smoothed with a third-order polynomial. The results are averaged over 100 simulations.

the steep change is driven by welfare damages from increased warming in the near future. Panel (b) highlights this finding: projected warming increases the lifetime value of the ability to migrate by just over four times, relative to the stationary historical climate scenario. Panel (c) combines the analyses in panels (a) and (b) to explore how the value of migration driven by climate change changes over time. I find a fairly flat representation for the value of migration in a no-warming scenario in the near future.[33] This result, combined with the acceleration in values under the business-as-usual scenario, are represented in panel (c). Projected warming stands to increase the lifetime value of migration by over seven times relative to a world without climate change

The first two columns of Table 8 decompose this finding across a variety of spatial and demographic dimensions. I find that the value of migration, for the average 17–19-year-old in 2020 across Mexico is \$21,139, or roughly 2 years of income, under the business-as-usual climate scenario — about four times higher than that under the limited climate change scenario. The results in these two columns additionally highlight the progressive nature of migration as an adapting tool for climate damages: migration is around 60% more valuable to those with the lowest relative levels of education than those with the highest.[34] The same can be said for individuals employed in

[33]This, mainly, is a model-driven result, as the only non-stationary state distribution is the joint climate distribution.

[34]The differential wage premium of urban employment for those with high education is well established in the

agriculture, relative to those employed in other sectors.

Table 8: Heterogenous Values of Migration

	Business-as-Usual Climate		
	New workers today	Children born today	Limited Climate Change
Average	21,139	36,026	5,023
Climate			
Dry	15,059	20,909	4,323
Temperate	5,812	5,415	5,292
Warm	75,687	130,389	5,197
Agricultural Worker			
Yes	35,699	62,206	4,998
No	19,210	31,959	5,026
Years of Education			
0–5	31,285	54,802	5,148
6–11	24,719	42,469	5,049
12+	19,304	32,458	5,007

Note: These are the average results of 100 simulations of the model. Values are for 17-19 year olds and are measured in 2024 dollars. The value of migration is defined as the difference between average lifetime welfare for an individual who can migrate and one who cannot. Business-as-usual refers to the SSP2-4.5 climate projections, limited climate change refers to a climate system centered around 1950–1979 averages. Children born today and new workers today refer to individuals who were less than 2-years-old and between 17–19-years-old during the 2020 Census count, respectively.

The upper panel of Table 8 presents the spatial heterogeneity value throughout Mexico. I find that individuals living in the temperate parts of the country face a relatively stable value of migration across the different climate scenarios, while those in the warm regions face an almost 15-fold increase in the value of migration under the business-as-usual scenario relative to the no-climate-change scenario. Decomposing these values across the demographic distribution of the population shows that low-skilled and agricultural workers stand to gain the most from the ability to move under a business-as-usual pathway, while the gains from migration are spread more evenly across the population under the stationary climate scenario.

I note that dynamic considerations are vital to analyzing the value of migration at limiting the damages of climate change. To see why this is true, I compare the estimated value of migration under a business-as-usual climate scenario with and without dynamics. Specifically, I use the static logit framework presented in Equation 1 to simulate the decisions of the male population

literature (c.f., Gould, 2007). I purposely abstract from such effects to aid in comparing how the relative values of migration will change with the climate across groups such as education.

of Mexico through 2080. I estimate a lifetime value of domestic migration that is two orders of magnitude lower than that given by the dynamic framework: \$520. This difference is driven by a lower static-predicted rate of migration. As displayed in Figure D5, the dynamic model predicts migration rates for young men in Mexico as being between 8% and 10% from 2020 through 2038. In comparison, the migration rate predicted by the static framework is between 3% and 5% over the same time frame, as seen in Figure D6. These differences arise from the fact that the dynamic model is able to capture the long-term benefits from migration accrued over one's lifetime, whereas the static model only captures the value from migration for a single year when an individual is deciding whether or not to move.

Stylized Framework to Incorporate Congestion Impacts on Wages

Given that migration is modeled in a partial equilibrium framework, the above analyses discount any general equilibrium effects that may lead to decreases in wages in desirable locations. For example, individuals are likely to be aware that local wages will respond to future changes in population. In other words, the places that are desirable today will attract higher portions of the population in the future, which will drive future wages down in these locations, making them less desirable today. This mechanism would work to decrease the value of migration.

It is not computationally feasible to allow for these dynamics in their base form in the dynamic model. I design a simple and stylized general equilibrium exercise to understand how serious of a concern this omission may be. To establish intuition, imagine a world in which individuals begin in one of two locations, a or b. For simplicity, suppose that individuals all agree that a is the more desirable location. However, individuals also understand that wages respond to the proportion of the population that is in a. In particular, take some individual, i, who is deciding whether to move to a. i's immediate reaction is that they would like to move to a; however, after pondering the decision, they recognize that many other people will likely also move to a, which will drive wages down. They are familiar with much of the literature on the topic and combine the expected change in populations with Borjas (2003)'s estimated elasticity of wages to labor supply to update what they expect their wages to be in location a: $w' < w$. But then they recognize that everybody else is likely performing the same set of calculations, and thus some fraction of the population will actually decide to stay in b. Consequently, wages in a will not fall all the way to w', but rather to $w'' > w'$. This process can be iterated on until $w^{(k)} \approx w^{(k+1)}$.

Convergence of this framework is similar in spirit to the level-k reasoning theory used in behavioral economics to explain why equilibrium theory fails to accurately predict outcomes in many games (Camerer et al., 2004). In each iteration, some mass of the individuals who were marginal to moving to a given location decide to instead either stay in their current location or move to another location. If all of these marginal individuals decide to stay in their current location and the set of marginal individuals is decreasing across iterations, then convergence, defined against changes in population counts in each location across iterations, is guaranteed. However, given that most individuals who decide to move are doing so not only because of the specific destination but also because of the origin, convergence is not guaranteed.

More concretely, I simulate behavior under the baseline scenario and track the percentage change in population levels in each of the locations from 2023 and 2038. I modify the wage equation to respond to these changes using an elasticity of -0.3 (Borjas, 2003), linearly from 2023 through 2038. This marks the end of the first iteration. In the second iteration, I resimulate behavior with the updated wage equation. I again keep track of population changes to arrive at a further updated wage equation to be used in the third iteration and so on. In practice, it does not take many iterations for the algorithm to converge: by iteration 12, the maximum difference in relative population changes across iterations is below 1%. The final column of Table D3 shows the updated values of migration after accounting for congestion impacts.[35] I find that the average option value across Mexico decreases by around 10%. Appendix F.4 provides more details on this exercise.

Migration Limits Climate Damages

So far, I have found that warming in the near future will lead to a stark increase in the value that individuals place on their ability to move domestically. I now ask how effective we should expect migration to be in limiting overall climate damages. I use Figure 8 to study this question. The figure plots average lifetime climate damages for the cohorts of 17–19-year-olds from 2020 through 2038. Average expected climate damages are calculated as

$$\frac{1}{n(y)} \sum_{i(y)} \left\{ \tilde{V}(\ell_i, \omega_i | \text{no cc}) - \tilde{V}(\ell_i, \omega_i | \text{BAU}) \right\}.$$

$\tilde{V}(\cdot|\text{no cc})$ represents the lifetime welfare of an individual, in 2024 dollars, when they are in a world with limited climate change, and $\tilde{V}(\cdot|\text{BAU})$ represents the same for an individual in a world governed by the SSP2-4.5 climate projections. The solid curve shows the trajectory of these damages through the 2038 cohort in the baseline scenario with estimated migration frictions. The dashed curve shows the same trajectory in the world where migration is completely restricted.

I find that climate damages are increasing over time, both under the status quo and under a world without migration. Taken together with the relative importance of amenities and wages in explaining the differential values of migration, this outcome is driven by the stark shift in the distribution of extreme heat events across Mexico depicted in Figure 2. Moreover, the figure shows that it is not only climate damages that are increasing with time, but also the gap between damages in a world

[35] The maximum difference is less than 0.09% across locations between the 16$^{\text{th}}$ and 17$^{\text{th}}$ iterations.

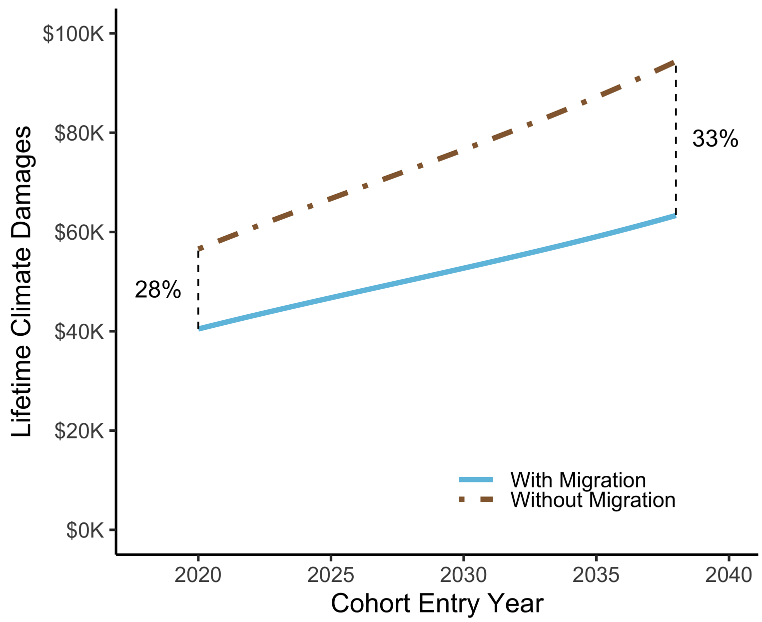


Figure 8: Migration and climate damages over time
Note: This figure plots model-predicted climate damages for 17–19-year-olds over time. The solid curve represents damages in a world with migration; the dashed curve represents damages in a world without. The results are averaged over 100 simulations.

with migration and one without: average climate damages experienced by a 17–19 year old in 2020 are 28% lower if they have the ability to migrate within Mexico. For an individual who was less than 2 years old in 2020, this migration-induced reduction in climate damages is projected to increase to 33%.

The first column of Table 9 directly measures the importance of migration in limiting climate damages for young men in 2020. The spatial heterogeneity in the decrease in damages from migration speaks directly to the mechanism through which migration limits climate damages. Specifically, individuals living in the temperate regions of the country face a 4% reduction in their expected lifetime climate damages due to their ability to migrate. This projection is in stark contrast to the 35% reduction in damages faced by those living in the warmest regions of the country who are able to move to temperate regions to escape damages.

The second column studies the same reduction in climate damages from the ability to migrate, but for the 2038 cohort. Comparing the results in the first two columns, I find that the decrease in damages from migration is expected to decline for the individuals living in temperate regions of the country, whereas it is expected to increase very slightly for those in the dry regions and to increase by more than 4 percentage points for those in the warm regions of the country. The differences in the change in the reduction in damages from migration across the country comes from two facts. First, even the temperate regions of the country are expected to face increased climate damages in

Table 9: Migration Reduces Expected Climate Damages

	% ↓ in Damages from Migration	
Age in 2020:	17–19-years-old	<2-years-old
Average	28%	33%
Climate		
Dry	27%	27%
Temperate	3%	1%
Warm	35%	40%
Agricultural Worker		
Yes	32%	37%
No	28%	32%
Years of Education		
0–5	32%	36%
6–11	30%	34%
12+	28%	32%

Note: These are the average results of 100 simulations of the model. Damages are measured in lifetime values. The percentage decrease in damages from migration is the average difference in climate damages from a world without migration to one with migration.

the future. Second, the other locations in an individual's choice set are also worsening, with the dry regions warming even faster than the temperate ones. Thus, the value that an individual who already lives in a temperate location will place on their ability to move within Mexico will attenuate over time with the degree of warming.

As an important caveat, I have assumed that the only way that one is able to adapt to climate change is through migration. Practically, this means that the estimated valuations of migration in limiting the damages of climate change will be biased upwards because some individuals who the model predicts to migrate would instead prefer to engage in in-situ forms of adaptation. That said, since I abstract away from these other forms of adapting behavior, I am underestimating the overall value of adaptation, generally defined, in limiting the damages of climate change. In this paper, I study the value of migration, a dynamic form of adaptation, at limiting the damages of climate change. Future work could study the role and limits of in-situ adaptation, as well as static forms of adaptation, to climate change. One way to do this would be to slightly modify the framework I design for this paper through allowing for a joint decision to migrate and switch occupations or to invest in additional education.

The Value of Climate Information to the Climate-Naive

The previous two sections have highlighted that migration will be an increasingly valuable form of adaptation to climate change. That said, given the dynamic interplays present in the decision to move, its value and effectiveness as an adapting strategy will depend on our understanding of future warming. Here, I use the estimated model to uncover the importance of understanding the climate system when deciding whether to dynamically adapt to climate change. I then present evidence that policy can be used to dampen the welfare losses due to naivety on the climate transition across the population.[36]

Table 10: Value of Climate Information to the Climate-Naive

	% ↓ in Climate Damages		Value of Climate Information in 2020		
	New workers today	Children born today	Dollar Value	% ↓ from $1,100 Dynamic Subsidy	% ↓ from $1,100 Static Subsidy
Average	2.2%	1.2%	1,099	18.7%	7.8%
Climate					
Dry	0.4%	0.6%	544	29.7%	12.8%
Temperate	1.3%	2.2%	298	49.0%	21.4%
Warm	3.2%	1.2%	4,300	10.4%	4.1%
Agricultural Worker					
Yes	2.9%	1.0%	2,019	13.7%	5.3%
No	2.1%	1.3%	976	20.2%	8.6%
Years of Education					
0–5	2.7%	1.1%	1,700	13.7%	6.0%
6–11	2.5%	1.2%	1,313	18.8%	7.0%
12+	2.1%	1.3%	987	19.2%	8.5%

Note: These are the average results of 100 simulations of the model. Values are for 17-19 year olds, measured in 2024 dollars. The percent decrease in climate damages (as in Table 9) represents that for the climate-naive individuals from becoming fully-informed. The value of climate information refers to the difference in average lifetime welfare between the population of fully-informed and climate-naive individuals, in 2020. The decrease in the value of climate information from a dynamic and static subsidy refers to the decrease of the dollar value of climate information from a one-time $1,100 subsidy to be used at the first time an individual moves and a take-it-or-leave-it offer, both in 2020.

In the first two columns of Table 10, I highlight the value to the climate-naive of becoming fully-informed on the climate system through the change in their expected lifetime climate damages. In 2020, I find that obtaining and using a full understanding of the climate system would lead to an expected decrease in lifetime climate damages of just over 2%. This corresponds to a dollar value of $1,099 over one's lifetime (column 3). By 2038, the expected reduction in lifetime climate damages is projected to fall to just over 1%. As Figure 2a highlights, this is because of the flattening

[36]In Appendix B.8, I derive a closed form, recursive, representation of the alternative-specific value function for the climate-naive individuals, which allows me to calculate the true welfare for the climate-naive.

of the rate of warming around 2040 under the business-as-usual climate scenario, which causes a narrowing in the gap between the decisions made by the climate-naive and the fully-informed, as seen in Figure D5.

Where do the increased climate damages due to naivety on the climate system come from? Compared to the decisions one would make with full information, I find that the climate-naive are over three-times as likely to mistakenly choose to stay in their current location as they are to mistakenly decide to move. Moreover, across the simulated time-frame, mistakenly deciding to remain in one's location induces average losses to lifetime welfare of \$9,006. The average lifetime-losses from making a mistake during a move are less than half of this: \$3,796. Taken together, this implies that subsidizing migration has the potential to be one way of reducing the heightened expected climate damages borne by the climate-naive.

The fourth and fifth columns of Table 10 highlight the effectiveness of subsidizing migration at reducing the welfare losses of the climate naive. I present the results from two potential policy options. The first is a dynamic subsidy which can be used at the first time an individual decides to move. The second is a static policy which is only available in the policy year. Both policies correspond to a subsidy roughly worth the average lifetime value to the climate naive of becoming fully-informed. Importantly, I highlight that it is possible that subsidizing migration could actually *increase* the welfare losses to the climate-naive if their losses outweighed their gains from their policy-induced moves. I find that this is not the case. Moreover, I find that dynamic policies are substantially more valuable than one-time offers.

The difference between the welfare implications of the two policy options is driven by the dynamic policy's influence on the value that each individual places on the ability to move at a reduced cost in the future. This increases the relative value of the dynamic policy through two channels. The first is that the dynamic policy will affect the welfare of everyone who considers the option of migration. Notice that even those who would not move in the policy year, with or without the policy, will receive positive value from the knowledge that they will have the option to move at a reduced cost in the future. The second is that the dynamic policy will generate a strictly higher value-per-mover than a static policy. In other words, the dynamic policy positively selects from the mass of the population marginal to moving under a static policy.

Figure 9 provides intuition for why dynamically-available policies will positively select from the set of individuals marginal to a static reduction in moving costs. I plot a conceptual distribution taken across the population. m, here, represents the expected cutoff of the maximum over the preference shocks, in the absence of policy, for which individuals would decide to move. Individuals to the left of m are inframarginal to the policy and therefore receive a transfer of s from any offered subsidy. The static subsidy of s shifts the portion of the population who decide to move to the right, as represented by m'. m' is determined by an individual's indifference condition: this represents

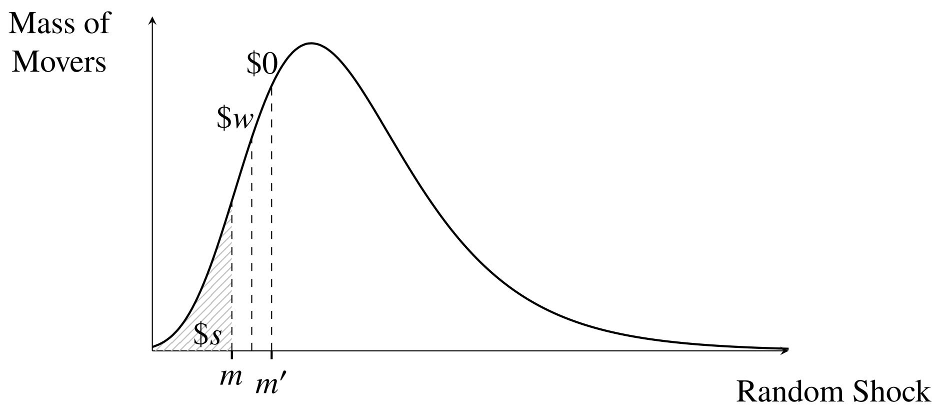


Figure 9: Dynamic and Static Policy
Note: m is the cutoff for inframarginal movers, s is the size of the subsidy, w is the size of the option value for future moving opportunities with a subsidy of s. The shaded area is the portion of the population inframarginal to any policy.

the individual who receives \$0 from moving with the subsidy, that is, they receive a welfare-loss of exactly s from moving without the policy. w represents the option value of future moves with a subsidy of s. Notice that under the dynamic policy, individuals would not move today if their value of moving was below w, as that is their value of withholding the subsidy into the future. The average value of the different subsidies to movers is the weighted average of these values across the population that decides to move, with the weights given by the mass of the population with that particular value for the policy. The fact that the dynamic policy has a positive option value is what drives the positive selection results I find.

Climate change will cause two simultaneous changes to the curve in Figure 9. First, m and m' will both shift to the right as the migration rate will increase due to warming. Second, w will shift closer to m and away from m', as represented by the increasing option value of future moves at a reduced cost. These results underscore that as the climate becomes progressively hostile, individuals will increasingly value the option to choose *when* to invoke the decision to move at a reduced cost. Figure D8 highlights this finding across time, from warming.

7 Conclusion

This paper is a first step in understanding the value of dynamic forms of adaptation to climate change, in this case, through migration. I find that incorporating the large and heterogeneous costs of such forms of dynamic adaptation are of first order importance when seeking to understand the role that adaptation will play in mitigating the damages of climate change. Even with such large migration cost estimates, I find that the value individuals place on the option to migrate is both substantial and increasing in the degree of warming. Put another way, as time progresses, individuals will be increasingly exposed to extreme heat from seemingly small shifts in average temperature. This fact will lead many to seek out less-affected areas, even within their own country. This process of increased migration in response to warming is apparent through the estimated option value that individuals place on the ability to migrate.

Although migration can be a valuable tool in humanity's fight to adapt to climate change, I emphasize that its dynamic nature implies that correct expectations of the climate system may be of first order concern. To see this, I exploit variation in rates of warming across Mexico to identify the proportions of the population forming fully-informed and naive expectations about the climate transition. I estimate that a non-negligible fraction of the population makes decisions consistent with a framework of naive expectations. This finding is supported by results from an external survey asking participants in Mexico about their understanding of climate change. I find that the value of migration would increase by 5%, and the expected climate damages would decrease by 2%, for these individuals with naive expectations over the course of their lifetime if they were to obtain accurate information on the climate transition system and use it when forming expectations about the climate to aid in adaptation. The average masks a high degree of heterogeneity: those living in the warm regions of the country and those who have been historically marginalized stand to gain the most from becoming better informed.

I find that much of the value of correct information on the climate system comes from correcting otherwise lower predicted rates of migration by the climate-naive. As such, one way to close this gap is to subsidize migration. I find that one-time reductions to moving costs at the level of the internality, taken when an individual first decides to move, reduce the loss of welfare to the climate-naive by 19%, whereas take-it-or-leave-it offers reduce the same by 8%. Given these large differences, I use the estimated model to highlight the selection effects that are caused by the different policies. In particular, dynamic reductions to moving costs generate option value from the ability to use them in the future, which reduces the contemporaneous set of individuals marginal to the policy, and in turn increases the average value of the policy.

An important avenue for future work is through studying the role of liquidity constraints on the overall cost of migration as adaptation to climate change. Past work has highlighted that warming can either increase or decrease the migration propensity (e.g., Cattaneo and Peri, 2016). Understanding the interplay between one's environment and their decision to migrate can better inform policy on how to target aid.

References

Aguirregabiria, V. and J. Jeon (2020): "Firms' Beliefs and Learning: Models, Identification, and Empirical Evidence," *Review of Industrial Organization*, 56, 203–235.

ARAGÓN, F. M., F. OTEIZA, AND J. P. RUD (2021): "Climate Change and Agriculture: Subsistence Farmers' Response to Extreme Heat," *American Economic Journal: Economic Policy*, 13, 1–35.

AUFFHAMMER, M., S. M. HSIANGY, W. SCHLENKER, AND A. SOBELZ (2013): "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change," *Review of Environmental Economics and Policy*, 7, 181–198.

BASSETTI, P. AND M. E. WESTGATE (1993): "Senescence and Receptivity of Maize Silks," *Crop Science*, 33.

BISHOP, K. C. (2008): "A Dynamic Model of Location Choice and Hedonic Valuation," .

BORJAS, G. J. (2003): "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *The Quarterly Journal of Economics*, 118, 1335–1374.

BROWNSTONE, D. AND K. TRAIN (1998): "Forecasting New Product Penetration with Flexible Substitution Patterns," *Journal of Econometrics*, 89, 109–129.

BURKE, M., J. DYKEMA, D. B. LOBELL, E. MIGUEL, AND S. SATYANATH (2015): "Incorporating Climate Uncertainty into Estimates of Climate Change Impacts," *The Review of Economics and Statistics*, 97, 461–471.

BURKE, M. AND K. EMERICK (2016): "Adaptation to Climate Change: Evidence from US Agriculture," *American Economic Journal: Economic Policy*, 8, 106–140.

BUSSE, M. R., C. R. KNITTEL, AND F. ZETTELMEYER (2013): "Are Consumers Myopic? Evidence from New and Used Car Purchases," *American Economic Review*, 103.

CAMERER, C. F., T.-H. HO, AND J.-K. CHONG (2004): "A Cognitive Hierarchy Model of Games," *The Quarterly Journal of Economics*, 119, 861–898.

CARLETON, T., E. DUFLO, B. K. JACK, AND G. ZAPPALÀ (2024): "Chapter 4 - Adaptation to Climate Change," North-Holland, vol. 1 of *Handbook of the Economics of Climate Change*, 143–248.

CARLETON, T., A. JINA, M. DELGADO, M. GREENSTONE, T. HOUSER, S. HSIANG, A. HULTGREN, R. E. KOPP, K. E. MCCUSKER, I. NATH, J. RISING, A. RODE, H. K. SEO, A. VIAENE, J. YUAN, AND A. T. ZHANG (2022): "Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits," *The Quarterly Journal of Economics*, 137, 2037–2105.

CARLETON, T. A. AND S. M. HSIANG (2016): "Social and Economic Impacts of Climate," *Science*, 353.

Carto, S. L., A. J. Weaver, R. Hetherington, Y. Lam, and E. C. Wiebe (2009): "Out of Africa and into an Ice Age: On the Role of Global Climate Change in the Late Pleistocene Migration of Early Modern Humans Out of Africa," *Journal of Human Evolution*, 56, 139–151.

Cattaneo, C. and G. Peri (2016): "The Migration Response to Increasing Temperatures," *Journal of Development Economics*, 122, 127–146.

Chipman, J. S. (1960): "The Foundations of Utility," *Econometrica*, 28, 193–224.

Cruz, J.-L. and E. Rossi-Hansberg (2023): "The Economic Geography of Global Warming," *The Review of Economic Studies*, 91, 899–939.

de Dios Ortuzar, J. (1983): "Nested Logit Models for Mixed-Mode Travel in Urban Corridors," *Transportation Research Part A: General*, 17, 283–299.

Dell, M., B. F. Jones, and B. A. Olken (2012): "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, 4, 66–95.

——— (2014): "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, 52, 740–798.

Deryugina, T., L. Kawano, and S. Levitt (2018): "The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns," *American Economic Journal: Applied Economics*, 10, 202–33.

Deschênes, O. and M. Greenstone (2007): "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *American Economic Review*, 97, 354–385.

Desmet, K. and E. Rossi-Hansberg (2015): "On the Spatial Economic Impact of Global Warming," *Journal of Urban Economics*, 88, 16–37.

Feng, S., A. B. Krueger, and M. Oppenheimer (2010): "Linkages Among Climate Change, Crop Yields and Mexico-Us Cross-Border Migration," *Proceedings of the National Academy of Sciences of the United States of America*, 107, 14257–14262.

García, E. (2004): *Modificaciones al Sistema de Clasificación Climática de Köppen (Para Adaptarlo a las Condiciones de la República Mexicana)*, no. 6 in Serie Libros, Instituto de Geografía, Universidad Nacional Autónoma de México, quinta edición: corregida y aumentada ed.

Gould, E. D. (2007): "Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium," *The Review of Economic Studies*, 74, 477–506.

GOWRISANKARAN, G., A. LANGER, AND M. REGUANT (2024): "Energy Transitions in Regulated Markets," Working Paper 32088, National Bureau of Economic Research.

HECKMAN, J. AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320.

HERRERO, M. P. AND R. R. JOHNSON (1980): "High Temperature Stress and Pollen Viability of Maize," *Crop Science*, 20, 796–800.

HOTZ, V. J. AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60, 497–529.

HSIANG, S. AND R. E. KOPP (2018): "An Economist's Guide to Climate Change Science," *Journal of Economic Perspectives*, 32, 3–32.

IFRACH, B. AND G. Y. WEINTRAUB (2016): "A Framework for Dynamic Oligopoly in Concentrated Industries," *The Review of Economic Studies*, 84, 1106–1150.

IPCC (2022): *Climate Change 2022: Impacts, Adaptation and Vulnerability*, Summary for Policymakers, Cambridge, UK and New York, USA: Cambridge University Press.

JESSOE, K., D. T. MANNING, AND J. E. TAYLOR (2018): "Climate Change and Labour Allocation in Rural Mexico: Evidence from Annual Fluctuations in Weather," *Economic Journal*, 128, 230–261.

KALA, N. (2017): "Learning, Adaptation, and Climate Uncertainty: Evidence from Indian Agriculture," .

KENNAN, J. (2008): "Logit Switching Costs," Accessed: 2024-01-05.

KENNAN, J. AND J. WALKER (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79, 211–251.

KOŞAR, G., T. RANSOM, AND W. VAN DER KLAAUW (2022): "Understanding Migration Aversion Using Elicited Counterfactual Choice Probabilities," *Journal of Econometrics*, 231, 123–147.

KÖPPEN, W. (1884): "Die Wärmezonen der Erde, nach der Dauer der heissen, gemässigten und kalten Zeit und nach der Wirkung der Wärme auf die organische Welt betrachtet," *Meteorologische Zeitschrift*, 1, 215–226.

LEISEROWITZ, A., J. CARMAN, N. BUTTERMORE, L. NEYENS, S. ROSENTHAL, J. MARLON, J. SCHNEIDER, AND K. MULCAHY (2022): "International Public Opinion on Climate Change, 2022," .

Lemoine, D. (2024): "Estimating the Consequences of Climate Change from Variation in Weather," Working Paper 25008, National Bureau of Economic Research.

Lessem, R. (2018): "Mexico-U.S. Immigration: Effects of Wages and Border Enforcement," *Review of Economic Studies*, 85, 2353–2388.

Liu, M., Y. Shamdasani, and V. Taraz (2023): "Climate Change and Labor Reallocation: Evidence from Six Decades of the Indian Census," *American Economic Journal: Economic Policy*, 15, 395–423.

Livneh, B., T. J. Bohn, D. W. Pierce, F. Munoz-Arriola, B. Nijssen, R. Vose, D. R. Cayan, and L. Brekke (2015): "A Spatially Comprehensive, Hydrometeorological Data Set for Mexico, The U.S., and Southern Canada 1950-2013," *Scientific Data*, 2.

Makridis, C. and T. Ransom (2018): "Beating the Heat: Temperature and Spatial Reallocation over the Short and Long-run," Working paper, SSRN.

Mathes, S. (2024): "The Dynamics of Residential Sorting and Health: Implications of Climate Change in the U.S." *International Economic Review*, forthcoming.

McFadden, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, 105–142.

McKee, T. B., N. J. Doesken, J. Kleist, et al. (1993): "The Relationship of Drought Frequency and Duration to Time Scales," in *Proceedings of the 8th Conference on Applied Climatology*, California, vol. 17, 179–183.

Mendelsohn, R. (2019): "An Examination of Recent Revealed Preference Valuation Methods and Results," *Review of Environmental Economics and Policy*, 13, 267–282.

Mendelsohn, R., W. D. Nordhaus, and D. Shaw (1994): "The Impact of Global Warming on Agriculture: A Ricardian Analysis," *The American Economic Review*, 84, 753–771.

Miller, S., K. Chua, J. Coggins, and H. Mohtadi (2021): "Heat Waves, Climate Change, and Economic Output," *Journal of the European Economic Association*, 19, 2658–2694.

Missirian, A. and W. Schlenker (2017): "Asylum Applications Respond to Temperature Fluctuations," *Science*, 358, 1610–1614.

Montoison, A., P. Pascal, and L. Salomon (2020): "NOMAD.jl: A Julia interface for the constrained blackbox solver NOMAD," https://github.com/bbopt/NOMAD.jl.

Mora, C., B. Dousset, I. R. Caldwell, F. E. Powell, R. C. Geronimo, C. R. Bielecki, C. W. W. Counsell, B. S. Dietrich, E. T. Johnston, L. V. Louis, M. P. Lucas, M. M. McKenzie, A. G. Shea, H. Tseng, T. W. Giambelluca, L. R. Leon, E. Hawkins, and C. Trauernicht (2017): "Global Risk of Deadly Heat," *Nature Climate Change*, 7, 501–506.

Nakamura, E., J. Sigurdsson, and J. Steinsson (2021): "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," *The Review of Economic Studies*, 89, 1557–1592.

Oswald, F. (2019): "The Effect of Homeownership on the Option Value of Regional Migration," *Quantitative Economics*, 10, 1453–1493.

Ransom, T. (2022): "Labor Market Frictions and Moving Costs of the Employed and Unemployed," *Journal of Human Resources*, 57, S137–S166.

Riahi, K., D. P. van Vuuren, E. Kriegler, J. Edmonds, B. C. O'Neill, S. Fujimori, N. Bauer, K. Calvin, R. Dellink, O. Fricko, W. Lutz, A. Popp, J. C. Cuaresma, S. KC, M. Leimbach, L. Jiang, T. Kram, S. Rao, J. Emmerling, K. Ebi, T. Hasegawa, P. Havlik, F. Humpenöder, L. A. Da Silva, S. Smith, E. Stehfest, V. Bosetti, J. Eom, D. Gernaat, T. Masui, J. Rogelj, J. Strefler, L. Drouet, V. Krey, G. Luderer, M. Harmsen, K. Takahashi, L. Baumstark, J. C. Doelman, M. Kainuma, Z. Klimont, G. Marangoni, H. Lotze-Campen, M. Obersteiner, A. Tabeau, and M. Tavoni (2017): "The Shared Socioeconomic Pathways and Their Energy, Land Use, and Greenhouse Gas Emissions Implications: An Overview," *Global Environmental Change*, 42, 153–168.

Rode, A., T. Carleton, M. Delgado, M. Greenstone, T. Houser, S. Hsiang, A. Hultgren, A. Jina, R. E. Kopp, K. E. McCusker, I. Nath, J. Rising, and J. Yuan (2021): "Estimating a Social Cost of Carbon for Global Energy Consumption," *Nature*, 598, 308–314.

Rosen, S. (1986): "Chapter 12 The Theory of Equalizing Differences," Elsevier, vol. 1 of *Handbook of Labor Economics*, 641–692.

Rudik, I., G. Lyn, W. Tan, and A. Ortiz-Bobea (2022): "The Economic Effects of Climate Change in Dynamic Spatial Equilibrium," .

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

——— (1988): "Maximum Likelihood Estimation of Discrete Control Processes," *SIAM Journal of Control and Optimization*, 26.

Sarvimäki, M., R. Uusitalo, and M. Jäntti (2022): "Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations," *Journal of the European Economic Association*, 20, 2497–2539.

Schlenker, W. and M. J. Roberts (2009): "Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields Under Climate Change," *Proceedings of the National Academy of Sciences of the United States of America*, 106, 15594–15598.

Seo, S. N. and R. Mendelsohn (2008): "An Analysis of Crop Choice: Adapting to Climate Change in South American Farms," *Ecological Economics*, 67, 109–116.

Sherwood, S. C. and M. Huber (2010): "An Adaptability Limit to Climate Change Due to Heat Stress," *Proceedings of the National Academy of Sciences*, 107, 9552–9555.

SIAP (2022): *Panorama Agroalimentario 2022*, Ciudad de México: Servicio de Información Agroalimentaria y Pesquera, edición 2022 ed.

Simon, H. A. (1955): "A Behavioral Model of Rational Choice," *The Quarterly Journal of Economics*, 69, 99–118.

Thom, K. (2010): "Repeated Circular Migration: Theory and Evidence from Undocumented Migrants," .

Thornton, M., R. Shrestha, Y. Wei, P. Thornton, S.-C. Kao, and B. Wilson (2022): "Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1," .

Thrasher, B., W. Wang, A. Michaelis, F. Melton, T. Lee, and R. Nemani (2022): "NASA Global Daily Downscaled Projections, CMIP6," *Scientific Data*, 9.

Timmermann, A. and T. Friedrich (2016): "Late Pleistocene Climate Drivers of Early Human Migration," *Nature*, 538, 92–95.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2010): "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index," *Journal of Climate*, 23, 1696 – 1718.

Willis, K. J., K. D. Bennett, D. Walker, and P. Forster (2004): "Ice Ages and the Mitochondrial DNA Chronology of Human Dispersals: A Review," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359, 255–264.

Wilson, L. and W. Barnett (1983): "Degree-days: An Aid in Crop and Pest Management," *California Agriculture*, 37, 4–7.

Wolpin, K. I. (1984): "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy*, 92, 852–874.

Zappalà, G. (2024): "Adapting to Climate Change Accounting for Individual Beliefs," *Journal of Development Economics*, 169, 103289.

Appendices

A Data

Table A1: Data Sources

Data Source	Purpose	Years
Mexican Migration Project (MMP)	Estimation (Life-histories)	1950–2019
ENIGH, ENE surveys	Estimation (Income)	1984–2019
Livneh et al. (2015)	Estimation (Weather)	1930–1979
Daymet	Estimation (Weather)	1980–2019
NASA NEX-GDDP CMIP6	Estimation/Simulations (Weather)	2020–2100
Mexican full-count Census	Simulations	2020

Table A1 presents the data sources used in this paper.

A.1 Mexican Migration Project

The MMP is a joint data collection effort between Princeton University and Universidad de Guadalajara that "aims to gather and maintain high-quality data on the characteristics and behavior of documented and undocumented Mexican migrants to the United States"[37]. To accomplish this, the MMP has sent surveyors out to between three and five pre-selected communities each Winter since 1982, where data are collected from a random sample of 200 households.

During these interviews, the surveyors collect various types of information from each household head. These include demographic and health information about both the head of the household and their spouse, detailed information on the first and last domestic and international migratory trips by both of them as well as any other migrants in their household, and a variety of different economic variables for these individuals, such as their current (time of the survey) wage. Importantly, each household head also provides the surveyor with their life history at the yearly level and that of their spouse. In this history, they are asked to provide information about where they have lived and their primary occupation during every year of their life. The publicly available dataset containing these variables censors some of the location data to preserve the confidentiality of those individuals surveyed;[38] however, I was provided access to the censoring algorithm and a crosswalk necessary for de-censoring the data so that I am able to verify the exact municipality of residence for the majority of individuals surveyed. This is the primary section of the survey that I exploit in this project.

The following restrictions are made in selecting the sample used to estimate the structural model:

- Birth date not before 1950

[37] Quote is taken from the MMP's home webpage, accessed 03/20/23

[38] 1.9% of person-year observations were either censored or left blank due to issues with recall on the part of the surveyed individual.

- Birth location within Mexico
- After the de-censoring algorithm is employed, known birth location
- After the de-censoring and imputation algorithms are employed[39], municipality of residence for each year of the life history is known and is in Mexico.
- No years with missing state variables

The first condition is for two reasons. First, I use the previous 30 years' of weather realizations when estimating state transitions. Restricting births to not before 1950 implies that I begin modeling choices in 1967, which means that I require weather data from 1930 (I discretize years into decades which transition stochastically). Moreover, it helps ensure that the estimated wage data are not used to predict too far out sample. Further, I require that each individual be born in Mexico and have a known municipality of birth because past work has shown that individuals strongly prefer to live in their home location (e.g., Kennan and Walker, 2011). The remaining requirements are enforced to guarantee a complete panel for each individual in the data, which is necessary to estimate the dynamic model.

The life history data has an annual frequency; however, the model is instead designed with a triennial structure. To convert the data to this frequency, I employ the following algorithm, given individual i at time-step t:

- If there are no moves during the three-year period, i's starting location at time $t+1$ is the same as their location at time t. Assign state variables according to the first year in this period.[40]
- If there is a move during the three-year period, i's starting location at time $t+1$ is the location of their first move during this period. Assign state variables according to the year of this move.

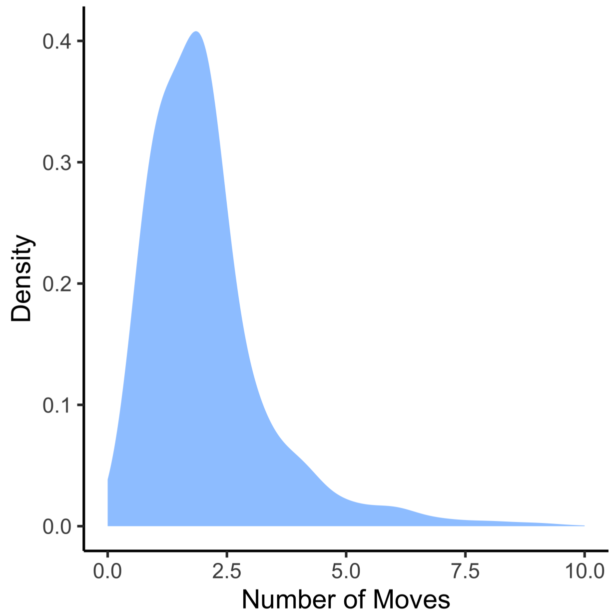
[39]Even after employing the de-censoring algorithm, there are still some individuals (834) who live in an unknown municipality at some point in their life. I drop any of these individuals who live in an unknown municipality and an unknown state. For the remaining individuals, given knowledge of the state, but not municipality, of residence, I use the following algorithm to impute their location in those years:

- If they live in a state that they do not again live in with a known municipality, drop them from the data
- If they live in a state where they either have previously lived or will live in the future, and at that time there is a single known municipality of residence, assign the years that they reside in that state, but with an unknown municipality, as the same known municipality in that state.
- If they live in a state where they either have previously lived or will live in the future, and also have lived/will live in multiple known municipalities in that state, assign the years that they reside in that state, but with an unknown municipality, with the municipality that they live in the year nearest the year with missing information.

This algorithm removes 517 individuals from the panel.

[40]For example, the first time period covers ages 17-20. If i does not move during these years, then their state variables for this period are those when they are 17.

Figure A1: Density of Number of Moves in the MMP Data, For Migrants



A.2 Wages

Given a lack of robust wage data in the MMP life histories, I use Census surveys to estimate a wage equation. More specifically, I use the 1984, 1989, 1992, 1994, 1996, 1998, 2000, 2002, 2004, and 2006 versions of La Encusta Nacional de Ingresos y Gastos de los Hogares (ENIGH), and the quarterly 2008 through 2019 versions of La Encuesta Nacional de Ocupación y Empleo (ENOE). I use these data to create a repeated cross section of wages for each of the above years. When forming this cross section, I restrict the sample to males, who are the head of their household between the ages of 17 and 61, with positive hours worked and wages, who are not self-employed or working for a cooperative, and who live in one of the municipalities visited by an individual in the MMP data. To adjust for outliers, I remove the top and bottom 5% of the wage data within each location-year. This leaves me with 3,626,618 observations. Table A2 presents the estimated coefficients corresponding to the weather impacts on wages.

A.3 Köppen Climates

The Köppen climate classification system is a widely used system for classifying locations in space by their differing local climates (Köppen, 1884). This system is global in scale, and as such is not able to capture local idiosyncrasies in climate. For this reason, El Instituto de Geografía and La

Table A2: Wage Equation Estimates, Weather

	Hourly Income
Ag worker	0.74
	(3.55)
Growing Degree Days	7.5***
	(2.8)
Growing Degree Days^2	-0.95**
	(0.44)
Drought	-0.38*
	(0.20)
Inundation	-0.036
	(0.310)
Ag worker × GDD	-3.0
	(2.8)
Ag worker × GDD^2	0.58
	(0.52)
Ag worker × Drought	0.38
	(0.25)
Ag worker × Inundation	-0.44*
	(0.24)
Temperate × GDD	-5.1
	(3.8)
Warm × GDD	-6.4
	(14.6)
Temperate × GDD^2	0.75
	(0.75)
Warm × GDD^2	0.87
	(2.17)
Temperate × Drought	0.77***
	(0.22)
Warm × Drought	0.66**
	(0.33)
Temperate × Inundation	-0.56*
	(0.32)
Warm × Inundation	-0.13
	(0.36)
Num.Obs.	3626618
R2	0.231
FE: Location, Year	X
Dep. Var. Mean	27.06

Note: Estimates correspond to the coefficients from Equation 7 in the main text. Wages are measured in 2010 hourly pesos. Standard errors, in parentheses, are clustered at the municipality level. *p < 0.1; **p < 0.05; ***p < 0.01

Universidad Nacional Autónoma de México have developed a more detailed classification system for Mexico (García, 2004). This system is based on the Köppen system, but is more detailed with respect to the local nuances of the climates within Mexico. I use this updated classification system in this paper, and I discuss the specifics of it here.

I assign each municipality their climate class based on the classification which contains the most area of the municipality. The classification system is broken into four main groups: warm (cálido), dry (seco), temperate (templado), and cold (frío). None of the municipalities have a majority of their area in the cold region, so I do not discuss this group further. Each of the three remaining groups is further broken down into their subgroup: warm into warm warm and warm semi-warm, dry into dry very dry, dry dry, and dry semi-dry, temperate into temperate semi-warm, temperate temperate and temperate semi-cold. The classification is then broken down one more time: warm and temperate groups into those with humid and subhumid climates, and the dry group into very warm, warm, semi-warm with warm winter, semi-warm with cool winter, temperate with warm summer, temperate with cool and long summer and semi-cold.

There are no municipalities in the warm semi-warm humid class, so there are three warm climates: warm warm humid, warm warm subhumid, and warm semi-warm subhumid.

I combine temperate semicold humid (one municipality) and temperate semicold subhumid (nine municipalities) into a single class, temperate semi-cold. There are therefore five temperate climates: temperate semi-warm humid, temperate semi-warm subhumid, temperate temperate humid, temperate temperate subhumid, and temperate semi-cold.

I combine all dry climates based on their second grouping with "warm" and "temperate" third level classifications (there are no temperate semi-cold municipalities). This results in six dry climates: dry warm semi-dry, dry temperate semi-dry, dry warm dry, dry temperate dry, dry warm very dry, and dry temperate very dry.

Figure 4, in the main text, plots the resulting climates throughout Mexico.

A.4 Metropolitan Zones

La Secretaría de Desarrollo Social, El Consejo Nacional de Población, and El Instituto Nacional de Estadística y Geografía have developed a classification system for municipalities in Mexico based on their population size and economic activity into metropolitan zones (las zonas metropolitanas). This system is explictly designed to determine which parts of the country are "urban". I use their 2004 definitions, based on 2000–2010 populations, to classify each municipality as either urban, if there is a metropolitan zone in the municipality, and rural, otherwise. There are 63 such metropolitan zones, covering 1,021 municipalities.

B Methods & Empirical Implementation

B.1 Climate & Weather

The climate is a high dimensional physical system upon which conventional measures of weather are drawn. Given the complexity of the climate system and the need for a high degree of parsimony in the model to maintain computational feasibility, I assume that individuals keep track of two summary statistics in each of the 3 climates in Mexico: the average of the daily temperature distribution and the total level of agricultural-season precipitation. In total, individuals are assumed to keep track of 6 weather statistics, each of which contribute to their observed state.[41] Even though individuals are assumed to keep track of summary statistics of the climates, past work has shown that it is oftentimes extremes which drive economic outcomes. For this reason, I use rich reduced-form mappings of the relationship between these summary statistics and the relevant extreme outcomes to ensure that the model is able to capture the relevant variation driving migration decisions.

Individuals use the following mapping to determine the expected value of degree days above 26 degrees Celsius, degree days below 14 degrees Celsius, and growing degree days:

$$DD_\ell = \mathbb{1}\{\alpha_\ell^D + \alpha_{1\ell}^D \text{temp}_{C(\ell)} > 0\} \times \left[\alpha_\ell^D + \alpha_{1\ell}^D \text{temp}_{C(\ell)}\right]$$

Where each location ℓ has a different set of coefficients, α_ℓ^D for each of the degree day types. These coefficients are estimated using a censored Tobit framework to account for the fact that the degree days are non-negative and censored at 0. Table B1 provides summary statistics for these mappings, showing the number of zeros in each climate for each type of degree day, the average fixed effect (α_ℓ^D), and the average linear coefficient ($\alpha_{1\ell}^D$). Similarly, individuals use the following mapping to determine the probability of either drought or flooding in each of the locations, as predicted by the level of precipitation in the agricultural season and the average temperature in the relevant level-one Köppen climate:

$$\text{prob}(R_\ell) = \Phi\left(\gamma_\ell^R + \gamma_1^R \text{P}_{C(\ell)} + \gamma_2^R \text{temp}_{C(\ell)}\right)$$

Where each location ℓ has its own fixed effect but a shared set of precipitation and temperature coefficients, for each type of extreme rain event. These coefficients are estimated using probit regression. Table B2 provides summary statistics for this mapping, showing the average fixed effect (γ_ℓ^R), and the shared precipitation and temperature coefficients.[42]

[41] This is similar, in spirit, to a moment-based Markov equilibrium (Ifrach and Weintraub, 2016), where individuals keep track of the states of a few dominant firms and a few summary statistics for the mass of fringe firms.

[42] It is important to note that drought and flood are defined using the SPEI, which accounts for differences in potential and realized evapotranspiration. This means that these events are defined relative to the local climate, rather than in absolute terms.

Table B1: Weather Mapping Statistics

	Growing	DD26	DD14
Dry			
N Zeros	0	1	0
Avg. FE	-580.75	-2222.66	1539.43
Avg. Temp.	173.14	115.92	-70.63
Temperate			
N Zeros	0	4	0
Avg. FE	-929.01	-1733.00	2035.91
Avg. Temp.	172.86	44.89	-96.88
Warm			
N Zeros	0	0	1
Avg. FE	-442.82	-1936.38	-192.89
Avg. Temp.	169.49	138.17	57.38

Note: Estimates correspond to the coefficients from the Tobit regression, as described in Section 3. N Zeros represents the number of locations with no degree days in that category. Average fixed effect (FE) and temperature are the average of the (non-zero) fixed effects and level-one climate impacts, respectively, for locations in that climate. DD26 and DD14 are degree days above 26 and below 14 degrees Celsius, respectively.

B.2 Transitions

I employ a two-stage estimation routine, where in the first stage I estimate the distribution of state variable transitions, which I feed into the second stage nested-full solution maximum likelihood algorithm. I discuss here the estimates from the first stage.

An individual's state is described by the following set:

$$\left\{ \underbrace{v^{\ell}, educ, ag,}_{\text{Constant}} \; \underbrace{age}_{\substack{\text{Deterministic}\\ \text{Trans.}}} \;, \underbrace{dec, k, \mathbf{w},}_{\substack{\text{Stochastic}\\ \text{Trans.}}} \; \underbrace{\ell}_{\text{Endogenous}} \;, \underbrace{\tau}_{\text{Unobserved}} \;, \varepsilon \right\}$$

Where v^{ℓ} is an individual's birth location, $educ$ is either low, medium, or high representing less than 6, between 6 and 12, and at least 12 years of education, respectively, ag is 1 if the individual works in agriculture at least 25% of their observed life history and 0 elsewise. dec transitions to the next decade with probability 0.3 and remains the same with probability 0.7. age transitions to the

Table B2: Rainfall Probability Estimates

	Drought	Flood
Precipitation	-1.34	2.31
Temperature	0.11	-0.15
Climate (Average FE)		
Dry	-2.54	1.29
Temperate	-2.09	0.52
Warm	-2.21	0.68

Note: Estimates correspond to the coefficients from the probit regressions of level-one climate variables on drought and flood risk, described in Section 3. The bottom panel reflects the average fixed effect across climates for the particular rainfall event.

Table B3: Weather Mapping Fit

Growing Degree Days		Degree Days above 26	Degree Days below 14	Drought	Flood
R^2: Full	R^2: No FE	Pseudo R^2			
0.985	0.647	0.239	0.239	0.041	0.101

Note: Growing degree days are for maize, and refer to degree days with a lower kink at 8°C and an upper kink at 32°C. Drought and flood refer to an SPEI below (above) 1.5. I use the McFadden pseudo R^2. All statistics refer to the mappings presented in the main text: Equation 5 and Equation 6.

next time-step with probability 1. τ represents the moving-belief type, ℓ the current location, and ε the vector of location-individual-time-specific T1EV draws.

k is an indicator for whether or not an individual is currently a parent. This is an absorbing state; that is, once an individual is a parent, they will always be a parent: $k_t = 1 \Rightarrow k_s = 1 \; \forall s \geq t$. If $k_t = 0$, then $k_{t+1} = 1$ with some probability p, where $p = \Phi(\gamma_0 + \gamma_1 age + \gamma_2 age^2)$. I estimate γ using the MMP data with a probit regression - Table B4 presents the estimated coefficients.

$\boldsymbol{w}$ represents a six-dimensional vector containing the average daily temperature and level of agricultural-season precipitation in each of the six climates in Mexico. I assume that this vector follows a multivariate Gaussian:

$$\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{\mu}_y, \Sigma_y)$$

Where $\mu_{iy} = \rho_{i0} + \rho_{i1} y + \rho_{i2} \mu_{it-1}$. The mean of each component of this Gaussian is assumed to be an AR(1) process, augmented with a time trend to account for climate change over time. Table B5 presents the estimates of ρ for the historical climate used in estimation. Σ represents the covariance matrix, obtained via a seemingly unrelated regression of each μ_i.

Table B4: First Stage Parenthood Transitions

	$\mathbb{1}\{\text{New Child}\}$
Intercept	-1.785***
	(0.018)
Time Step	0.377***
	(0.011)
Time Step Squared	-0.0373***
	(0.0014)
Num.Obs.	82256
RMSE	0.32

Notes: New Child is a binary variable representing whether an individual becomes a new parent in the following period. That is, they have no children in the current period and at least one child in the following. Estimation is done using a probit regression on the subsample of non-parents in the MMP data. $^{***}p < 0.01$.

B.3 Identification of Preference Parameters

Section 4 of the main text provides examples of the sources of variation in the estimation data which identify the preference parameter estimates. Here, I highlight two sources of variation, whose existence can be verified using reduced form statistics. In particular, the first column of Table B6 highlights differences in average migration rates across differences in wages, age, experienced weather, whether an individual has migrated in the past, and whether they were born in their current location. All of these differences are strongly significant conditional on the others and differences across rural and urban locations. The second column of Table B6 highlights the relationship between the length of stay for migrants, measured in years, and the same set of variables. This set of comparisons highlights a source of identification for the preference for living in an urban location, conditional on all other variables, including expected wages.

B.4 Moving Costs

The estimated structural model allows for a calculation of frictions associated with any particular move seen in the estimation data. Section 5 discusses three different ways of contextualizing such estimates. The moving cost for an individual observed moving from location ℓ to location ℓ', with current-period state variables given by ω[43] ($\bar{\Delta}(\ell, \ell', \omega)$), the moving cost, is defined as the difference

[43]Note, here I include an individual's time-step, t in ω. This is different from the definition given the main text.

in flow utilities, inclusive of the estimated moving cost, given by 3, net of the expected difference in the preference shock, conditional on making choice ℓ'. Mathematically:

$$\bar{\Delta}(\ell,\ell',\omega) = u(\ell,\ell,\omega) - u(\ell,\ell',\omega) - \mathbb{E}[\varepsilon_{\ell'} - \varepsilon_{\ell}|c = \ell',\ell,\omega] \tag{10}$$

Borrowing from the conditional choice probability literature (Hotz and Miller, 1993), Kennan and Walker (2011) show that Equation 10 can be rewritten as something equivalent to:

$$\bar{\Delta}(\ell,\ell',\omega) = \frac{log(p(\ell,\ell,\omega))}{1 - p(\ell,\ell,\omega)} + \\ \delta^3 \int \left[\log\left(\sum_{k=1}^{27} \exp(\bar{v}_{t+1}(\ell',k,\omega')) \right) - \log\left(\sum_{k=1}^{27} \exp(\bar{v}_{t+1}(\ell,k,\omega')) \right) \right] dF_{\omega'|\omega,\tau} \tag{11}$$

Equation 11 contains three ingredients worth discussing. The first of these is $p(\ell,\ell,\omega)$. This is the model-predicted probability of an individual residing in location ℓ not moving in their current period, conditional on their current state. The moving cost is increasing in this value. Intuitively, the probability of not moving is positively related to the observable difference in flow utility from staying in the current location and moving to any other location, including ℓ'. For this reason, it accounts for part of the first difference in Equation 10. The second and third ingredient represent the discounted continuation value from starting next period in ℓ' and ℓ, conditional on ω, respectively. The moving cost is increasing in the continuation value of starting in location ℓ'. To see why this is, note that conditional on a given difference in flow utilities and ε_{ℓ}, a higher expected continuation value from choosing location ℓ' means that the conditional ε' has less work to do to convince an individual to move to ℓ'. This causes Equation 10 to be larger, again, all else held equal. Similar logic can explain the presence of the final term: the continuation value associated with remaining in ℓ. Table B7 provides statistics on the expected moving costs of the moves used in estimating the structural model.

B.5 Estimation

I discuss here a few technical details regarding estimation of the dynamic model. The first of these is the discretization procedure used. As discussed in Appendix B.2, the only continuous state variable in the model is the set of weather states. I discretize each of the continuous weather variables into 3 bins, where the lower and upper limits are determined such that practically all of the observed data lie within them. The dry and temperate climates each have a temperature bin-width of 1 degree Celsius, while the warm climate has a bin-width of 0.5 degrees Celsius.[44] Similarly, the dry climate

[44]There is much less variation in mean temperature in the warm climate than in the other two.

has an agricultural-season bin-width of 0.1 meters, while the temperate and warm climates have a bin-width of 0.2 meters.

I use multilinear interpolation to interpolate the value function outside/inside of these bins. In particular, for a given weather state, $\boldsymbol{w}$, I find the nearest bin above and below w_i for each i in 1 through 6.[45] I then form the set of all combinations of these grid points, there are 64 (2^6) such combinations. I then calculate the Euclidean distance from $\boldsymbol{w}$ to each combination, and set the relevant value, at $\boldsymbol{w}$, as the inverse-distance weighted average of the value at each combination.

The above discretization leads to a state space with roughly 4.6 billion different points. The model solution explicitly solves the Bellman at each of these points. The MLE framework requires that the model be solved thousands of times.[46] The large and non-stationary state space is necessary to study two intrinsically dynamic objects: migration and climate change. In estimation, I exploit separability across many large swathes of the state space to parallelize the computation across a cluster of 9 high performance GPU nodes on the University of Arizona's High Performance Computing cluster. I also exploit sparsity of many sections of the transition distribution to further reduce the computational burden[47]. Moreover, I derive a closed form representation of the model solution which avoids numerical integration over the preference shocks. Using NVIDIA P100 GPUs, each likelihood evaluation takes around three minutes, with the likelihood maximized using a highly-effective black-box solver (Montoison et al., 2020).

B.6 Standard Errors

To obtain standard errors on the structural parameter estimates, I calculate central differences around the optimal parameter vector. To be clear, I calculate standard errors using the following algorithm:[48]

1. Perturb θ_k^* above and below by $|\theta_k^*|\sqrt{\epsilon}$
2. Calculate $\log(prob_i|\theta_k^p)$ for each perturbation direction (p), k, and i
3. Calculate $\frac{\partial log(prob_i)}{\partial \theta}$ using finite central differences
4. Create individual contributions to the score as $\frac{\partial log(prob_i)}{\partial \theta}\frac{\partial log(prob_i)}{\partial \theta}'$
5. Calculate information matrix as $\mathcal{I} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial log(prob_i)}{\partial \theta}\frac{\partial log(prob_i)}{\partial \theta}'$
6. Invert $\mathcal{I}$

[45] If w_i is above (below) the upper (lower) limit, then I use the upper (lower) limit for both of the points described.

[46] For example, the primary version of the model required 2,639 evaluations of the solution procedure.

[47] I employ a custom compressed row storage (CRS) routine for the transition matrix. In particular, I modify the generic CRS algorithm to allow for a sparse tensor. This transition object is a tensor rather than a matrix as each decade has its own 2-d transition matrix.

[48] ϵ, here, represents machine precision for 16-bit floats

7. Divide $\mathcal{I}^{-1}$ by the number of people (n)
8. The square root of the resulting diagonal converges to the vector of standard errors

The resulting standard errors are clustered at the individual level, due to the finite-mixture structure imposed on the problem.

B.7 Translation to Dollars

I discuss here the mapping I use to translate values, as measured in utils, into dollars. Mathematically, suppose that we have some difference in value functions, u, which is measured in utils. Notice that the estimated marginal utility of income, θ_1 from Equation 2, is measured in utils per 2010 peso per hour per year. Then normalizing u by θ_1 delivers u_1, which is measured in 2010 pesos per hour per year. I multiply this by the average number of hours worked by an individual living in Mexico in 2010: 2244.[49] This gives the value of u_2, which is measured in 2010 pesos. I bring u_2 to 2010 dollars using the exchange rate on January 1, 2010: 12.8096 pesos/dollar[50] and then to 2024 dollars using the consumer price index for all urban consumers: 1 2010 dollar to 1.45 2024 dollars.

B.8 Calculating Biased Welfare

I discuss here how I calculate the welfare of individuals forming naive expectations of the climate system. Before doing so, I highlight that for the fully-informed individuals, their remaining lifetime welfare is given by the sum of their alternative specific value function associated with the state and particular period t choice and the logit draw for that decision. This is because their expectations mirror the true climate system. On the other hand, the naive individuals make their choices based off a biased Bellman, and as such, their welfare will be lower than that given by this Bellman.

I highlight that in the terminal period a naive and fully-informed individual share the same true alternative-specific value function. This is because after this period there are no further choices to be made. I use this fact to recursively calculate the welfare for naive individuals. Mathematically, at time T, a naive individual's true welfare (net of their current period stochastic shock) is given by:

$$\overline{V}_T^{nt}(\ell, \ell', \omega) = \overline{V}_T(\ell, \ell', \omega),$$

where $\overline{V}_T(\ell, \ell', \omega)$ is as given by Equation 9. For time $t < T$, the naive individuals welfare (net of

[49] I get this figure from: OECD Data, accessed 7/9/2024.

[50] FRED Data, accessed 7/9/2024.

the stochastic logit draw) is given by:

$$
\begin{aligned}
\overline{V}_t^{nt}(\ell,\ell',\omega) &= u(\ell,\ell',\omega) + \delta^3 \mathbb{E}\left[V_{t+1}^{nt}(\ell',\omega')|\omega\right] \\
&= u(\ell,\ell',\omega) + \delta^3 \iint \left\{\overline{V}_{t+1}^{nt}(\ell',j,\omega') + \mathbb{E}\left[\varepsilon_j | d^n = j\right]\right\} dG^n(j|\ell',\omega')dF(\omega'|\omega) \\
&= u(\ell,\ell',\omega) + \delta^3 \iint \Bigg\{\overline{V}_{t+1}^{nt}(\ell',j,\omega') + \gamma \\
&\qquad + \log\left(\sum_{k=1}^{27} \exp\left(\overline{V}_{t+1}^{n}(\ell',k,\omega') - \overline{V}_{t+1}^{n}(\ell',j,\omega')\right)\right)\Bigg\} dG^n(j|\ell',\omega')dF(\omega'|\omega),
\end{aligned}
$$

where $\overline{V}^n$ is the biased alternative specific value function for naive individuals and d^n is the decision made based on this biased view of the world. The expectation operator on the first line of this equation is taken over the joint distribution of the state transitions (F), the stochastic shocks, and the naive choice probabilities (G^n). The second line uses the definition of the alternative-specific value function to rewrite the true value of a naive individual from living in location ℓ' with state ω' as a weighted sum of the true expected value of living in location ℓ' with state ω' and choosing location j, where the weight is given by the naive choice probability.[51] The third line borrows from Kennan (2008) to rewrite the expected value of the logit draw, conditional on the choice made based on the biased alternative-specific value function ($\overline{V}^n$), in the form given above.

[51] Another way to see this is through an application of the law of iterated expectations taken across the joint distribution of the naive choice probabilities and the state transitions.

Table B5: Average First Stage Historical Climate Transitions

	Temperature	Precipitation
Dry		
Intercept	5.854507***	0.297038***
	(1.4242)	(0.0321)
Time Trend	0.002950**	-0.000977***
	(0.0013)	(0.0002)
Lag	0.687234***	0.067172
	(0.0760)	(0.0975)
Temperate		
Intercept	3.361396***	0.641326***
	(0.9733)	(0.0800)
Time Trend	0.003898**	-0.000001
	(0.0016)	(0.0005)
Lag	0.801928***	0.154281
	(0.0574)	(0.1027)
Warm		
Intercept	8.793137***	0.672060***
	(1.8607)	(0.0893)
Time Trend	0.001561	0.000517
	(0.0012)	(0.0006)
Lag	0.624611***	0.245836**
	(0.0794)	(0.0977)

Notes: Temperature is the average of the daily temperature distribution in the relevant climate. Precipitation is the average the total level of agricultural-season precipitation across the locations composing the relevant climate. The time trend is measured in years from 1950. The data used in estimation come from Livneh et al. (2015) for years 1950-1979 and Thornton et al. (2022) for years 1980-2020.

Table B6: Comparisons for Identification

	$\mathbb{1}$\{Migrate\}	Length of Stay
Previous Migrant	0.05262***	-0.64***
	(0.00086)	(0.23)
$\mathbb{1}$\{Origin is Birth Location\}	-0.04401***	-0.53**
	(0.00092)	(0.22)
Predicted Wage	-0.000185***	0.163***
	(0.000049)	(0.017)
$\mathbb{1}$\{Origin is Urban\}	-0.00083	1.50***
	(0.00070)	(0.22)
Age	-0.000929***	-0.108***
	(0.000036)	(0.017)
Degree days above 26°C	0.00418***	0.48**
	(0.00066)	(0.23)
Degree days above 26°C^2	-0.00068***	-0.043
	(0.00013)	(0.048)
Degree Days below 14°C	0.0016***	0.58***
	(0.0004)	(0.13)
Degree Days below 14°C^2	-0.000271***	-0.024
	(0.000046)	(0.015)
Num.Obs.	260472	6012
R2	0.043	0.069

Column 1 represents a linear probability model for individual migration decisions, using the MMP data. Column 2 represents the OLS estimates for the number of years of a given move, recorded in the MMP data, as a function of the characteristics of the first year of that move. *p < 0.1; **p < 0.05; ***p < 0.01

Table B7: Expected Moving Costs

	Expected Cost		
	Average	Minimum	Maximum
Overall	-50,194	-247,082	152,622
Towards Birth Location	29,167	-32,009	152,622
Away From Birth Location	-96,618	-247,082	-33,453
Number of Moves	4,262		

Notes: Costs are measured in 2024 dollars. The expected cost is defined in Appendix B.4. Negative costs represent net positive expected contemporaneous utility.

C Stylized Policy Example

Suppose that we have some individual i who lives for two periods and, without loss, begins in location b with the opportunity to move at the beginning of each period, t. Further, suppose for simplicity that they do not discount the future ($\delta = 1$). They receive utility u^a from living in location a and u^b from living in location b. Assume that moving costs in period t are distributed according to some distribution: $c_t \sim F(\cdot)$, and that they make their period t moving decision after observing c_t, but before receiving flow utility for period t. We can solve for the period 1 decision using backward recursion. In particular, notice that if they find themselves in location b at the start of $t = 2$, they will move to location a if:

$$u^a - c_2 > u^b, \tag{12}$$

and if they find themselves in location a at the beginning of period 2, they will move back to location b if:

$$u^b - c_2 > u^a, \tag{13}$$

We can use Inequality 12 and Inequality 13 to solve for their period 1 choice. In particular, i will move to location a in period 1 if:

$$\begin{aligned} &u^a - c_1 + \overbrace{u^a[1 - F(u^b - u^a)] + (u^b - E[c_2])F(u^b - u^a)}^{E[V(a)]} > \\ &u^b + \underbrace{u^b(1 - F(u^a - u^b)) + (u^a - E[c_2])F(u^a - u^b)}_{E[V(b)]}. \end{aligned} \tag{14}$$

Where $E[V(\cdot)]$ represents the expected value of beginning period 2 in location $\cdot$. Inequality 14 holds when:

$$c_1 < (u^a + E[V(a)]) - \left(u^b + E[V(b)]\right).$$

Or when the current period moving cost is less than the gain from moving. We can use this to define the ex ante probability that i moves in the first period from location b to location a as:

$$p(m_1) = \int_{-\infty}^{(u^a + E[V(a)]) - \left(u^b + E[V(b)]\right)} dF.$$

Let $z = (u^a + E[V(a)]) - \left(u^b + E[V(b)]\right)$. Notice that we can write the expected number of moves as:

$$\underbrace{F(z)}_{\text{Period 1 moves}} \quad + \underbrace{F(u^a - u^b)\left[1 - F(z)\right]}_{\text{Period 2 moves from } b} + \underbrace{F(u^b - u^a)F(z)}_{\text{Period 2 moves from } a}$$
$$= F(z) + \underbrace{F(u^a - u^b) + F(z)\left[F(u^b - u^a) - F(u^a - u^b)\right]}_{\text{Period 2 moves}}. \tag{15}$$

Where m_t represents movers in period t.

Static Policy in Period One

Imagine now that there is a static reduction in moving costs for the first period of $s > 0$. This only shifts the left hand side of Equation 14 to the right by s. In particular, they now move to a in period 1 if:

$$c_1 < (u^a + E[V(a)]) - \Big(u^b + E[V(b)]\Big) + s,$$

which gives the ex ante probability that i moves in the first period as:

$$\begin{aligned} p(m_1|s_1) &= \int_{-\infty}^{(u^a+E[V(a)])-\left(u^b+E[V(b)]\right)+s} dF \\ &= \underbrace{\int_{-\infty}^{(u^a+E[V(a)])-\left(u^b+E[V(b)]\right)} dF}_{\text{Inframarginal movers}} + \underbrace{\int_{(u^a+E[V(a)])-\left(u^b+E[V(b)]\right)}^{(u^a+E[V(a)])-\left(u^b+E[V(b)]\right)+s} dF}_{\text{Marginal movers}}. \end{aligned}$$

We can write the expected number of moves as:

$$\underbrace{F(z+s)}_{\text{Period 1 moves}} \quad + \underbrace{F(u^a - u^b)\left[1 - F(z+s)\right]}_{\text{Period 2 moves from } b} + \underbrace{F(u^b - u^a)F(z+s)}_{\text{Period 2 moves from } a}$$
$$= F(z+s) + \underbrace{F(u^a - u^b) + F(z+s)\left[F(u^b - u^a) - F(u^a - u^b)\right]}_{\text{Period 2 moves}}. \tag{16}$$

Given that $s > 0$, the static policy will induce weakly more moves than the base setting with no policy intervention. Moreover, we can see from the second part of Equation 16 that the static policy detracts from the overall number of second period migrants from b when $F(z+s) > F(z)$. That said, it also increases the number of second period migrants back to b, by increasing the number of individuals in a. When $u^b > u^a$, the second channel overpowers the first: the static subsidy will induce weakly more second period migrants than the baseline. When $u^a > u^b$, the difference in second period migrants between the static policy and the baseline is ambiguously signed.

Static Policy in Period Two

Now suppose instead that there is a static reduction in *second* period moving costs of $s > 0$. i's expected value of moving to location a in period 1 is given by:

$$u^a + \underbrace{u^a[1 - F(u^b - u^a + s)] + (u^b - E[c_2] + s)F(u^b - u^a + s)}_{E[V(a)|s_2]} - c_1,$$

and of staying in location b by:

$$u^b + \underbrace{u^b(1 - F(u^a - u^b + s)) + (u^a - E[c_2] + s)F(u^a - u^b + s)}_{E[V(b)|s_2]}.$$

They will move to location a at the beginning of the first period if:

$$\begin{aligned}
c_1 < {} & (u^a + E[V(a)|s_2]) - \left(u^b + E[V(b)|s_2]\right) \\
= {} & (u^a + u^a[1 - [F(u^b - u^a) + p(u^b - u^a \le c_2 \le u^b - u^a + s)] \\
& + (u^b - E[c_2] + s)[F(u^b - u^a) + p(u^b - u^a \le c_2 \le u^b - u^a + s)])) \\
& - (u^b + u^b(1 - [F(u^a - u^b) + p(u^a - u^b \le c_2 \le u^a - u^b + s)]) \\
& + (u^a - E[c_2])[F(u^a - u^b) + p(u^a - u^b \le c_2 \le u^a - u^b + s)] + sF(u^a - u^b + s)) \\
= {} & (u^a + u^a[1 - F(u^b - u^a)] + (u^b - E[c_2])F(u^b - u^a)) \\
& - (u^b + u^b[1 - F(u^a - u^b)] + (u^a - E[c_2])F(u^a - u^b)) \\
& + (-u^a p(u^b - u^a \le c_2 \le u^b - u^a + s) \\
& + (u^b - E[c_2])p(u^b - u^a \le c_2 \le u^b - u^a + s) + sF(u^b - u^a + s)) \\
& - (-u^b p(u^a - u^b \le c_2 \le u^a - u^b + s) \\
& + (u^a - E[c_2])p(u^a - u^b \le c_2 \le u^a - u^b + s) + sF(u^a - u^b + s)) \\
= {} & \underbrace{(u^a + E[V(a)]) - (u^b + E[V(b)])}_{\text{Inframarginal movers}} \\
& + \underbrace{((u^b - u^a - E[c_2])p(u^b - u^a \le c_2 \le u^b - u^a + s) + sF(u^b - u^a + s))}_{\text{Increase in period 1 movers, from subsidy to move back}} \\
& - \underbrace{((u^a - u^b - E[c_2])p(u^a - u^b \le c_2 \le u^a - u^b + s) + sF(u^a - u^b + s))}_{\text{Reduction in period 1 movers, from subsidy in period 2}}.
\end{aligned}$$

This highlights that there are two countervailing forces at play with a static policy in the second period. In particular, some individuals will decide to move in the first period because they know they can move back at a reduced cost. On the other hand, some individuals will decide not to move in the first period because they would prefer to wait until the following period to move to a when the subsidy is available.

One-Time Dynamic Policy

Now instead suppose that there is a dynamic subsidy which is only available to be used at the first time an individual decides to move. In this case, i's expected value from moving to a in the first period is given by:

$$u^a + \underbrace{u^a[1 - F(u^b - u^a)] + (u^b - E[c_2])F(u^b - u^a)}_{E[V(a)]} - c_1 + s,$$

and their expected value from staying in location b is given by:

$$u^b + \underbrace{u^b(1 - F(u^a - u^b + s)) + (u^a - E[c_2] + s)F(u^a - u^b + s)}_{E[V(b)|s_2]}.$$

Notice that the expected value of moving to location a in the first period is the same as that of the static offering in the first period, while their expected value of staying in location b is the same as that of the static offering in the second period. This means that they will move when:

$$c_1 < (u^a + E[V(a)]) - \left(u^b + E[V(b)|s_2]\right) + s,$$

or when:

$$c_1 < \underbrace{(u^a + E[V(a)]) - (u^b + E[V(b)]) + s}_{\text{Movers under static policy in Period 1}}$$
$$- \underbrace{((u^a - u^b - E[c_2])p(u^a - u^b \leq c_2 \leq u^a - u^b + s) + sF(u^a - u^b + s))}_{\text{Change in period 1 movers, from subsidy in period 2}}.$$

The ex ante probability of i choosing to move in the first period under this policy is therefore given by:

$$p(m_1|s_{d1}) = \underbrace{\int_{-\infty}^{z} dF}_{\substack{\text{Inframarginal} \\ \text{movers}}} + \underbrace{\int_{z}^{z+s} dF}_{\substack{\text{Marginal movers} \\ \text{under static policy}}} + \underbrace{\int_{z+s}^{z+s-w} dF}_{\substack{\text{Change in movers} \\ \text{under dynamic policy}}},$$

where $z = (u^a + E[V(a)]) - (u^b + E[V(b)])$ and $w = ((u^a - u^b - E[c_2])p(u^a - u^b \leq c_2 \leq u^a - u^b + s) + sF(u^a - u^b + s))$. Note the following:

$$w > 0 \Leftrightarrow \underbrace{(u^a - u^b - E[c_2])}_{\substack{\text{Expected gain (or loss)}\\ \text{from period 2 move}}} \underbrace{\left[F(u^a - u^b + s) - F(u^a - u^b)\right]}_{\substack{\Delta \text{ in the probability}\\ \text{of moving because of the subsidy}}} + \underbrace{sF(u^a - u^b + s)}_{\substack{\text{Expected value of}\\ \text{the subsidy}}} > 0. \tag{17}$$

When the value of moving is larger than the expected cost, a dynamic policy will always induce fewer individuals to move in the first period. This is because some individuals who received a poor cost draw understand that their next period draw will likely be more favorable. How many individuals find themselves in this position depends both on how strongly the subsidy shifts their behavior and the size of the gain from migration.

Determining the sign of w is more nuanced when the value of moving is *smaller* than the expected cost. Suppose that the expected change in utility from moving in period 2 ($E[\Delta u]$) is less than 0. Notice that w is positive when:[52]

$$\begin{aligned}
E[\Delta u]\,[p(m) - p(m|s)] &< sp(m|s) \\
\frac{p(m) - p(m|s)}{p(m|s)} &> \frac{s}{E[\Delta u]} \\
\underbrace{\text{abs}\,(\%\Delta p(m|s))}_{\substack{\text{Percentage change}\\ \text{in period 2 movers}\\ \text{from losing the subsidy}}} &< \underbrace{\text{abs}\left(\frac{s}{E[\Delta u]}\right)}_{\substack{\text{Size of subsidy, relative to}\\ \text{expected loss from moving}}} .
\end{aligned}$$

That is, w is positive when the shift in the probability of moving in the second period, as brought about by removing subsidy, is less than the value of the subsidy relative to the expected loss from moving. For a sufficiently small subsidy, this highlights that the slope of the density of moving costs at the endpoint of the inframarginal individuals drives whether a dynamic subsidy will induce more or less individuals, than a static policy, to move in the first period. This will also determine whether a dynamic policy will induce positive or negative selection as compared to a static policy.

We can write the expected number of moves as:

$$\begin{aligned}
&\underbrace{F(z+s-w)}_{\text{Period 1 moves}} + \underbrace{F(u^a - u^b)\,[1 - F(z+s-w)]}_{\text{Period 2 moves from } b} + \underbrace{F(u^b - u^a)F(z+s-w)}_{\text{Period 2 moves from } a} \\
&= F(z+s-w) + \underbrace{F(u^a - u^b) + F(z+s-w)\left[F(u^b - u^a) - F(u^a - u^b)\right]}_{\text{Period 2 moves}}
\end{aligned} \tag{18}$$

[52]The first line only rewrites 17. The second line uses the fact that $E[\Delta u] < 0$ when flipping the inequality. The third line multiplies each side by -1 to make both sides positive.

Notice that $w > 0$ implies that the dynamic policy will reduce the number of first period movers and the number of second period movers from a back to b. Mathematically, this is because $F(z+s-w)\left[1+F(u^b-u^a)\right] \leq F(z+s)\left[1+F(u^b-u^a)\right]$. Intuitively, this is because the dynamic policy induces some individuals who would have moved from b to a in period 1 under the static policy to instead spend their period in location b in the hopes that next period's draw will be better. That said, these individuals are not entirely giving up on the prospect of moving to a - they are only delaying the expected move. This means that the dynamic policy will induce *more* individuals to move to a in the second period, as compared to the static offering. Mathematically, this enters through the fact that $F(u^a-u^b)[1-F(z+s-w)] \geq F(u^a-u^b)[1-F(z+s)]$, for $w > 0$. Which of these two forces prevails depends on the particular distribution from which moving costs are drawn and the gain from migration.

Further, notice that the dynamic subsidy will weakly generate higher welfare than the first-period static policy. To see why this is, notice that the dynamic policy can be broken apart into two distinct policies. In the first period, individuals can take exactly the static policy, however they further have the option of choosing the static policy in the second period (albeit only from a starting position of b, rather than a). This ability to choose generates option value, in itself. Moreover, when $w > 0$, the reduction in the fraction of the population marginal to moving gives the selection results I highlight in the main text. In particular, the individuals who decide not to move under the dynamic policy are those who were marginal to the policy, which means that they received a relatively high cost draw in the current period, which is why they required the subsidy to induce their move. The differential effects of selection into the policy is what drives the large differences in the welfare outcomes of movers under the different dynamic structures of policy.

Finally, notice that w is exactly the option value of the dynamic subsidy. To see why this is true, I highlight that the value of having the option to migrate in the second period is given by:

$$\underbrace{F(u^a-u^b+s(p))}_{\substack{\text{Probability of moving}\\ \text{in period 2}}}\underbrace{[u^a-u^b-E[c_2]+s(p)]}_{\substack{\text{Value of moving}\\ \text{in period 2}}},$$

where $s(p)$ is equal to s for the dynamic policy and 0 for the static policy and the baseline. The option value of the dynamic subsidy is then the difference between the option value of moving in period two with the subsidy from the option value of moving in period 2 without the subsidy.

Mathematically:

$$
\begin{aligned}
&F(u^a - u^b + s)[u^a - u^b - E[c_2] + s] - F(u^a - u^b)[u^a - u^b - E[c_2]] \\
=&[F(u^a - u^b + s) - F(u^a - u^b)][u^a - u^b - E[c_2]] + sF(u^a - u^b + s) \\
=&w.
\end{aligned}
$$

This tells us that the dynamic policy will induce positive selection from those marginal to moving exactly when the option value of moving tomorrow, as brought about by the subsidy, is positive. The magnitude of this selection is increasing with the magnitude of the option value from the subsidy. Notice that we can very simply move away from a two period model and retain similar intuition to the above. In particular, we can generalize the above where now the value of moving to a in period t is given by $u^a + EV_a^{t+1}$ and the value of staying in b is given by $u^b + EV_b^{t+1}$. The value of having the option to move in the following period is therefore given by:

$$
F((u^a + EV_a^{t+2}) - (u^b + EV_b^{t+2}) + s(p))[(u^a + EV_a^{t+2}) - (u^b + EV_b^{t+2}) - E[c_2] + s(p)].
$$

The remainder of the calculations follow through as given above.

Fully Dynamic Policy

For completeness, now suppose instead that there is a policy which reduces one's moving costs by $s > 0$ in both periods. Under this scenario, i values moving to a in period 1 at:

$$
u^a + \underbrace{u^a[1 - F(u^b - u^a + s)] + (u^b - E[c_2] + s)F(u^b - u^a + s)}_{E[V(a)|s_2]} - c_1 + s,
$$

and of staying in b in period 1 at:

$$
u^b + \underbrace{u^b(1 - F(u^a - u^b + s)) + (u^a - E[c_2] + s)F(u^a - u^b + s)}_{E[V(b)|s_2]}.
$$

Notice that the expected value of moving to a in the first period is the same as that of the static policy in the second period, net of the subsidy receieved from moving. Moreover, the value of staying in b is exactly the same as that of the static policy in the second period. We can use this fact to derive the cutoff to move to a in the first period as:

$$\begin{aligned}
c_1 &< (u^a + E[V(a)|s_2]) - \left(u^b + E[V(b)|s_2]\right) + s \\
&= \underbrace{(u^a + E[V(a)]) - (u^b + E[V(b)]) + s}_{\text{Marginal movers w/ static policy}} \\
&\quad + \underbrace{((u_b - u_a - E[c_2])p(u^b - u^a \leq c_2 \leq u^b - u^a + s) + sF(u^b - u^a + s))}_{\text{Change in period 1 movers, from subsidy to move back}} \\
&\quad - \underbrace{((u^a - u_b - E[c_2])p(u^a - u^b \leq c_2 \leq u^a - u^b + s) + sF(u^a - u^b + s))}_{\text{Change in period 1 movers, from subsidy in period 2}}.
\end{aligned}$$

The ex ante probability of i deciding to move in the first period is therefore given by:

$$p(m_1|s_d) = \underbrace{\int_{-\infty}^{z} dF}_{\star_1} + \underbrace{\int_{z}^{z+s} dF}_{\star_2} + \underbrace{\int_{z+s}^{z+s+v} dF}_{\star_3} + \underbrace{\int_{z+s+v}^{z+s+v-w} dF}_{\star_4},$$

where $\star_1$ are the inframarginal movers, $\star_2$ are the movers induced by the static first-period subsidy, $\star_3$ represents the movers induced to move by the prospect of a reduced return moving cost, and $\star_4$ represent individuals who prefer to wait until the following period to potentially move, $z = (u^a+E[V(a)])-(u^b+E[V(b)])$, $v = (u_b-u_a-E[c_2])p(u^b-u^a \leq c_2 \leq u^b-u^a+s)+sF(u^b-u^a+s)$, and $w = (u^a - u_b - E[c_2])p(u^a - u^b \leq c_2 \leq u^a - u^b + s) + sF(u^a - u^b + s)$.

D Robustness & Additional Results

D.1 Additional Model Fit Statistics

Table D1 shows the data and model-predicted moments discussed in the main text, Figures D1 and D2 replicate the relevant figures in the main text, when simulating the entire life-history of each individual in the estimation data.

D.2 Additional Results

Table D4 presents the estimated preference and moving cost parameters when allowing for a different moving cost for agricultural workers. Figure D3 present the value of migration for 17–19-year-olds, broken apart by climate, urban/rural location of birth, education level, and sector of employment. Figure D4 presents the climate damages for the same group, broken apart by the same characteristics.

Table D1: Model Fit - Full Lifecycle

	Migration Rate	
	Model	Data
Overall	4.1%	4.7%
Children		
Yes	3.9%	3.4%
No	4.1%	7.2%
Agricultural Worker		
Yes	3.5%	3.5%
No	4.3%	5.3%
Years of Education		
0-5	3.8%	4.2%
6-11	4.0%	3.8%
12+	4.6%	7.5%

Note: For each category, I calculate the average migration predicted by 100 model simulations of the life trajectories of the individual-year observations used in estimation. The data migration rates are calculated as the average migration rate for each category.

What is driving the large differences in the value of migration across the two climate pathways? The model incorporates two primary channels through which climate can affect welfare: directly through the flow utility of temperature as an amenity and indirectly through local wages. I use the model to understand how much of the reported heterogeneous values are driven by the separate channels. To do so, I individually turn on and off the wage and amenity channels along with the ability to migrate. For each of these counterfactuals, I resimulate behavior under the business-as-usual climate scenario and present the results in Table D3 in Appendix D.2. The first column (Only Wages) looks at the value of migration when the amenity value of weather is turned off, i.e., by setting θ_2–θ_6 equal to zero. I find that the resulting option values are much more homogeneous across space and quite similar to those reported for the no climate change scenario. The second column (Only Amenities) instead turns off the wage impact in flow utilities, i.e., by setting θ_1 equal to 0. I find values that are much closer to those in the business-as-usual scenario. This result highlights the high premium that individuals place on the amenities of their location and reaffirms past findings using data from the United States (e.g., Kennan and Walker, 2011; Ransom, 2022). Moreover, I find that it is through amenities, rather than wages, that the heterogeneous value of migration is most noticeable. This outcome is likely driven by past sorting behavior of non-agricultural workers and those with higher levels of education to areas less affected by climate change.

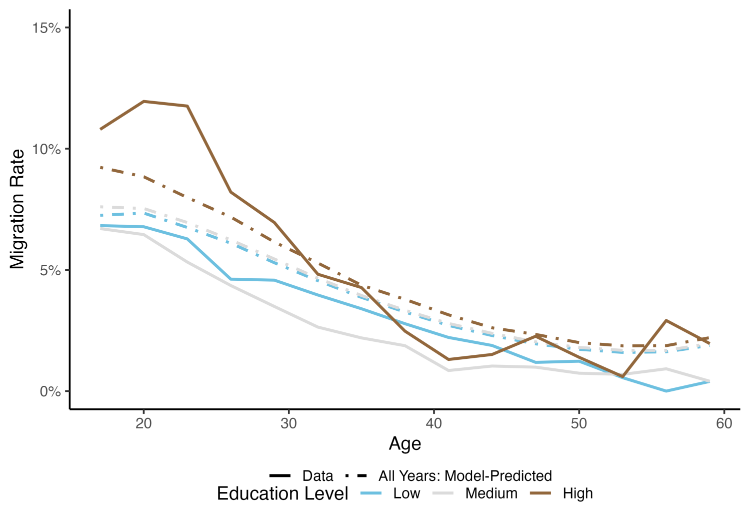


Figure D1: Model-Predicted Migration Rates by Education
Notes: This figure plots the average migration rates in the data against those predicted by the model, broken apart by education levels. It represents the model-simulations for the entire lifecycle for all individuals used in estimation. Low, mid, and high-education refers to less than 6 years, 6–11 years, and 12 or more years of education, respectively

D.2.1 Additional Policy Results

I discuss here the different effects on migration across the age distribution from the dynamic structure of policy-induced reductions in moving costs. I study the effectiveness of a 1% reduction in moving costs, broken apart across the age distribution. Figure D9 plots the results of this exercise. The solid curve represents the change in the ex ante probability than an individual migrates in 2020 from a permanent 1% reduction in moving costs. The dashed curve plots the same for a one-time 1% reduction in moving costs; that is, a policy which reduces moving costs by 1% in 2020, but for which all future costs would return to the baseline levels.

I find that an individual's relative responsiveness to the two schemes changes throughout their life-cycle. This is because of two conflicting forces. On the one hand, individuals are more likely to move when they know that they will be able to leave their destination at a reduced cost in the future. On the other hand, individuals understand that the cost of future moves will be higher than they are today, simply because they will be older in the future than they are today. Under the one-time subsidy, moving costs will be even higher in the future, relative to the year of the subsidy. The same is not true for the permanent subsidy. If individuals are planning on making a move at some point in their lives, this would lead to a higher proportion of migrants under the one-time subsidy than under the permanent subsidy. At the beginning of an individual's life, the first force dominates the second, leading to a higher rate of migration under the permanent subsidy. However, as individuals age, the probability that they will make more than one additional move in their lifetime falls, as signified by the decreasing difference between the two curves. Eventually, the second force dominates the first,

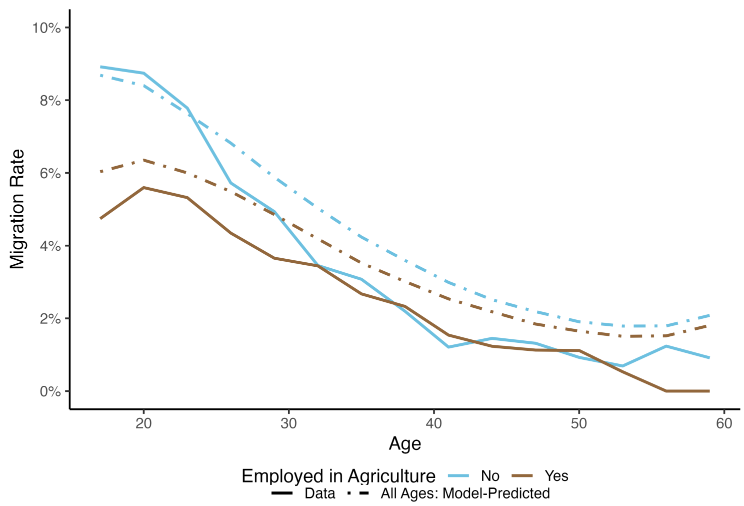


Figure D2: Model-Predicted Migration Rates by Sector of Employment
Notes: This figure plots the average migration rates in the data against those predicted by the model, broken apart by sector of employment. It represents the model-simulations for the entire lifecycle for all individuals used in estimation. Low, mid, and high-education refers to less than 6 years, 6–11 years, and 12 or more years of education, respectively

leading to a higher rate of migration under the one-time subsidy.

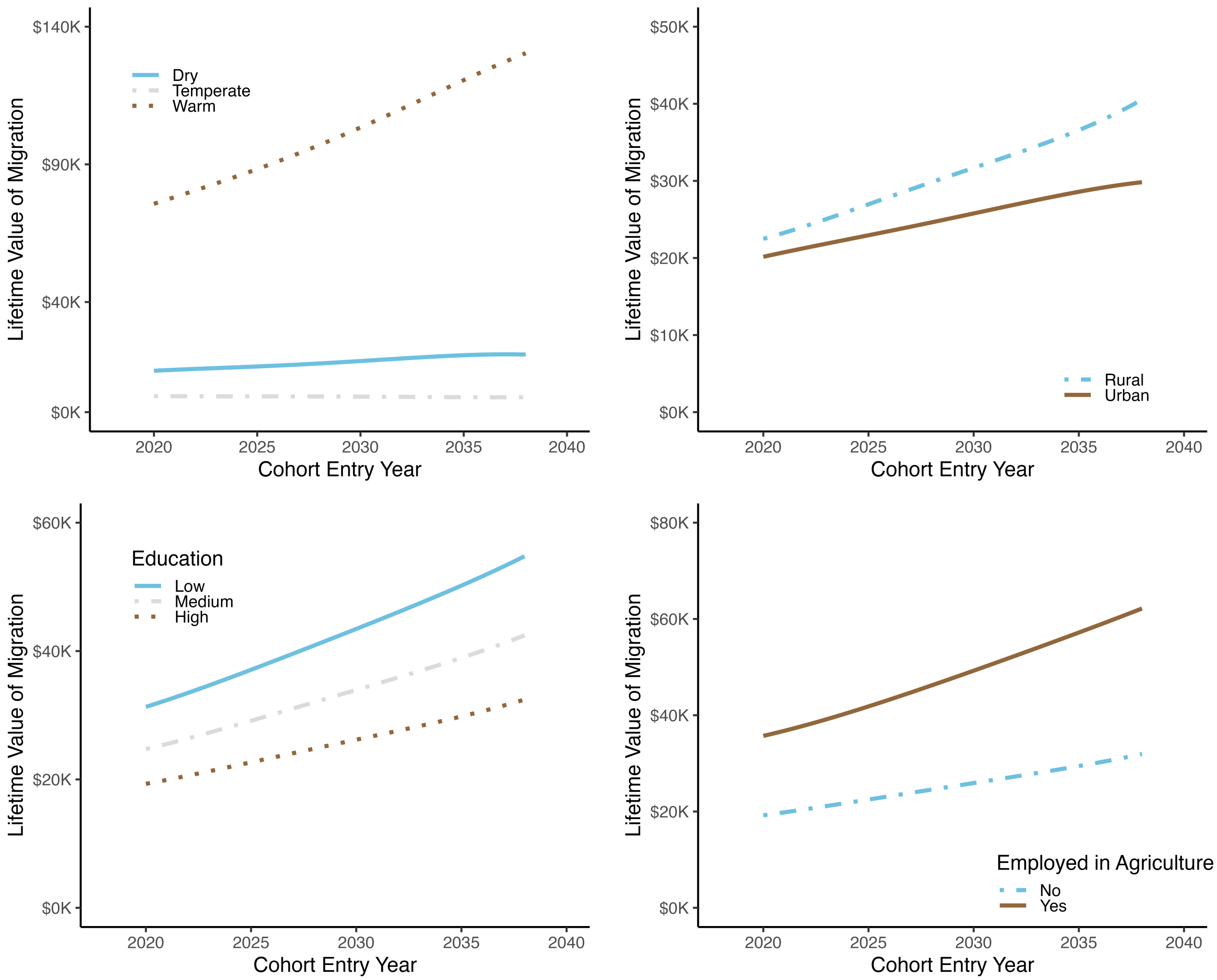


Figure D3: This figure plots the average value of migration for 17–19-year-olds, from 2020 through 2038. (a) breaks this value apart by climate of birth, (b) by urban/rural status of the location of birth, (c) by education level, and (d) by sector of employment

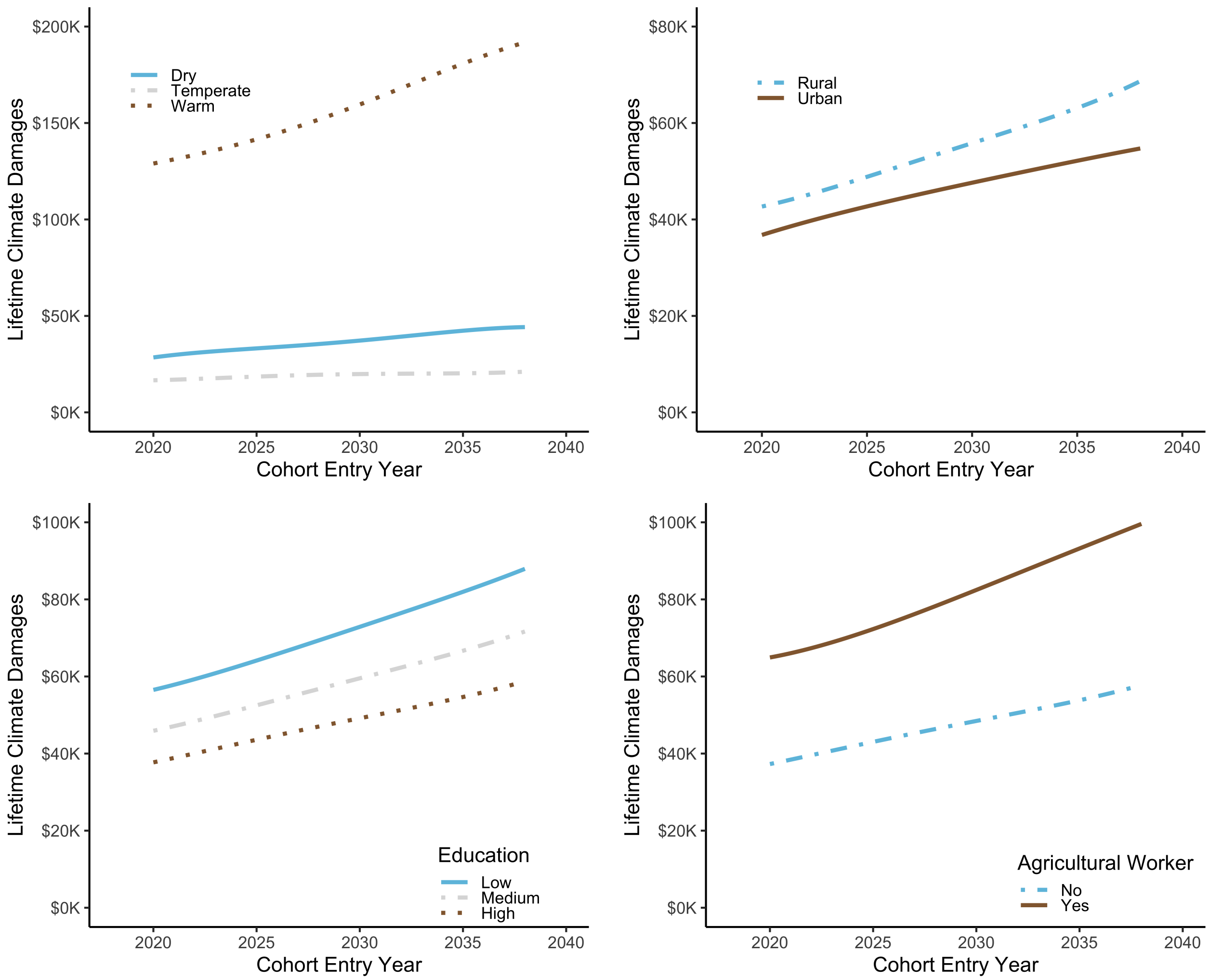


Figure D4: This figure plots the average level of climate damages for 17–19-year-olds, from 2020 through 2038. (a) breaks this value apart by climate of birth, (b) by urban/rural status of the location of birth, (c) by education level, and (d) by sector of employment

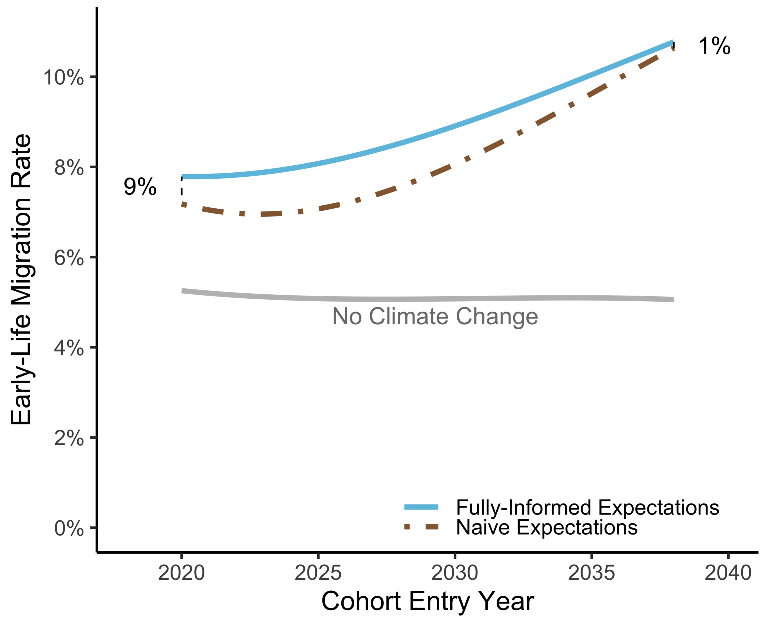


Figure D5: Migration rates across time and beliefs
Notes: This figure plots model-predicted migration rates, for 17–19-year-olds by time. The solid curve represents the migration rate for fully-informed individuals; the dashed curve for climate-naive individuals, and the light solid curve the average migration rate in the limited climate change scenario. The results are averaged over 100 simulations.

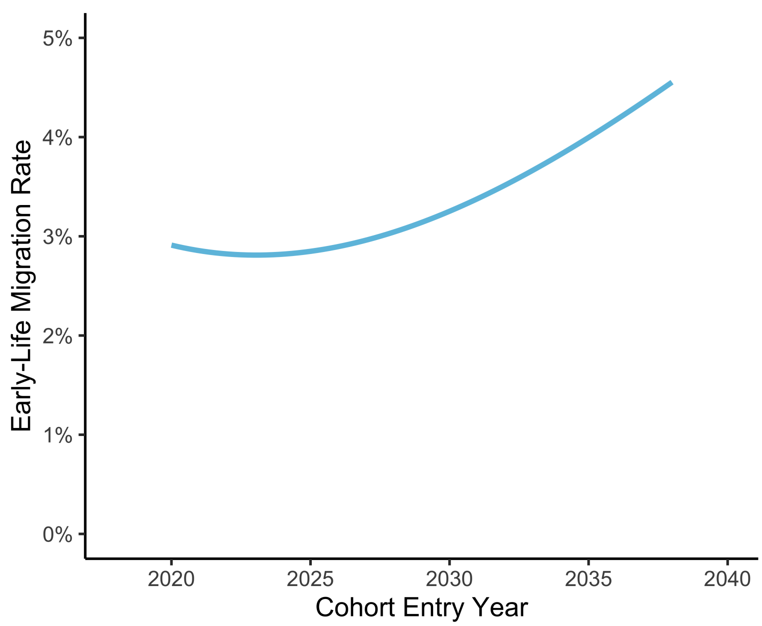


Figure D6: Migration rates from static logit
Notes: This figure plots the static logit model-predicted migration rates, for 17–19-year-olds by time, under a business-as-usual warming scenario. The results are averaged over 100 simulations.

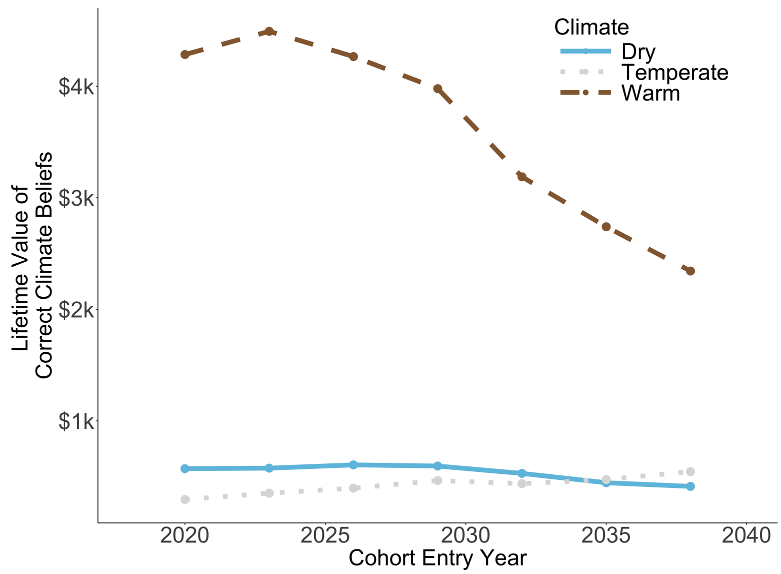


Figure D7: Value of information to climate-naive

Notes: This figure plots the value of becoming fully-informed, to the climate-naive, over time and space. The solid curve represents the value to those living in dry location, the dashed curve in warm locations, and the dotted curve in temperate locations. Results are averaged over 100 simulations.

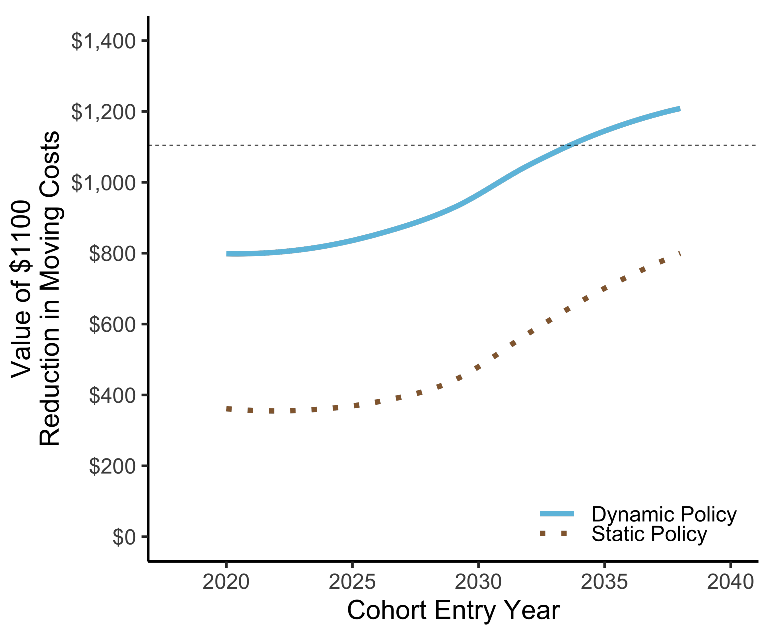


Figure D8: Value of migration subsidy, across time and policy-structure

Notes: This figure plots the value, to 17–19-year-old movers, of $1,100 migration subsidies, broken apart by the dynamic structure of the policy. The solid curve represents the value of a dynamic policy which is used at the first time an individual decides to move. The dotted curve represents the value of a static policy which is only available in the given year. The dashed line represents the level of the subsidy. Results are averaged over 100 simulations.

Table D2: Static Logit Parameters: Forward Looking Evidence

Flow Utility		Moving Costs	
Income	0.063	Moving Intercept	2.911
	(0.006)		(0.095)
Degree Days above 26°C	0.023	Distance	0.136
	(0.042)		(0.014)
Degree Days above 26°C^2	-0.070	Children	0.128
	(0.014)		(0.044)
Degree Days below 14°C	0.148	Age	0.209
	(0.014)		(0.009)
Degree Days below 14°C^2	-0.024		
	(0.002)		
Average Temperature Next Decade	0.019		
	(0.006)		
Urban Location	-0.349		
	(0.027)		
Birth Location	2.450		
	(0.022)		
Log-Likelihood: -25255.83			

Note: Income is measured in annual hourly 2010 pesos, degree days are divided by 100, distance is measured in log kilometers. Average temperature over the next decade refers to the mean of the daily temperature distribution at the level-one Köppen climate level. Future temperatures are drawn from the historical data through 2020 and the business-as-usual simulated weather data for 2020–2030. Standard errors are computed with the inverse Hessian.

Table D3: Supplemental Values of Migration

	Value of Migration		
	Only Wages	Only Amenities	General Equilibrium
Average	5,773	21,271	19,205
Climate			
Dry	4,947	16,587	13,660
Temperate	6,306	5,909	5,363
Warm	5,336	73,963	68,544
Agricultural Worker			
Yes	5,187	34,604	32,652
No	5,851	19,504	17,423
Years of Education			
0–5	5,501	30,530	29,527
6–11	5,682	24,649	23,136
12+	5,821	19,555	17,232

Note: These are the average results of 100 simulations of the model. Values are for 17-19 year olds forming fully-informed climate expectations, in 2020, measured in 2024 dollars. The value of migration is defined as the difference between average lifetime welfare for an individual who can migrate and one who cannot. All values are for a business-as-usual climate scenario: the SSP2-4.5 climate projections. General equilibrium refers to the business-as-usual scenario with a congestion cost in the income equation, iterated to convergence. The amenity impact is from a counterfactual where weather does not impact wages. The wage impact is from a counterfactual where weather does not impact amenities.

Table D4: Utility Parameters, Agricultural Movers

Flow Utility		Moving Costs	
Income	0.006	Moving Intercept	2.781
	(0.0007)		(0.074)
Degree Days above 26°C	0.023	Distance	0.109
	(0.005)		(0.011)
Degree Days above 26°C^2	-0.010	Children	0.091
	(0.002)		(0.030)
Degree Days below 14°C	0.021	Age	0.172
	(0.002)		(0.006)
Degree Days below 14°C	-0.004	Agricultural Worker	0.078
	(0.0003)		(0.027)
Urban Location	-0.041		
	(0.003)		
Birth Location	0.179		
	(0.002)		
Mass of Movers	0.627		
	(0.013)		
Log-Likelihood: -24866.35			

Note: Income is measured in annual hourly 2010 pesos. Degree days are divided by 100, and measured in the destination location. Distance is measured in log kilometers. The likelihood contains 90,578 individual-year observations from 11,194 individuals. Standard errors in parentheses.

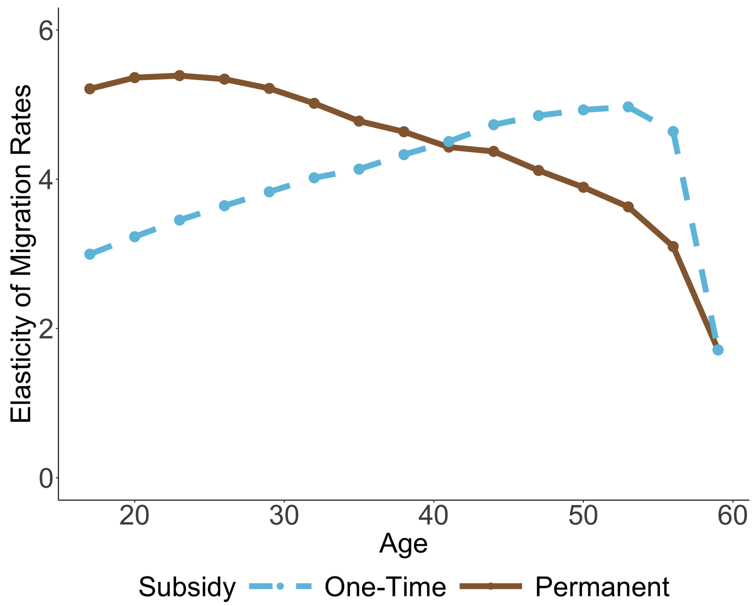


Figure D9: Responsiveness of Migration Rates to Migration Subsidies
Notes: This figure plots the percent increase in migration rates from a 1% decrease in migration costs, across the age distribution and duration of the subsidy. The solid curve represents a permanent subsidy, while the dashed curve represents a one-time subsidy.

E Extensions

E.1 Within-Location Across-Municipality Moves

Table E1: Utility Parameters: Within Location Moves

Flow Utility		Moving Costs	
Income	0.006	Moving Intercept	3.185
	(0.0006)		(0.066)
Degree Days above 26°C	0.059	Within Location	4.486
	(0.005)		(0.025)
Degree Days above 26°C^2	-0.023	Distance	0.166
	(0.002)		(0.010)
Degree Days below 14°C	0.027	Children	0.375
	(0.002)		(0.283)
Degree Days below 14°C^2	-0.004	Age	0.226
	(0.0003)		(0.050)
Urban Location	-0.043		
	(0.003)		
Birth Location	0.214		
	(0.002)		
Mass of Movers	0.641		
	(0.013)		
Mass of Believers	0.766		
	(0.176)		
Log-Likelihood: -30858.60			

Note: Income is measured in annual hourly 2010 pesos. Degree days are divided by 100 and measured in the destination location. Distance is measured in log kilometers. The likelihood contains 90,578 individual-year observations from 11,194 individuals. Standard errors in parentheses.

Table E1 presents the estimated preference and moving cost parameters when additionally modeling the decision to move across municipalities but within the same location - that is, individuals choose over 28 options: 26 across location moves, 1 within location move, and the option to remain in the same municipality-location. In particular, I modify the baseline moving costs as follows:

$$\tilde{c}(\ell, \ell', \omega) = \tilde{\gamma}_{1\omega} + \tilde{\gamma}_2 \mathbb{1}\{\ell = \ell' \wedge \text{move}\} + \tilde{\gamma}_3 d(\ell, \ell') + \tilde{\gamma}_4 \mathbb{1}\{k_\omega \geq 1\} + \tilde{\gamma}_5 \text{age}_\omega,$$

where, now, $\tilde{\gamma}_2$ represents the within location moving cost. $\tilde{\gamma}_3$ is identified only from moves across locations, and $\tilde{\gamma}_4$ and $\tilde{\gamma}_5$ are identified from both within and across location moves. The estimated moving costs are higher than in the baseline specification, but the preference parameters are quite similar. The within location moving cost is quite high. This is because most moves are long-distance in the estimation data.

E.2 Intensification Effects

The model discussed in the primary section of this paper assumes that migration frictions in the future will resemble those of today. This is an assumption that is ubiquitous in the literature. Past work has pointed out that the role that migration (or, more broadly, adaptation) plays in limiting future climate damages will depend on whether intensification effects of climate change outweigh technological progress related to migration costs. In this section, I present a modified version of the model to account for these conflicting forces.

Throughout the duration of the MMP survey used in estimation, 1960s-2010s, Mexico experienced significant technological change as well as a noticeable rightward shift in its climate distribution. I modify the moving cost function to account for potential differences in the moving cost arising because of differences in local agricultural weather (Cattaneo and Peri, 2016). The baseline model remains identical to that presented in Section 3, with the exception that now the moving cost additionally is allowed to depend on a quadratic in the level of local growing degree days in the origin. Mathematically,

$$\bar{c}(\ell, \ell', \omega) = \gamma_{1\omega} + \gamma_2 GDD(\ell, \omega) + \gamma_3 GDD(\ell, \omega)^2 + \gamma_4 d(\ell, \ell') + \gamma_5 \mathbb{1}\{k_\omega \geq 1\} + \gamma_6 \text{age}_\omega,$$

Identification of γ_2 and γ_3 relies on differences in migration rates away from areas with different levels of growing degree days and on the choices of where individuals move to, acknowledging that future moving costs will be a function of expected future growing degree days in the destination today. I remain agnostic as to the sign and magnitude of the linear and quadratic terms in the moving cost specification. In particular, one could imagine that a good agricultural harvest, as proxied for by growing degree days, could lead to either increased migration because of increased incomes (e.g., Liu et al., 2023) or decreased migration due to a lower need to supplement current incomes. Given that liquidity constraints enter into the model through the moving cost, it is also possible that sufficiently poor agricultural outcomes, again, as proxied for by growing degree days may be associated with lower migration rates.

Table E2 presents the estimated preference parameters from this updated model specification. Similar to the baseline model for the amenity values of extreme heat, this model allows for a bliss point interpretation of the coefficients for growing degree days in the moving costs. The model estimated bliss point is 2,463 growing degree days, with a marginal degree day below this value

Table E2: Utility Parameters, Intensification

Flow Utility		Moving Costs	
Income	0.006	Moving Intercept	3.018
	(0.0007)		(0.118)
Degree Days above 26°C	0.040	Growing Degree Days	-0.202
	(0.005)		(0.084)
Degree Days above 26°C^2	-0.016	Growing Degree Days^2	0.041
	(0.002)		(0.016)
Degree Days below 14°C	0.023	Distance	0.111
	(0.002)		(0.011)
Degree Days below 14°C^2	-0.004	Children	0.072
	(0.0003)		(0.028)
Urban Location	-0.040	Age	0.172
	(0.003)		(0.006)
Birth Location	0.183		
	(0.002)		
Mass of Movers	0.629		
	(0.013)		
Log-Likelihood: -24859.345			

Note: Income is measured in annual hourly 2010 pesos. Degree days are divided by 100 and measured in the destination location. Growing degree days have a lower discontinuity at 8 degrees Celsius and an upper discontinuity at 32 degrees Celsius, are divided by 1000, and are measured in the origin location. Distance is measured in log kilometers. The likelihood contains 90,578 individual-year observations from 11,194 individuals. Standard errors in parentheses.

reducing the migration cost and one above it increasing this cost. 40% of the data used in estimation experienced growing degree days above this threshold; 30% of those living in dry locations, 15% in temperate locations, and 86% in warm locations. If these same individuals were transplanted to 2050, the share of individuals living in dry locations who experience growing degree days beyond this threshold would increase to 97%, the share living in temperate locations to 45%, and the share living in warm locations to 100%.

I find that accounting for this mechanism leads to a 10% increase in lifetime climate damages experienced by the average 17–19-year-old in 2020. Those living in temperate climates stand to relatively lose the most - with an increase in lifetime damages of 27%, while those living in warm climates experience an 8% increase, and those living in dry climates experience an 5% increase. Individuals living in temperate climates face the highest increase to climate damages because of their destinations. Particularly, some of those living in these locations decide to move to locations in the dry and warm regions of the country, which experience the brunt of climate change. The model now separately incorporates the effect of climate change through reductions in their lifetime value of moving to these places, but also through the increased cost of moving away. This causes their lifetime damages to increase the most, relative to the modest baseline damages. On the other hand, I find that climate damages in the warm parts of the country increase to 40% of lifetime value, or a 3 percentage point increase on the baseline. This is because the model accurately captures that not only are they unhappy in their current location, but they are also less likely to be able to leave.

F Simulations

F.1 Census Data

The counterfactual simulations use data on the entire male population living in Mexico, as of 2020. I discuss here how I use publicly available Census data to create an individual-level dataset to simulate population movements.

I use the 2020 full-count Census (El Censo de Población y Vivienda 2020) to recover municipality-level distributions of education, age, and agricultural employment. More specifically, I use the características económicas module to recover education levels in each municipality. These data contain information on the count of individuals, by sex, in each of the following education categories: no schooling, complete primary education, incomplete secondary education, complete secondary education, complete technical education, completed upper secondary, and completed higher education. I define low education as those with no schooling (less than 6 years), medium education as those with complete primary or incomplete secondary education (6–11 years), and high education as those with complete secondary or higher education (12 or more years). When forming the municipality-level education distribution, I subset to only include males.

The above survey is based on the basic questionnaire, which is administered to all individuals living in Mexico. The basic survey does not contain information on the sector of employment. For this reason, I additionally use the extended questionnaire (Tabulados del Cuestionario Apliado) to recover the proportion of each municipality's population that works in agriculture, livestock, forestry, fishing, and hunting. I use this proportion to define agricultural employment. A drawback of this extended questionnaire is that it only covers the population living in private households; it misses all individuals living in collective housing, the homeless, and those in the Foreign Service. Moreover, there are three municipalities in Mexico that are not covered by the extended questionnaire. I assume that their agricultural employment rates are the same as the weighted average of the other municipalities in their assigned location.

I use the Características de la Población module to recover raw population counts, broken apart by sex and age, at the municipality level. I subset these data to only males who are less than 62 years old (the terminal age group in the model). I keep all children younger than 17 in the data, as I will bring these individuals into the model as they come of age.

The above three data sources are merged to create a set of municipality-level state-variable distributions and population-by-age counts. I use the Köppen maps, discussed in Appendix A.3, and the metropolitan-zones, discussed in Appendix A.4 to aggregate this information to the 27 model-locations. In particular, I sum population-ages across each municipality in a given location, and use the total population, at the municipality level, to weight the education and agricultural employment distributions.

This framework allows me to create an individual-level dataset. To do so, for each location and age group, I draw the number of individuals as given by that location-age grouping. I assign their agricultural employment status from a Bernoulli draw with a location-specific success probability. I then assign their education level from a multinomial draw with a location-specific probability vector. Since the census data does not contain information on individual birth municipalities, I assume that everyone in the data currently resides in their birth location. Finally, I assign children to parents using the population counts below 17 and fertility rates (number of children by age) from the MMP data. I hold the resulting dataset fixed for the counterfactual simulations, across both counterfactual scenarios and within scenarios and across iterations.

Table 1, in the main text, presents statistics on the relevant variables across the 2020 census and the MMP data.

F.2 Simulation Framework

I use the individual-level dataset described above to simulate the migration choices of the entire male population through 2080 under a variety of counterfactual scenarios. Within each counterfactual presented, I simulate behavior for 100 iterations. Within each iteration, I draw new shocks from:

- Moving types, from a Bernoulli distribution with probability given it's MLE estimate
- Belief types, from a Bernoulli distribution with probability given it's MLE estimate
- Weather shocks across the three primary Köppen climates for the six weather states, from a mean-zero Gaussian with covariance given by that from the historical weather data
- Preference shocks for each individual-year, from a T1EV distribution
- Fertility shocks for each individual-year that is not a parent, from a standard normal

These shocks, along with the individual-level data described above, completely characterize the state of each simulated person. I use the MLE parameter estimates to solve the relevant Bellman equations for each individual to simulate forward their life choices.

F.3 Counterfactual Scenarios

As discussed in the paper, I model a business-as-usual climate scenario as one following the SSP2-4.5 scenario. This scenario is characterized by a moderate degree of warming under current climate policies. On the other hand, I model a scenario of limited climate change as a climate centered around recent historical averages, those from 1950 to 1980. Table F1 presents the estimates of ρ for the business-as-usual scenario, averaged across the entire period.[53]

F.4 General Equilibrium

This paper models migration in a partial equilibrium setting. One potential concern is that general equilibrium forces will respond to shifts in populations through space through a suppression of the wage level in desirable locations. Some percent of the population modeled as migrating will sufficiently marginal in their decision such that this reduction would be sufficient to reverse their decision; either causing them to move to a different location or not to move at all. This would imply that both my estimated value of migration, my migration rate elasticity would be biased upwards, and the estimated percent reduction in climate damages because of migration, would be biased upwards. I find that these forces are relevant, but that the primary conclusions of the paper remain outstanding after accounting for the described congestion impacts.

Before describing the relevant algorithm, I first discuss how such impacts should enter into the model. In particular, define ξ as the elasticity of wages w with respect to labor supply Q. Then if Q_ℓ rises by x%, w_ℓ will increase by $x\xi$%. I calibrate $\xi = -0.3$ (Borjas, 2003).

I iterate on the following algorithm until population shares converge:

1. Simulate migration decisions with the estimated model and wage equation through 2038
2. Calculate the difference in population levels from 2023 to 2038 throughout Mexico
3. Assume that wages respond to this change in population[54]
4. Resimulate behavior under the new wage regime

For example, in the first step of this process I use the estimated model and the setup described in Section 6 to simulate the behavior of each male in Mexico, beginning in 2020 and ending in 2038. This allows me to calculate the percent change in populations between 2023 and 2038 for each of the

[53] Each year of the simulation has its own set of coefficients corresponding to the previous 30 years of weather outcomes. I present the coefficients when estimating the transition across the entire simulation period. These are not the coefficients I use in the counterfactuals.

[54] In particular, I assume that wages respond linearly to this change in populations up to 2038, and then remain at 2038 levels.

locations in the choice set.[55] I use these changes in populations to calculate updated wages, where the new wage is calculated according to the change in populations and the calibrated elasticity. I then simulate behavior according to this update in wages, and repeat steps 2-4 until the maximum difference in population differences, across locations, between iterations, is sufficiently small.

More formally, define $m^{(k)}(t) \in \mathbb{R}^{27\times 27}$ as the migration matrix at year t in iteration k. Then, $m^{(k)}(t)_{ij} = \int \mathbb{1}\{\ell = i \cap \ell' = j\} dP(k)$, where $P(k)$ is the distribution of individuals at iteration k. Let $n^{(k)}(t) \in \mathbb{R}^{27}$ be the population vector at year t in iteration k. Then, $n_i^{(k)}(t) = \sum_j m_{ij}^{(k)}(t)$. Finally, define $a^{(k)} \in \mathbb{R}^{27}$ as the percentage change in populations from 2023 to 2038 in iteration k. Then, $a_i^{(k)} = \frac{n_i^{(k)}(2038) - n_i^{(k)}(2023)}{n_i^{(k)}(2023)}$. This algorithm converges when $d(a^{(k)}, a^{(k-1)}) < \varepsilon$, where $n^{(k)}$ comes from simulated population movements, as described in Appendix F, with wages updated such that $w^{(k)} = w^{(k-1)} * (1 + \xi a(k-1))$. A sufficient condition for convergence is that $\frac{\partial a_p^{(k+1)}}{\partial a_j^{(k)}} \to 0 \ \forall p, j$. If all marginal movers across simulations decide to either move to their initial destination or to not move at all, then this condition will be met. In reality, most chosen destinations are close substitutes to other locations within the choice set. This implies that the sufficient condition is unlikely to hold in practice.

[55] Beginning in 2023 only aids in speeding up convergence of this algorithm. Similar results are drawn when using 2020 as the base year.

Table F1: Average First Stage SSP2–4.5 Climate Transitions

	Temperature	Precipitation
Dry		
Intercept	16.264189***	0.291821***
	(3.6974)	(0.0588)
Time Trend	0.028378***	-0.000175
	(0.0069)	(0.0007)
Lag	0.199050	0.142953
	(0.1824)	(0.1732)
Temperate		
Intercept	16.632479***	0.989206***
	(2.9162)	(0.1339)
Time Trend	0.031055***	-0.002027
	(0.0063)	(0.0013)
Lag	0.112766	-0.159349
	(0.1559)	(0.1550)
Warm		
Intercept	18.750155***	0.996126***
	(2.5778)	(0.1177)
Time Trend	0.018471***	-0.001678***
	(0.0028)	(0.0004)
Lag	0.238298**	0.087531
	(0.1049)	(0.1069)

Notes: Temperature is the average of the daily temperature distribution in the relevant climate. Precipitation is the average the total level of agricultural-season precipitation across the locations composing the relevant climate. The time trend is measured in years from 2020. The data used in estimation come from Thrasher et al. (2022), are subsetted to years 2020–2050.