NBER WORKING PAPER SERIES

WEATHER, CLIMATE, AND MIGRATION IN THE UNITED STATES

Jamie T. Mullins
Prashant Bharadwaj

Working Paper 28614
<http://www.nber.org/papers/w28614>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

Much thanks to Teevrat Garg and seminar participants at various conferences and seminars for helpful feedback. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Jamie T. Mullins and Prashant Bharadwaj. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Weather, Climate, and Migration in the United States
Jamie T. Mullins and Prashant Bharadwaj
NBER Working Paper No. 28614
March 2021
JEL No. Q0

ABSTRACT

Do people move as a result of temperature shocks? Documenting weather as a push factor for migration is crucial for informing policy related to climate change and adaptation. This paper studies the impacts of high-temperature days on out-migration from counties in the US. We find that outmigration responds to long-term variation in temperatures, but not to the short-term temperature variations that are commonly leveraged in the literature. We provide evidence consistent with the idea that the effect of long-term variation in temperature is driven by changes in expectations regarding future conditions, and specifically climate change.

Jamie T. Mullins
Department of Resource Economics
University of Massachusetts Amherst
Stockbridge Hall
80 Campus Center Way
Amherst, MA 01035
jmullins@umass.edu

Prashant Bharadwaj
Department of Economics
University of California, San Diego
9500 Gilman Drive #0508
La Jolla, CA 92093
and NBER
prbharadwaj@ucsd.edu

1 Introduction

Climate change is expected to impact many aspects of the global socio-economic system, including the distribution of global populations. Human migration has been widely identified as both one of the greatest potential problems driven by climate change and also one of the most promising adaptive mechanisms for those facing increased risks under climate change (Black et al., 2011; IPCC, 2014; Biden, 2021). There has been substantial research into the likely impacts of climate change on migration in recent years (see Kaczan and Orgill-Meyer, 2020 and Cattaneo et al., 2019 for recent reviews), however this work has focused on migration within and out of developing countries as the populations of these countries face the gravest risks. There has also been more attention paid to migratory responses to catastrophic events than the slower-onset effects of climate change such as droughts and temperature shifts, this is true in both poor and rich countries (Kaczan and Orgill-Meyer, 2020; Laczko et al., 2009). We aim to fill an important gap in the literature with this investigation by characterizing the responsiveness of internal migration within a developed country - the United States - to slowly evolving temperature changes, and showing that such responsiveness is best understood as adaptation to expectations regarding future local conditions under climate change.

As a number of recent papers have done, we focus on the distinction between responses to short- and long-term changes in temperatures as critical to identifying impacts relevant for thinking about climate change (e.g., Kolstad and Moore, 2020). Through the introduction of an extension to the Random Utility Model commonly leveraged in the study of migration, we show that short-term (or transitory) temperature increases are not expected to lead to substantial outmigration, while longer-term changes in local conditions (or perceptions thereof) could prompt outmigration, holding constant other adaptation mechanisms. These predictions stand in contrast to results for other, previously investigated outcomes which do respond to short-term variations in weather, for example mortality, morbidity, and crop yields (Deschênes and Greenstone, 2011; White, 2017; Schlenker and Roberts, 2009). Our empirical results bear out the predictions of our model, with migration responding to long-term, but not temporary, temperature variation. We also provide evidence that the active channel for the identified temperature-responsiveness of migration is changes in expectations regarding future conditions, and in particular expectations regarding the likely local implications of climate change. Taken together, our results demonstrate that migration can serve as a mechanism for adaptation to climate change, and is already actively functioning as such in the United States.

Our empirical analysis leverages annual data from the Internal Revenue Service (IRS)

on county-to-county migration in the United States from 1983-2017 in order to shed light on whether people move due to weather and climate changes in their origin counties. To examine short-term “weather” shocks, we estimate a basic model of county-to-county migration flows with fixed effects for origin-destination pairs and state by year fixed effects. This specification is derived from a Random Utility Model of migration decisions adapted from prior work (Beine and Parsons, 2015; Beine et al., 2016). Measures of temperature serve as the independent variables of interest throughout this paper. As a result of the state by year fixed effects, the coefficient on temperature in this specification is identified solely from differential deviations from local mean conditions between counties within the same state and year. As such, these estimates reflect strictly the “short-run/weather/no-adaptation response” (Auffhammer, 2018b; Kolstad and Moore, 2020). We find that such short-term “weather” shocks do not result in any statistically significant or economically meaningful outmigration.

We next re-estimate a similar empirical model, again using origin-destination fixed effects, but now we allow inter-annual (and longer) variation in temperatures to contribute to our identification, and find substantial and significant effects of extreme temperatures on outmigration rates. Temperature coefficients estimates from this specification are identified from both the short-term and the long-term variation and imply economically meaningful and statistically significant outmigration responses to higher temperatures. In particular, we find that each additional day with a mean temperature between 80-90°F increases annual outmigration of households by 0.43% relative to a day with a mean temperature between 60-70°F, while a single additional day >90°F increases yearly outgoing migration flows by 0.96%.

We provide two additional lines of evidence that these estimates are driven by long-term variation in temperatures. First, we estimate the main model after filtering the data (both independent and dependent variables) using high-pass time-series filters which limit considered variation to frequencies below the specified filter frequencies (Hsiang, 2016). Second, we undertake a long-differences estimation in which long-term changes – over 30 years – in smoothed data (5-year averages of the independent and dependent variables) are used to estimate the effects of temperature deviations on migration (Burke and Emerick, 2016). Both approaches confirm that long-term variation drives the identified migratory-responsiveness to temperatures.

We also investigate expectations about future conditions as an important channel for *why* long-term variation might affect migratory decisions. First, we estimate a Common Correlated Effects (CCE) model based on the work of Pesaran (2006) and more specifically Bertoli et al. (2016) to test how the exclusion of the role of expectations of future conditions

impacts estimates of temperature effects on migration. The basis of this approach is the observation that migration decisions are likely impacted by “the confounding influence exerted by the future attractiveness of various options in the choice set”, and that expectations of such future attractiveness function as unobserved common factors which are addressable via CCE methods (Bertoli et al., 2016).

Our findings when we estimate this CCE-based model suggest that changes in expectations about the future are driving the migratory responses to temperature that we identify in our main analyses. Second, we find that high-temperature effects on migration are larger in areas where belief in global warming is stronger, and that such beliefs in global warming also respond to previous experiences with high-temperature realizations. Together these results suggest that expectations of future weather conditions – and in particular of the local effects of climate change – are important factors underlying the identified temperature responsiveness of outmigration.

While migration due to climate change has been offered as an important adaptation mechanism, micro evidence on whether people move due to higher temperatures in developed countries has only become a focus of research very recently. The burgeoning literature in this area has focused largely on extreme events that cause severe disruptions to the economy and well-being. For example, Mahajan and Yang (2017) show that people move away from places that experience devastating hurricanes, Hornbeck (2012) finds that migration was the primary adaptation mechanism used by those affected by the American Dust Bowl in the late 1930s, and Boustan et al. (2012) study migration responses to natural disasters in the 20th century. Additionally, there is increasing evidence from developing countries where incomes are much more closely tied to weather due to a higher reliance on agriculture. For example, Cai et al. (2016), Cattaneo and Peri (2016), and Beine and Parsons (2015, 2017) all study the link between temperatures and international migration in multi-country analyses focusing on developing economies. Bohra-Mishra et al. (2014) find migratory responses to weather shocks in Indonesia; Mueller et al. (2014) examine the relationship between heat stress and long-term migration in Pakistan; Marchiori et al. (2012) provide evidence on weather anomalies and migration in sub-Saharan Africa; and Missirian and Schlenker (2017) provide evidence on refugee applications increasing as a result of temperature shocks in developing countries. While such papers are indeed important in identifying a weather-migration connection, it is also useful to consider how internal migration responds to non-catastrophic sources of weather and climate variation in contexts where other coping mechanisms might be present (these can range from direct adaptation mechanisms such as air conditioning to income smoothing resources such as social safety nets or well-functioning credit markets). We are also able to shed light on climate expectations as an important channel for why long-term

variation matters for migration in our context.

Such internal migration in the US context provides a unique opportunity to examine migration decisions which are unconstrained by the policy and administrative barriers which define international migration flows. This is yet another important avenue by which we contribute to the extant literature investigating the likely impact of climate change on migration using international migratory flows (Beine and Parsons, 2015, 2017; Cai et al., 2016; Cattaneo and Peri, 2016; Cattaneo et al., 2019). Additionally, migration within the United States - in the modern era we study - is unlikely to be driven by concerns for survival, and can instead be characterized as investment or preference-driven migration which is more likely to be well-characterized as an outcome of an optimization problem rather than driven by liquidity constraints (as in Bazzi, 2017). Closely related to this paper is the work of Feng, Oppenheimer, and Schlenker (2012) who examine net-migration flows in the United States due to changes in crop yields (which in turn are affected by the weather) in the Corn Belt. While agricultural incomes no doubt play a role in determining push and pull factors, we show that migratory responsiveness to extreme temperatures are not limited to agriculturally-dependent areas. Our paper also adds to this knowledge space by focusing on short-term versus long-term changes in weather and climate, and relates these changes with perceptions about climate change as a deciding push factor in migration.

The remainder of this paper is structured as follows. In Section 2 we describe the data used for the analyses and the way in which migration as an outcome variable fits into the temperature-effects literature in economics. Section 3 presents the conceptual model and introduces an analog to the Permanent Income Hypothesis that will frame the discussion of our analytical results. These results, as well as the details of the employed empirical methods, are laid out in Section 4. Section 5 examines heterogeneity and robustness of our results and presents a descriptive assessment of migratory destination characteristics. Section 6 concludes.

2 Data and Setting

2.1 Data

In order to examine internal migratory flows within the US, we rely upon annual Statistics of Income (SOI) data from the Internal Revenue Service of the United States (IRS). By matching filing addresses from tax returns in sequential years, these data provide directional county-to-county migratory flows back to the early 1980s. Currently, the IRS provides access

to data beginning in 1990.¹ This is the data, and thus the time period from 1990-2017, which we use for our main sample. For several analyses - for which longer panels are of particular value - we leverage additional SOI migration data going back to 1983 as archived and made available by the University of Michigan ICPSR project.²

The SOI migration data includes directional, county-to-county flows of three measures of migration: 1.) filed personal tax returns, 2.) claimed exemptions, and 3.) aggregate income claimed. The count of returns serves as a proxy for the number of households that migrate, while the number of claimed exemptions proxies for the number of individuals that migrate.³ Both these measures are available for all years in the data, while inclusion of aggregate income begins in 1992.

Weather conditions, temperatures and precipitation, are merged on to each migratory flow for both origin and destination counties. Weather data is from NOAA’s Global Historical Climatology Network and assigned to counties based on the inverse distance weighted average from all monitors within 300km of the centroid of the relevant county. The temperature measures used in this analysis are annual counts of days with reported mean temperatures falling in each of eight distinct temperature bins in a given year for a given county. The bins capture counts of days with mean temperatures in 10-degree Fahrenheit ranges from 30°F to 90°F, with the remaining bins counting days <30° and >90°F. Following the order of operations laid out in Dell et al. (2014) to maintain variability in temperatures through the aggregation process, these counts are calculated for each weather monitor in the data and then aggregated to the county level. Total precipitation and mean daily temperatures are also calculated for each monitor and aggregated to the county level based on the same inverse distance weighting procedure. Only data from origin counties in the continental United States are included in the analytic sample.

Additional county-year data for controls and supporting analyses was taken from the Economic Profile by County data product (CAINC30) provided by the Bureau of Economic Analysis and the Yale Program on Climate Change Communication.⁴ For the details of our sample construction and data idiosyncrasies, please see Appendix Section A.2.

Table 1 summarizes the annual county-to-county migration flows and non-migrant totals for each of the three IRS SOI outcomes. The table also reports the average annual county-of-origin temperature conditions in our main sample.

¹Data available: <https://www.irs.gov/statistics/soi-tax-stats-migration-data>, accessed Feb. 12, 2021.

²Data available: <https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/2937>, accessed Feb. 12, 2021.

³See the SOI website for complete information: <https://www.irs.gov/statistics/soi-tax-stats-migration-data>.

⁴Available: <https://apps.bea.gov/regional/downloadzip.cfm> and <https://climatecommunication.yale.edu>, respectively. Both accessed Feb. 23, 2021

The implied total, annual out-migration rates for households, individuals, and income in the data are 5.04%, 4.39%, and 3.90% respectively. These estimates do not exactly reflect actual out-migration rates because low county-to-county flows do not appear in the data for privacy reasons, and we do not account for outflows to international destinations (or U.S. destinations not included in the sample). Nevertheless, these rates serve as the baselines against which our estimated marginal effects can be assessed. Straightforward comparison of these rates to one another suggests that smaller, lower-income households are more likely to migrate.

2.2 Migration as an outcome

The identification of adaption to weather variation and ultimately warming due to climate change is typically done via the comparison of dose-response functions of temperature-impacted outcomes over time or across space (Carleton and Hsiang, 2016). Smaller-magnitude responses to temperature variation at a given time or place are usually taken to indicate effective adaptation to high temperatures (e.g., Barreca et al., 2016; Burke and Emerick, 2016).

Migration decisions are different from most other outcomes of interest in the study of temperature impacts and climate change adaption as migration potentially represent a direct measure of adaptive action. Just as the crop switching studied by Hornbeck (2012) or changes in planting time studied by Kala (2017) represent adaptive responses to different precipitation patterns (drought and shifting monsoon start, respectively), we study migration as an “explicit measure” of adaptation to higher realized local temperatures (Carleton and Hsiang, 2016).

Kolstad and Moore (2020) separate effects of climate change into two categories: short- or long-term responses. In this framing, short-term responses are those for which “expectations, beliefs, and... investments [are] held fixed”, while long-term responses to changes in climate conditions take into account “both expectations and capital investments”. Auffhammer (2018b) presents a related framing that additionally identifies short-term responses as those taking place on the intensive margin of activity, while long-term responses occur on the extensive margin. Hsiang (2016) presents yet another 2-category framing of the effects of climate as “direct” and “belief” effects, in which direct effects arise directly from the weather realizations (which are functions of the climate) and belief effects result from “individuals’ beliefs” regarding climate “affect[ing] their decisions and resulting outcomes”. In each case, the first category of effects of climate change can be reasonably proxied - in the absence of the second category of effects - by the estimated impacts of “unexpected and temporary”

weather variation (Kolstad and Moore, 2020).⁵ We will develop a model that assumes that climate change impacts on migration (if any) would be the second type of effect of climate change in each framing, that is a long-term, extensive margin, belief effect. Such effects do not arise from - and therefore cannot be identified based on - “transient” weather shocks. As a means of investigating the validity of our modeling assumption, we lay out a working hypothesis that adapts the conclusions of the Permanent Income Hypothesis for amenity flows, and then undertake an empirical investigation of this hypothesis. Thus, we characterize, weigh evidence for, and ultimately conclude that migratory responses to high temperature realizations are in fact responses to perceived long-term (or “permanent”) shocks to local weather conditions arising from anthropogenic climate change. Put another way, migration is already an active channel for climate change adaptation.

3 Conceptual Model

The aim of this investigation is to characterize the impact that local temperatures have on internal migration decisions in the United States, and in particular, how the sensitivity of such migratory responses to temperatures might differ depending on the periodicity - and thus the underlying source - of the relevant temperature variation. Our empirical context allows us to observe household decisions which can reasonably be characterized as the outcomes of individuals weighing the advantages and disadvantages of staying in their current place of residence versus migrating to another location.⁶ Our focus will be on how realized temperatures impact migratory decisions, and whether migration motivating temperature variation can be thought of as short-term (i.e., weather variation) or long-term (i.e., climate change).

Using the notation of the Random Utility Model, in each period, t , individuals beginning in origin county, j , choose which destination county, k , to live in for the period. Non-migrators are modeled as selecting $k = j$, with utility in year t of:

$$u_{ijjt} = A_{jt} + \epsilon_{ijjt} \quad (1)$$

For all others, $k \neq j$. Such migrators face moving costs, denoted M_{jkt} , and thus utility:⁷

⁵In the terminology of Hsiang (2016), the “marginal treatment comparability assumption... that the effect of a marginal change in the distribution of weather (relative to expectation) is the same as the effect of an analogous marginal change in climate” is true for short-term, intensive, and direct effects.

⁶Our model and derivations are adapted from those of Beine and Parsons (2015) and Beine et al. (2016).

⁷Those familiar with dynamic choice models will note the silence of this approach with regard to implications of the present decision on future migration decisions. Consideration of the discounted expected values of optimal future strategies for staying and migrating can be added to this framework (see Beine et al., 2016, section 2.b.ii). We consider such an extension in Section 4.3.3.

$$u_{ijkt} = A_{kt} - M_{jkt} + \epsilon_{ijkt} \quad (2)$$

In both expressions, A_{kt} represents the utility derived from the amenities of the county of residence for the period. This term is often referred to as the deterministic component of the utility, while ϵ_{ijkt} is the individual-specific stochastic term. We next assume the stochastic term follows an iid, type-1 extreme value distribution and apply the results of McFadden et al. (1973) in taking the ratio of movers to stayers in a given county.⁸ Taking natural logs of both sides of the resulting statement yields the following linear relationship - in expectation - between the natural log of the migration rate (herein, the ratio of migrators to non-migrators), amenity levels in the origin and destination, and costs associated with migration. See Appendix Section A.1 or Beine and Parsons (2015) for the details of these steps.

$$\ln \left(\frac{N_{jkt}}{N_{jkt}} \right) = A_{kt} - A_{jt} - M_{jkt} \quad (3)$$

We now substitute in a generic outcome variable to represent the migration rate and split amenities into weather and non-weather amenities, W_{jt} and N_{jt} , respectively.

$$Y_{jkt} = W_{kt} + N_{kt} - W_{jt} - N_{jt} - M_{jkt} \quad (4)$$

One aim of this investigation is the empirical estimation of the relationship between realized temperatures and the out-migration rate, that is dY_{jkt}/dW_{jt} . The second aim of this paper is to decompose the response to realized weather into short- and long-term components. We now think about realized weather - and in particular, the realization of the measures of specific dimensions of weather - as being a draw from a (potentially joint-) probability distribution which characterizes the local climate (e.g., Hsiang, 2016; Kolstad and Moore, 2020). Put another way, local weather can be characterized as the sum of the expected value of the local climate distribution, C_j , and a stochastic term, $\varepsilon_{jt} \sim (0, \sigma_j^2)$, such that we write: $W_{jt} = C_j + \varepsilon_{jt}$. For our setting and interest in temperature then, we assume that climate is characterized by the first moment (mean) of the weather distribution, and the realized temperature is simply a draw from that distribution (Auffhammer, 2018b). Rather than

⁸This distributional assumption, first laid out by McFadden et al. (1973), implies independence of irrelevant alternatives, and thus drops consideration of so-called “multilateral resistance” (the idea that changes in non-origin, non-destination locations impact flows between the origin and destination) from the model. This has long been standard in both the trade and migration literatures, and we adopt it here under the basic assumption that factors impacting the origin-destination migration decision, but occurring in non-origin, non-destination counties, are not correlated with our variables of interest, namely origin-county temperatures (see: Beine and Parsons, 2015).

assuming local weather distributions - and therefore climates - are fixed, however, we want to allow for the potential change of expectations regarding local climates, and we therefore write $W_{jt} = C_{jT} + \varepsilon_{jt}$, where T is multi-year period that includes year t .

$$Y_{jkt} = C_{kT} + \varepsilon_{kt} + N_{kt} - C_{jT} + \varepsilon_{jt} - N_{jt} - M_{jkt} \quad (5)$$

The separation of the weather impacts on out-migration into short- versus long-term can therefore be thought of as the attribution of any identified out-migration effects between the random weather variation term, ε_{jt} , and the long-term changes in local conditions embodied by variation in C_{jT} .

An important distinction between this model of temperature responsiveness and those of other outcomes, is that the above characterizations (in Equations 3-5) are only true in expectation. This means that the ε_{jt} and ε_{kt} terms drop out due to their zero values in expectation. In practical terms, this implies that random, or short-term, weather variation should not impact migration rates. If responses to realized weather are observed, the model implies that such effects would be acting through changes in expectations regarding climatic conditions, $E(C_{jT})$ for origin counties, which in turn impact location choices and migration. It is important to recognize that this characterization is a direct outgrowth of the assumptions of our modeling framework,⁹ though the intuition is also straightforward. When an outcome is a function of an expectation rather than the realization of a state, the only way that observed variation can influence the outcome is through changes in the expectation. In the context of this study, migration decisions depend only on expectations regarding future amenity flows and conditions, and for temperature, only long-run variation should influence expectations regarding future conditions. Therefore, only long-run temperature variation should impact migration. We will spend the rest of the paper empirically investigating the validity of our model's assumptions and implications.

As made explicit in the model above, we expect climate change to impact migratory decisions, but short-term variation in weather to have no effect. This framing suggests an analogy to the Permanent Income Hypothesis ("PIH") put forward originally by Friedman (1957). Just as the standard PIH postulates that changes in consumption patterns are a function of permanent, but not transitory, changes in income, we hypothesize that our observable indicator of a permanent change in consumption - namely migration - ought only respond to

⁹Models of temperature responses of short-term, intensive, and/or direct effects of climate change depend directly on realized exposure, without expectations. For instance, heat-stroke deaths and corn-yields respond to temperature exposure, not expectations of conditions as is implicitly assumed in the RUM framework for migration.

permanent (or perceived permanent), but not transitory, changes in amenity flows.¹⁰ Given the timescale over which climate change might be undone, we equate long-term (in the sense used by Kolstad and Moore, 2020) and permanent time span descriptors throughout our discussion. The amenity of interest in this setting is weather, and in particular temperatures, so our working PIH-analog suggests that migration responds to permanent, but not transitory, changes in temperature.

Our outcome of interest is behavioral in nature, and as such, estimated effects of temperature variation are informative regarding perceptions of the connection between locally-unusual temperature realizations and changes in beliefs regarding the underlying long-term weather distribution (i.e.- climate) rather than the reality of such a link.¹¹

Just as the PIH has been tested empirically (see Hall, 1978 and many others since), we will bring our working hypothesis to the data. The evidence for or against our working hypothesis is of interest for better understanding migratory flows and migration decisions (Molloy et al., 2011; Cattaneo et al., 2019); the consideration of modes of adaptation to climate change (Bohra-Mishra et al., 2014; Beine and Parsons, 2015; Cai et al., 2016); how we think of responses to weather variation in the context of a changing climate (Burke and Emerick, 2016; Taraz, 2017; Kala, 2017); and untangling responses to short versus long-term variation in weather outcomes (Dell et al., 2012; Hsiang, 2016; Auffhammer, 2018b; Lemoine, 2018; Kolstad and Moore, 2020).

4 Empirical Framework and Results

Because we cannot separately observe climate and weather variation, our empirical model builds from Equation 4 with a focus on temperature as our weather-measure of interest. In order to isolate quasi-random variation in realized-weather, we leverage fine-geographic fixed effects following the approach laid out by Deschênes and Greenstone (2007) and Dell et al. (2014). In particular, we will use directional county-pair fixed effects - denoted α_{jk} - in most

¹⁰Migration could conceivably be considered a means of consumption smoothly rather than an indicator of changes in consumption patterns. This alternative framing would lead to the prediction that migratory responses would occur in response to short-, but not long-term, shocks, which is not borne out by the data in the current setting.

¹¹This is again related to the conceptualization, put forward by Hsiang (2016), that weather deviations have both direct and belief effects. In this setting, we assume (and will test to see) that there is no direct effect of short-term temperature variation on outmigration. Our working hypothesis therefore implies that our estimates are driven completely by “belief effects” in Hsiang’s framework. In the framework of Kolstad and Moore (2020), we hypothesize that migration is analogous to a change in “fixed capital investments” which is a “long-run response” that adjusts “to *permanent* changes in climate, but... cannot adjust in response to short-term variation in weather.” Finally, in the framing of Auffhammer (2018b), our working hypothesis states that migration is an “extensive margin adaptation” because it is not “a short-run weather response, [but] a long-run climate response”.

of our regression analyses. These fixed effects absorb all time-constant features of the origin AND destination counties as well as the relationship between them. We will also control for location-specific time trends using various functions of time and geography for both the origin and destination locations. We indicate the general location-specific functions of time for the origin and destination as $f_j(t)$ and $f_k(t)$ respectively. We will consider functions of origin-temperature, $f(W_{jt})$, and also include a vector of origin-relevant variables, $\mathbf{X}_{jt}$, which will always include (at least) precipitation controls. We can represent this general form as follows:

$$Y_{jkt} = f(W_{jt}) + \mathbf{X}_{jt}\boldsymbol{\mu} + \alpha_{jk} + f_j(t) + f_k(t) + \epsilon_{jkt} \quad (6)$$

There are several things worth noting in this general empirical specification. Regressors for specific, time-varying components of the destination: N_{kt} and W_{kt} , are explicitly omitted due to their endogeneity as outcomes of agent decisions. Non-weather amenities in the origin location, N_{jt} , are also excluded from the regression to avoid bias from “bad controls”, which arises when variables that could themselves be impacted by the treatment of interest are included as control variables in a regression analysis (Angrist and Pischke, 2008). Such “bad controls” are especially likely in the context of temperature effects given that so many potential control variables may actually vary with weather (Hsiang et al., 2013; Cattaneo and Peri, 2016).

Following Barreca et al. (2016), unusual local precipitation in the origin-county is controlled for in all specifications using indicators for county-specific total annual precipitation above the 75th percentile or under the 25th percentile for the county. These controls are included in the origin-control vector $\mathbf{X}_{jt}$.

For the outcome variable, Y_{jkt} , we opt for the inverse hyperbolic sine transformation (“IHS” hereafter) of the migration rate rather than the natural log because of the many annual dyadic flows with zero values in the data. Regression estimated marginal effects on IHS-transformed outcomes can be interpreted approximately as percent changes in the same way that is common for natural-logarithm-transformed outcome variables. Through a series of robustness checks, we will also show that the character of the results for our main specification is unchanged by the use of the migration rate in levels or $\ln((1 + N_{jkt})/N_{jjt})$ as the outcome of interest (see Table A3 for estimates).

Our empirical specification allows for the consideration of each of the three migration rates that can be calculated from the IRS data: 1.) Tax Returns; 2.) Exemptions; and 3.) Aggregate Income on filed returns. We will present relevant estimates for each of the three, but ultimately we focus on tax returns as our main measure of interest given that returns are the core unit of observation in the IRS data and all three measures are closely

correlated. Thus, unless otherwise indicated, the outcome being considered is the inverse hyperbolic sine of the tax return migration rate defined as the ratio of migrating tax returns to non-migrating (or staying) tax returns in the relevant year.

Unless otherwise noted, standard errors are clustered at the level of origin-county, as that is the level at which temperature variation exists in our data.

4.1 Short-Term, Temporary Shocks

Turning to the empirical investigation of our working hypothesis, we begin with a model that is linear in temperature and includes granular spatial and temporal fixed effects.

$$Y_{jkt} = \beta W_{jt} + \mathbf{X}_{jt}\boldsymbol{\mu} + \alpha_{jk} + \delta_{Jt} + \epsilon_{jkt} \quad (7)$$

This model includes the direction-specific, county-pair fixed effects, α_{jk} ; state-by-year fixed effects for the state-of-origin, δ_{Jt} (where J indicates the state of county j); and our origin-county precipitation controls, $\mathbf{X}_{jt}$. The granular geographic fixed effects absorb cross-sectional variation while the year fixed effects absorb inter-annual variation. As a result, the coefficient on temperature in this specification is identified solely from differential deviations from local mean conditions between counties within the same state and year. That is, this specification is identified using only components of ε_{jt} from Equation 5 and not from any components of C_{jT} . As such, estimates reflect strictly the “short-run/weather/no-adaptation response” (Auffhammer, 2018b; Kolstad and Moore, 2020). Under the working hypothesis, we therefore expect the estimates of β to be small and statistically insignificant. As reported in Table 2, this is the case across all three measures of migration (returns, exemptions, and aggregate income). While this model includes state-by-year fixed effects, estimates are similarly small in magnitude (and insignificant) when year fixed effects (δ_t) or climate-zone-by-year fixed effects (δ_{CZt}) are used in place of δ_{Jt} . See Table A1 for relevant estimates using different geographic-unit-interactions with the year fixed effects and based on a more flexible specification of temperature. Similar to the estimates in Table 2, coefficients estimates are generally small and statistically insignificant.

4.2 Main Specification: Including Short- and Long-Term Variation

Opening up the variation used for identifying temperature effects, we now present our main specification:

$$Y_{jkt} = \sum_{b=1}^B \beta_b P_{bjt} + \mathbf{X}_{jt} \boldsymbol{\mu} + \alpha_{jk} + \delta_j t + \mu_k t + \epsilon_{jkt} \quad (8)$$

In place of the linear temperature measure used previously, we adopt a series of variables which are counts of days in calendar year t on which county j 's mean temperature fell into each of eight 10°F-wide bins. Six of these bins cover the temperature range from 30°F to 90°F with the remaining bins counting days <30°F and >90°F. The 60-70°F temperature bin is the omitted category in our main “seven-bin” analyses for which $B = 7$. The average number of days in each bin in each year is reported in Table 1, note that days with mean temperatures >90°F are quite uncommon on average. We also estimate specifications in which only two “extreme” temperature bins – counts of days <30°F and >80°F – are included (in these cases $B = 2$). The omitted category in these “two-bin” specifications is days with mean temperatures between 30-80°F. Estimates for each of the three outcome measures are presented in Table 3 and graphically in Figure 1. All models include the origin-precipitation controls and county-pair fixed effects.

This specification includes origin- and destination-specific linear time trends rather than year-fixed effects. This allows for (detrended) inter-annual variation in independent variables to contribute to our identification, and thereby yields temperature coefficient estimates that are identified from both short- and long-term variation. While this approach potentially allows for other, time-varying factors to influence estimates, the inclusion of the rich set of spatial fixed effects and time trends in the specification limits this risk.

Allowing for both short- and long-term variation to contribute to temperature-coefficient identification yields significant, positive, and increasing estimates for higher temperatures across all three migration outcomes. Similar patterns are clearly visible across the three outcomes, namely high temperatures lead to outmigration. We see also that an increase in the incidence of cold days (those with a daily average temperature below 30°F) is associated with reductions in outmigration whether measured in terms of households, individuals, or earnings. Going forward we will focus on tax returns as the main outcome of interest, but generally our results are comparable across the other outcome measures. The resulting coefficient estimates suggest that each additional day with a mean temperature between 80-90°F increases outmigration of households (proxied by tax returns) by 0.43% relative to a day with a mean temperature between 60-70°F. A day >90°F similarly increases outgoing migrant flows by 0.96%.

Given we find no impact from short-term temperature variations on out-migration, and now have identified significant effects of high-temperatures when both short- and long-term variation are considered, these estimates suggest that long-term temperature variation im-

pacts outmigration. Thus, we have now presented empirical evidence consistent with both parts of our working hypothesis. Namely, it appears as though migration - which embodies the decision to permanently change the amenity flows which one experiences - is not affected by short-term (or transitory) temperature changes, but is impacted by long-term (or effectively permanent) changes in temperature.

4.3 Exploring Long-Term Variation

Another means by which to separate the impacts of short- and long-term variation in temperatures is through isolation of the frequencies of the variation leveraged in identification. We have already seen that estimated effects of temperatures identified only with intra-year variation (with year-fixed effects) - which is the shortest frequency variation in our annual data - are small and insignificant. We now explore two methods of limiting identifying variation to lower frequencies and thus, longer time scales.

First, we will filter the data (both independent and dependent variables) using high-pass time-series filters which limit considered variation to frequencies below the specified filter frequencies (Hsiang, 2016). Second, we will undertake a long-differences estimation in which long-term changes in smoothed data (5-year averages of the independent and dependent variables) are used to estimate the effects of temperature deviations on migration (Burke and Emerick, 2016). We use the difference across the longest period over which our data allow for the comparison of 5 year moving averages, 30 years.

We follow these analyses up with a series of estimations which illuminate the roles of expectations and beliefs regarding climate change in the identified effects.

4.3.1 Filtered Analyses

We saw that specifications with year fixed effects – which leverage only intra-annual temperature variation – yield small and insignificant estimates of the impacts of temperature on migration, see Tables 2 and A1. A rough method to reduce temporal limits on the variation considered would be through expanding the period length of the fixed effects, for instance imposing 3-year-period or 5-year-period fixed effects. This approach would however, be sensitive to the specific years grouped together (e.g., whether 1990-1992 or 1991-1993 were the first 3-year-period considered). The application of time series filters is a related, but more robust approach to limiting identifying variation by interval length (similar to procedures described by Hsiang, 2016). Figure 2 and Table 4 report estimates from the main, seven-bin specification based on time-series filtered outcome and temperature variables. Specifically, a Butterworth high-pass filter is applied at expanding time-horizons. This approach follows

the continuum of temporal variation period from the year-fixed effects model to the main specification by moving from the consideration of only short-term variation to the consideration of all variation for identification. The time-series filters allow us to make this transition in smaller steps, but the same general story emerges. When only very short-term variation is used (0-3 year periodicity), we do not see positive effects of high temperatures on migration, but as variation of lower frequencies (i.e., occurring over longer periods) is allowed through the filters, and thus used for identification, the estimated effects of high temperatures approach the main, unfiltered estimates (also reported for comparison). This suggests again that it is not short-term variation in temperatures that is driving our main results, but instead low-frequency, longer-term variation.

It is worth pointing out that even when variation with period length up to 15 years is analyzed, the full estimates of our main regressions are not realized. This suggests that the relevant time-interval over which temperature variations should be considered in order to capture all contributions of temperatures to the migration outcome is longer than 15 years. The commonly considered 30-year-average definition of climate (e.g., Pachauri et al., 2014) is therefore consistent with our results.

4.3.2 Long-Differences

As a means of directly estimated the effects of long-term temperature variation, we also undertake a long-differences estimation of the impacts of temperature on migration. This approach regresses long-term changes in migration rates onto long-term changes in temperatures to estimate the response of migration to long-term changes in temperature realizations. Following Burke and Emerick (2016), we examine changes in 5-year averages in temperature realizations and migration rates across the longest possible time frame in our data, which is 30-years. It is worth noting that the change in local-climate temperature distributions due to human-induced climate change has probably not been large enough over this time frame (1987-2017) to be driving these estimates (Kolstad and Moore, 2020), however, in our setting (in contrast to settings concerned with direct effects of temperatures such as Burke and Emerick, 2016) it is perceptions of a shift in the climate distribution that will drive immigration decisions, and there has certainly been enough time (and media attention) for expectations of temperatures to have changed.

Figure 3 reports seven-bin model estimates for 30-year changes in 5-year averages of temperature-bin counts. Note a “marginal change” in this setting represents an average of one extra day in a temperature bin over a 5 year period, which is a larger marginal change than is considered by the rest of our analyses. That said, we see that higher temperatures are clearly associated with higher out-migration rates.

The marginal shock in this setting embodies a stronger signal of a shift in climate both because the statistics are 5-year averages (and a change in the five year average is more informative than the equivalent change observed in a single year) and because these changes are observed across a thirty-year time period, making inattention much less likely.

Looking at the long differences results together with the estimates from our main regressions suggests that migration may be driven by the binary perception that one's local climate distribution has changed rather than by the perceived magnitude of the change. We see that all hot-temperature bins in the long-differences analysis drive similar out-migration responses suggesting that perceptions of a local shift in the climate distribution may drive outmigration. In contrast, our main results show higher out-migration responses associated with higher temperature bins, suggesting the magnitude of changes in the perception of long-term temperature changes may be larger when one-year temperature anomalies are larger.

4.3.3 Expectations as a Channel

Changes perceived to be long-term or permanent necessarily impact expectations of the future. Our model suggests that temperature shocks can impact migration only through expectations, and the idea of expectations operating as the active channel underlying our results is bolstered by the observation that migration provides no amelioration of harm or protection from experienced temperature shocks in either the past or present.

To explicitly consider the channel of expectations of future conditions in the conceptual model, we add a term to the utility functions at the foundation of the RUM model to account for the discounted expected flow of future utility from a time t location choice, $\beta * V_{t+1}(k)$, as laid out by Beine et al. (2016) and Bertoli et al. (2016). Equation 2 thus becomes:

$$u_{ijkt} = A_{kt} - M_{jkt} + \beta * V_{t+1}(k) + \epsilon_{ijkt} \quad (9)$$

Following the steps in Section 3 (and detailed in Appendix Section A.1) and the work of Bertoli et al. (2016) leads us to a version of Equation 3 that additionally includes expected future utility flows from migrating to county k and staying in county j as follows:

$$\ln \left(\frac{N_{jkt}}{N_{jkt}} \right) = A_{kt} - A_{jt} - M_{jkt} + \beta [V_{t+1}(k) - V_{t+1}(j)] \quad (10)$$

Estimation of the empirical counterpart to Equation 10 without accounting for expectations of future conditions – as we have done up to this point – necessarily rolls any impact of contemporaneous temperatures acting through expectations into the estimated coefficients

on contemporaneous temperatures. What follows is a description of one approach to empirically separate the influence of temperature shocks acting through expectations from effects coming through other channels.

One subtle aspect of Equation 10 is that the functions of future amenity flows, $V_{t+1}(k)$ and $V_{t+1}(j)$, depend on expectations of future conditions in *all* counties in the choice set, rather than only those in the origin and destination counties. This is because future moves could be made to other counties beyond the origin and destination of the current move, and thereby impact future amenity flows.¹² Importantly, changes in expectations regarding future conditions are likely to be correlated across space, and can therefore be viewed as common, but unobservable, shocks, resulting in a multifactor error structure for the empirical counterpart to Equation 10 of the type described by Pesaran (2006). As laid out by Bertoli et al. (2016), the Common Correlated Effects (CCE) estimation procedure - which was first presented by Pesaran (2006) and extended in Chudik and Pesaran (2015) - can therefore remove the "influence exerted by [expectations of] the future" from the estimation of coefficients on the origin characteristics of interest by treating the $V_{t+1}(k)$ and $V_{t+1}(j)$ terms in Equation 10 as common factors impacting all county-to-county migration flows.¹³

If, as we have been assuming, migratory responses to high temperature are driven by changes in people's perceptions regarding local climate norms, and thus expectations of future meteorological conditions in their areas, we should see the impact of high temperatures shrink or disappear when the influence of expectations is controlled for via the CCE estimation procedure. Figure 4 shows the coefficient estimates from the seven temperature-bin model based on the CCE specification. We see that the effects of high temperature days on out-migration disappear when the influence of expectations is removed from consideration. Taken with the evidence in the previous sections, and the impotence of migration to protect against past or current temperature shocks, these results suggest that long-term variation in the realization of high temperatures drives outmigration through alterations in residents' expectations regarding the future.

4.4 Additional Evidence

Having shown a number of avenues of evidence for our working hypothesis and the interpretation of our main results as embodying long-term, adaptive responses to changes in temperature, we now present several other pieces of evidence which are consistent with this framing. As a group, the results of these analyses bolster the evidence already presented for

¹²This is true unless either agents are totally myopic or face zero migration costs, and even if the assumption that ϵ_{ijkt} follows an iid, type-1 extreme value distribution holds (Bertoli et al., 2016).

¹³See Appendix Section A.3 for details of our implementation.

the validity of our working hypothesis.

First, we find that migratory responsiveness to high temperatures increases over time. Column 1 of Table 5 shows relevant estimates from a model in which we add the interaction between year and the count of hot days in the two-temperature-bin specification. We see the coefficient on the interaction is positive and significant, showing that migration responses to high temperatures have grown in magnitude over the time of our sample. This relationship is also shown in Figure 5 which reports separate estimates for the effects of hot days by 3 year period. This general growth in the marginal effects of unusually hot days over time coincides with the growth in understanding and awareness of climate change across the United States. This increase in marginal effects is consistent with our working hypothesis if – as seems reasonable – understanding and awareness of climate change increases the likelihood that high temperature realizations are perceived as indicative of long-term changes in local norms rather than simple anomalies. This connection is visible more directly in the estimates in column 2 of Table 5 which show that larger migratory responses to hot days are observed where rates of belief in global warming are higher.

Related to the above, idiosyncratic weather variation will tend to obscure long-term or permanent changes in climate. Long-term changes would therefore be harder to discern from normal weather variation in places in which year-to-year variation in weather was relatively large. If migration responds to long-term changes in climate, we would expect to see the strongest responses in locales where such changes are more easily discernible, that is in places where year-to-year variation in weather is lower on average. By interacting a measure of local weather variability – the mean absolute deviation in yearly days $>80^{\circ}\text{F}$ – in the pre-analysis period (1983-1989) with the count of hot days each year, we find that areas with greater baseline variability in temperatures respond less strongly to the long-term variation from which our main estimates are identified. Relevant estimates are presented in column 3 of Table 5. Similar evidence is visible in Figure 6, where the marginal effect of one additional hot day falls to zero as the average annual frequency of such hot days during the preanalysis period increases. One additional hot day is a very small deviation from area norms in places that regularly see >67 such days, making such days more easily interpretable as transitory variation rather than informative of long-term (or permanent) shocks to weather.

Finally, if changes in perceptions of local long-term weather norms are driving migration in our analysis through changes in expectations about the future, we might expect to find the same variation would drive changes in stated beliefs about climate change. Figure 7 reports the regression-estimated relationship between survey-reported belief that global warming is happening and the incidence of high temperatures in the preceding year. In models with and without year fixed effects, we see that more hot days, and especially very hot days, are

associated with higher rates of reported belief in global warming the following year. Both migration and belief in global warming respond to long-term variation in the realization of hot days. This is consistent with a link between our working hypothesis and perceptions of climate change.

Our estimates of the effects of short-term and long-term temperature variation, in conjunction with the additional evidence presented in this and the preceding sections, suggest that observed internal migratory patterns are consistent with our working hypothesis and that the perceived permanence of observed temperature changes is related to understanding and perceptions of anthropogenic climate change. This appears to be the case even if the actual driving variation in temperatures is arising from low-frequency variation in the climate system that isn't attributable to humans (Deser et al., 2012; Hawkins and Sutton, 2012).

5 Heterogeneity and Robustness in Main Specification

Proceeding under the assumption that our main estimates represent responses to perceived local increases in long-term temperatures driven by anthropogenic climate change, we now investigate the robustness and heterogeneity of these estimates.

5.1 Robustness

Identification in this setting comes from the heavy reliance on fixed effects. County-of-origin fixed effects narrow the scope of temperature variation which is leveraged for identification to deviations of temperature from long-term local norms. Such variation is considered random, leading to estimates which can be interpreted causally. Table A2 reports regression estimates which build from the simplest specification with a causal interpretation on temperature coefficients to our main specification which (effectively) includes not only origin-county fixed effects, but origin-destination pair fixed effects and both origin and destination linear time trends. Estimates from a specification with origin and destination quadratic time trends are also presented. Across the table we see highly stable estimates of the impacts of origin-county temperature shocks on out-migration suggesting the main thrust of our results is remarkably robust to the details of the specification.

Our baseline results are also potentially sensitive to the choice of transformation of the outcome variable and to weighting schemes. Table A3 reports estimates from the main regression when the untransformed migration rate or the natural log of the “adjusted” (i.e., numerator plus one) migration rate are used as alternative outcome variables (columns 1 and 2 respectively). Columns 3 and 4 report estimates from the main regression when regression

weights of logged population and population are applied. Finally, Column 5 reports the results of a Poisson Pseudo-Maximum Likelihood estimation procedure that addresses the large number of zeros in our data appropriately. Generally, what we see across columns in Table A3, is that high temperatures are associated with significant outmigration while the lowest temperatures reduce outmigration. For the specifications where the magnitudes are directly comparable to our main results, the point estimates are generally quite similar to our baseline results.

5.2 Heterogeneity

Having established the robustness of our main estimates linking anomalous local temperatures to outmigration, we now consider heterogeneity in these impacts. We begin with a simple comparison across outcome measures. As noted above and reported in Table 1, the data used for our analysis imply total annual out-migration rates for households, individuals, and income of 5.04%, 4.39%, and 3.90% respectively. From Table 3, our estimates imply that an additional day $>90^{\circ}\text{F}$ increases the likelihood of outmigration of households by 0.048 percentage points (5.04×0.0096), of individuals by 0.040 percentage points (4.39×0.0091), and 0.031 percentage points (3.90×0.00793) for income, all relative to a day between $60\text{--}70^{\circ}\text{F}$. Given that the same shock (an additional day $>90^{\circ}\text{F}$) changes the anticipated average migration rate of households by more than rates for individuals or income, these results suggest that high temperatures tend to drive out smaller, lower-income households at higher rates.

Responses to temperature shocks are also not uniform across geographies. Again focusing only on tax returns, Figure 8 reports coefficients on hot days ($>80^{\circ}\text{F}$) and cold days ($<30^{\circ}\text{F}$) from the two-bin model separately for each of the nine NOAA Climate Regions. Generally, the pattern of hot days driving outmigration and cold days reducing outmigration holds, but notably not in the West, Southwest, and Northwest regions.¹⁴ The Climate Regions are quite broad, both in terms of geographic extent and in terms of baseline climates.

We next break our sample up into ten equal groups based on the average number of $>80^{\circ}\text{F}$ days each year in the origin county over the pre-analysis period: 1983-1989. Figure 6 presents the estimates of the coefficients on the $>80^{\circ}\text{F}$ temperature bins from the two-bin model. The presented estimates are from a single regression in which group indicator variables are interacted with the temperature bins. The estimates reveal fairly consistent magnitudes of the effects of hot temperatures across areas that tended to have few hot days in the past. However, as the number of hot days historically seen by a location climbs higher, the marginal effect of an additional hot day falls to zero. This suggests decreasing marginal

¹⁴The West Region covers CA and NV. The Northwest Region includes ID, OR, and WA. The Southwest Region includes UT, CO, AZ, and NM.

effects of hot temperatures, as places that saw the most high-temperature days on average tend to respond the least to an additional such day. As noted earlier, this suggests that marginal changes in long-term conditions are less discernible - and therefore less likely to be acted upon - in locations where such a marginal change is relatively small compared to the average number of hot days. Additionally, while the act of migration in response to unusual temperatures perceived as harbingers of future conditions is itself adaptation, the reduced response to high-temperature days in locations with more such days historically suggests other dimensions of adaptation or preference-based sorting are also active.

High temperatures have been shown to impact agricultural yields (e.g., Schlenker and Roberts, 2009), and both Feng et al. (2012) and Cai et al. (2016) find that temperatures drive outmigration (internally and internationally, respectively) through reduced agricultural yields.¹⁵ We probe this relationship in our data using a number of different measures of the centrality of agriculture across counties. In Table 6, we present coefficient estimates from the separate re-estimation of the main specification for origin-counties: 1.) in Corn Belt states and not; 2.) with above and below the median share of land dedicated to agricultural production; and 3.) with below and above the median population density. These results reveal that the impacts of high temperatures have larger magnitudes in the Corn Belt, in more Ag-dependent areas, but in more densely populated areas. The differences in magnitudes are substantial for the Corn Belt vs. Non-Corn Belt comparison, somewhat less so for high- versus low-ag counties, and insubstantial based on the density categorization. Corn Belt results mirror those reported by Feng et al. (2012), though they do more to link observed temperature-driven outmigration to economic losses attributable to reduced agricultural yields. Viswanathan and Kumar (2015) do something similar in India. It is important to notice however, that this cannot be the only active channel driving our results given that significant impacts of extreme temperatures exist for outmigration from places which are densely populated, have little agricultural land, and/or are outside the Corn Belt.

We also briefly consider whether affluence might modulate the population sensitivity to high temperature exposures. In particular, we interact the share of the over-25 population that finished high school and county-level per capita income (both from BEA data) with the counts of hot and cold days. Estimates reported in columns 4 and 5 of Table 5 suggest that out-migration rates in more-educated and higher earning counties are more responsive to high-temperature realizations. These results are in line with those of (Fan et al., 2016), which leverage a sorting model to find "that extreme temperatures have a greater [migration] impact on households with higher educational attainment". Both results are consistent with

¹⁵Others have shown that negative weather-shocks can reduce outmigration by tightening liquidity constraints, though this channel is unlikely to be as central in the U.S. context (Hirvonen, 2016; Bazzi, 2017).

greater sensitivity to high temperature realizations among more affluent populations.

5.3 Destinations

Finally, if people decide to leave an area because of perceived increases in long-term temperatures, to what types of places do they move? The addition of variables for the long-term average weather conditions in destination counties to our main two-bin specification can provide some relevant evidence.¹⁶ However, the characteristics of destination counties are determined by migrant decisions, and are therefore clearly endogenous. Such coefficient estimates have no causal content and should therefore be interpreted as purely descriptive. Given that we find migrators are driven by perceptions of warming in their origin counties, it would make sense that they would be more likely to move to places where hot days are perceived to be less frequent. This is precisely what we see in the regression estimates presented in Table 7. Destinations that have historically had (and are therefore likely to be perceived to have) fewer hot days, are more likely to be migration destinations in our data. Again, this is consistent with our overarching story that migration is driven by perceived increases in local long-term temperatures.

6 Conclusion

Migration is often posited as a likely response to changing climatic conditions, and the reallocation of populations and human capital is an important consideration when assessing the implications of global climate change. To date however, most research linking climatic conditions (and shifts thereof) to migration are focused on international flows, developing countries, and catastrophic events. This investigation has focused on the responses of internal migratory flows to perceived local temperature shifts in the context of a developed county, namely the United States.

Our central empirical result is that more instances of hot-days lead to higher rates of out-migration in terms of households, individuals, and total earnings. Compared to studies which look at international migration in response to temperature shocks, our results are in line with those of Cai et al. (2016) (though they find that high-temperature outmigration is driven by reduced agricultural yields), but stand in contrast to those of Beine and Parsons (2015) which show no effect of “climate factors” (including high temperatures) on international migration. Our results fit with three other studies that find higher temperatures are associated with

¹⁶Origin-destination-pair fixed effects are replaced with origin-county fixed effects to allow for the estimation of coefficients on static, destination-county variables.

higher-rates of outmigration to destinations within the same country. Feng et al. (2012) find that outmigration is driven by “climate-driven changes in crop yields” in the United States Corn Belt, Viswanathan and Kumar (2015) find similar results in India, and Bohra-Mishra et al. (2014) find that local high-temperatures drive out-migration in Indonesia.

Beyond this straightforward identification of the link between temperature anomalies and internal migratory flows, we go further to decompose the the temporal character of the variation in temperatures that underlies our main result. The Random Utility Model framework that commonly serves as the conceptual basis for migration analyses implies that migration should only respond to changes in expected weather, and not idiosyncratic weather realization. We formalize this implication of the RUM framework and develop a working hypothesis that migration will respond to permanent, but not transitory, changes in temperatures. This analog to the Permanent Income Hypothesis could similarly be stated as identifying migration as a long-term, extensive, or belief effect of climate change (in the framings of Kolstad and Moore (2020), Auffhammer (2018a), and Hsiang (2016) respectively). In all cases, the underlying claim of our working hypothesis is that migration is undertaken as an adaptive mechanism in response to long-term changes (or perceptions thereof) in the underlying distribution of local weather.

We present empirical evidence that short-term weather variation is not driving our main estimates (using year fixed effects and high-pass time series filters), that long-term temperature variation is driving our main estimates (using high-pass time series filters and long-differences estimation), and that these effects of long-term variation on migration are acting through the channel of updated expectations regarding local weather conditions. We also present several empirical results suggesting that expectation-changing, long-term variation in temperatures that drivers our main results is associated with climate change. Finally, we present descriptive results that even as outmigration is driven by increased incidence of high temperature days, migration destinations tend to be those that experience (on average) fewer such hot days. Taken together, our result suggest that migration is already serving as an active mechanism for adaptation to climate change within the United States.

While we provide initial answers to the questions of whether climate change will induce migration among those living in developed economies (yes) and when such responses might start (already have), there is much more work to be done to understand how climate change will impact people’s decisions of where to live and whether and where to move. In particular, our estimates do not factor in the effects of increased natural disasters or coastal flooding also expected under climate change, and together these factors could contribute to significant levels of geographic displacement, even in a developed-country setting like the United States.

References

- Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2008.
- Auffhammer, Maximilian, “Climate adaptive response estimation: Short and long run impacts of climate change on residential electricity and natural gas consumption using big data,” Technical Report, National Bureau of Economic Research 2018.
- , “Quantifying economic damages from climate change,” *Journal of Economic Perspectives*, 2018, *32* (4), 33–52.
- Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro, “Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century,” *Journal of Political Economy*, 2016, *124* (1), 105–159.
- Bazzi, Samuel, “Wealth heterogeneity and the income elasticity of migration,” *American Economic Journal: Applied Economics*, 2017, *9* (2), 219–55.
- Beine, Michel and Christopher Parsons, “Climatic factors as determinants of international migration,” *The Scandinavian Journal of Economics*, 2015, *117* (2), 723–767.
- and Christopher R Parsons, “Climatic factors as determinants of international migration: Redux,” *CESifo Economic Studies*, 2017, *63* (4), 386–402.
- , Simone Bertoli, and Jesús Fernández-Huertas Moraga, “A practitioners’ guide to gravity models of international migration,” *The World Economy*, 2016, *39* (4), 496–512.
- Bertoli, Simone and Jesús Fernández-Huertas Moraga, “Multilateral resistance to migration,” *Journal of development economics*, 2013, *102*, 79–100.
- , Herbert Brücker, and Jesús Fernández-Huertas Moraga, “The European crisis and migration to Germany,” *Regional Science and Urban Economics*, 2016, *60*, 61–72.
- Biden, Joseph R., “Executive Order on Rebuilding and Enhancing Programs to Resettle Refugees and Planning for the Impact of Climate Change on Migration,” *Federal Register Volume 86, Issue 25*, February 2021, (14013).
- Black, Richard, Stephen RG Bennett, Sandy M Thomas, and John R Beddington, “Migration as adaptation,” *Nature*, 2011, *478* (7370), 447–449.
- Bohra-Mishra, Pratikshya, Michael Oppenheimer, and Solomon M Hsiang, “Non-linear permanent migration response to climatic variations but minimal response to disasters,” *Proceedings of the National Academy of Sciences*, 2014, *111* (27), 9780–9785.
- Boustan, Leah Platt, Matthew E Kahn, and Paul W Rhode, “Moving to higher ground: Migration response to natural disasters in the early twentieth century,” *American Economic Review*, 2012, *102* (3), 238–44.

- Burke, Marshall and Kyle Emerick**, “Adaptation to climate change: Evidence from US agriculture,” *American Economic Journal: Economic Policy*, 2016, 8 (3), 106–40.
- Cai, Ruohong, Shuaizhang Feng, Michael Oppenheimer, and Mariola Pytlikova**, “Climate variability and international migration: The importance of the agricultural linkage,” *Journal of Environmental Economics and Management*, 2016, 79, 135–151.
- Carleton, Tamma A and Solomon M Hsiang**, “Social and economic impacts of climate,” *Science*, 2016, 353 (6304), aad9837.
- Cattaneo, Cristina and Giovanni Peri**, “The migration response to increasing temperatures,” *Journal of Development Economics*, 2016, 122, 127–146.
- , **Michel Beine, Christiane J Fröhlich, Dominic Kniveton, Inmaculada Martinez-Zarzoso, Marina Mastrorillo, Katrin Millock, Etienne Piguet, and Benjamin Schraven**, “Human migration in the era of climate change,” *Review of Environmental Economics and Policy*, 2019, 13 (2), 189–206.
- Chudik, Alexander and M Hashem Pesaran**, “Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors,” *Journal of Econometrics*, 2015, 188 (2), 393–420.
- Dell, Melissa, Benjamin F Jones, and Benjamin A Olken**, “Temperature shocks and economic growth: Evidence from the last half century,” *American Economic Journal: Macroeconomics*, 2012, 4 (3), 66–95.
- , —, and —, “What do we learn from the weather? The new climate-economy literature,” *Journal of Economic Literature*, 2014, 52 (3), 740–98.
- Deschênes, Olivier and Michael Greenstone**, “The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather,” *American Economic Review*, 2007, 97 (1), 354–385.
- and —, “Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US,” *American Economic Journal: Applied Economics*, 2011, 3 (4), 152–85.
- Deser, Clara, Reto Knutti, Susan Solomon, and Adam S Phillips**, “Communication of the role of natural variability in future North American climate,” *Nature Climate Change*, 2012, 2 (11), 775–779.
- Fan, Qin, H Allen Klaiber, and Karen Fisher-Vanden**, “Does extreme weather drive interregional brain drain in the US? Evidence from a sorting model,” *Land Economics*, 2016, 92 (2), 363–388.
- Feng, Shuaizhang, Michael Oppenheimer, and Wolfram Schlenker**, “Climate change, crop yields, and internal migration in the United States,” Technical Report, National Bureau of Economic Research 2012.

- Friedman, Milton**, “The permanent income hypothesis,” in “A theory of the consumption function,” Princeton University Press, 1957, pp. 20–37.
- Gross, Emily**, “Internal revenue service area-to-area migration data: strengths, limitations, and current uses,” *Statistics of Income. SOI Bulletin*, 2005, 25 (3), 159–160.
- Hall, Robert E**, “Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence,” *Journal of Political Economy*, 1978, 86 (6), 971–987.
- Hawkins, Ed and Rowan Sutton**, “Time of emergence of climate signals,” *Geophysical Research Letters*, 2012, 39 (1).
- Hirvonen, Kalle**, “Temperature changes, household consumption, and internal migration: Evidence from tanzania,” *American Journal of Agricultural Economics*, 2016, 98 (4), 1230–1249.
- Hornbeck, Richard**, “The enduring impact of the American Dust Bowl: Short-and long-run adjustments to environmental catastrophe,” *American Economic Review*, 2012, 102 (4), 1477–1507.
- Hsiang, Solomon**, “Climate econometrics,” *Annual Review of Resource Economics*, 2016, 8, 43–75.
- Hsiang, Solomon M, Marshall Burke, and Edward Miguel**, “Quantifying the influence of climate on human conflict,” *Science*, 2013, 341 (6151), 1235367.
- IPCC**, *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, 2014.
- Kaczan, David J and Jennifer Orgill-Meyer**, “The impact of climate change on migration: a synthesis of recent empirical insights,” *Climatic Change*, 2020, 158 (3), 281–300.
- Kala, Namrata**, “Learning, Adaptation, and Climate Uncertainty: Evidence from Indian Agriculture,” *MIT Center for Energy and Environmental Policy Research Working Paper*, 2017, 23.
- Kolstad, Charles D and Frances C Moore**, “Estimating the economic impacts of climate change using weather observations,” *Review of Environmental Economics and Policy*, 2020, 14 (1), 1–24.
- Laczko, Frank, Christine Aghazarm et al.**, *Migration, Environment and Climate Change: assessing the evidence.*, International Organization for Migration (IOM), 2009.
- Lemoine, Derek**, “Estimating the Consequences of Climate Change from Variation in Weather,” *NBER Working Paper*, 2018, (w25008).
- Mahajan, Parag and Dean Yang**, “Taken by storm: Hurricanes, migrant networks, and us immigration,” Technical Report, National Bureau of Economic Research 2017.

- Marchiori, Luca, Jean-François Maystadt, and Ingmar Schumacher**, “The impact of weather anomalies on migration in sub-Saharan Africa,” *Journal of Environmental Economics and Management*, 2012, *63* (3), 355–374.
- McFadden, Daniel et al.**, “Conditional logit analysis of qualitative choice behavior,” 1973.
- Missirian, Anouch and Wolfram Schlenker**, “Asylum applications respond to temperature fluctuations,” *Science*, 2017, *358* (6370), 1610–1614.
- Molloy, Raven, Christopher L Smith, and Abigail Wozniak**, “Internal migration in the United States,” *Journal of Economic perspectives*, 2011, *25* (3), 173–96.
- Mueller, Valerie, Clark Gray, and Katrina Kosec**, “Heat stress increases long-term human migration in rural Pakistan,” *Nature climate change*, 2014, *4* (3), 182.
- Pachauri, Rajendra K, Myles R Allen, Vicente R Barros, John Broome, Wolfgang Cramer, Renate Christ, John A Church, Leon Clarke, Qin Dahe, Purnamita Dasgupta et al.**, *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*, Intergovernmental Panel on Climate Change, 2014.
- Pesaran, M Hashem**, “Estimation and inference in large heterogeneous panels with a multifactor error structure,” *Econometrica*, 2006, *74* (4), 967–1012.
- Pierce, Kevin**, “SOI migration data: a new approach: Methodological improvements for SOI’s United States Population Migration Data, calendar years 2011-2012,” *Statistics of Income. SOI Bulletin*, 2015, *35* (1).
- Schlenker, Wolfram and Michael J Roberts**, “Nonlinear temperature effects indicate severe damages to US crop yields under climate change,” *Proceedings of the National Academy of sciences*, 2009, *106* (37), 15594–15598.
- Taraz, Vis**, “Adaptation to climate change: Historical evidence from the Indian monsoon,” *Environment and Development Economics*, 2017, *22* (5), 517–545.
- Viswanathan, Brinda and KS Kavi Kumar**, “Weather, agriculture and rural migration: evidence from state and district level migration in India,” *Environment and Development Economics*, 2015, *20* (4), 469–492.
- White, Corey**, “The dynamic relationship between temperature and morbidity,” *Journal of the Association of Environmental and Resource Economists*, 2017, *4* (4), 1155–1198.

Figures

Figure 1: Estimates for Main Specification from Three IRS SOI Outcome Measures

Notes: Red dots denote point estimates. Gray-shading denotes 95% confidence intervals based on robust standard errors clustered by county-of-origin. The correlations between the numbers of migrating returns and exemptions is 0.991, while the correlation between migrating returns and aggregate income is 0.923. Outcomes are IHS migration ratios of each measure. The analyzed sample is 1990-2017 for tax returns and tax exemptions. It is 1992-2017 for aggregate income.

Figure 2: Limiting Temporal Variation - Filtered Time Series

Notes: Markers denote point estimates. Whiskers denote 95% confidence intervals based on robust standard errors clustered by county-of-origin. Estimates are based on high-pass filtered (Butterworth Filter) time series of temperature-bin counts and the IHS migration rate of tax returns. Such filtering leaves variation with periodicity in the ranges indicated in the time series, while removing variation with other periods of variation. Outcomes are filtered IHS tax return migration ratios and sample period is 1990-2017.

Figure 3: Long Differences Estimates

Notes: Markers denote point estimates. Shaded area denotes 95% confidence intervals based on robust standard errors clustered by county-of-origin. Estimates are based on the 30-year difference (the longest our sample will allow) of 5-year moving averages of the temperature-bin counts (and precipitation controls) and IHS migration rate of tax returns. Origin and destination fixed effects are included and standard errors are clustered by origin county. The coefficient estimates now capture the marginal effect of a change in the 5-year average number of days in each bin, and are thus much larger in magnitude than the estimated marginal effects of a change in the 1-year count of days in each bin. Outcome is the change in the 5-year average of the IHS tax return migration ratio from 1987 and 2017. The regression is based on 160,473 origin-destination observations.

Figure 4: Influence of Expectations Removed

Notes: Markers denote point estimates. Shaded area denotes 95% confidence intervals based on robust standard errors clustered by county-of-origin. Estimates are based on the Common Correlated Effects (CCE) method applied to our main results. Effectively, cross-sectional averages of the dependent and independent variables are interacted with the origin-destination fixed effects and included as independent variables in the regression. Origin and destination fixed effects are also included directly and standard errors are clustered by origin county. The coefficient estimates now capture the marginal effect of an additional day in each temperature bin relative to a day in the 60-70°F temperature range after the influence of such a realization acting through expectations regarding the future is removed. Outcome is the IHS tax return migration ratio and the sample runs 1990-2017.

Figure 5: Coefficients on Hot Temperature Days Over Time

Notes: Markers denote point estimates. Whiskers denote 95% confidence intervals based on robust standard errors clustered by county-of-origin. Period-specific estimates are taken from the interaction term between an indicator for each 3-year time period and the count variable of days with mean temperatures $>80^{\circ}\text{F}$. Time-period indicators are also included directly. Outcome is the IHS tax return migration ratio and the analyzed sample runs 1983-2017.

Figure 6: High-Temperature Coefficients by Norms in County-of-Origin

Notes: Markers denote point estimates. Whiskers denote 95% confidence intervals based on robust standard errors clustered by county-of-origin. Outcome is the IHS tax return migration ratio and the sample period is 1990-2017. Baseline grouping specific estimates are taken from the interaction term between an indicator for grouping (based on the average number of days with mean temperatures $>80^{\circ}\text{F}$ in the years from 1983-1989) and the count variable of days with mean temperatures $>80^{\circ}\text{F}$ in the two bin specification of the main model.

Figure 7: Temperature and Belief in Global Warming

Notes: Red dots denote point estimates. Gray-shading denotes 95% confidence intervals based on robust standard errors clustered by county. Outcome is share of county population that indicate their belief that global warming is happening as estimated by the Yale Climate Opinion Maps (available: <https://climatecommunication.yale.edu/visualizations-data>, accessed: 3/1/2020). Sample includes all survey years: 2014, 2016, 2018, and 2019. Precipitation controls are also included in both regressions. State time trends are linear in year, but results are similar with state by cubic-year time trends. Results are for temperature exposures in year prior to survey.

Figure 8: Estimates of Extreme Temperature Bins by NOAA Climate Zone

Notes: Markers denote point estimates. Whiskers denote 95% confidence intervals based on robust standard errors clustered by county-of-origin. Outcome is the IHS tax return migration ratio and the sample period is 1990-2017. Estimates are from a single regression with a vector of indicators for NOAA Climate Zone interacted with the temperature bin variables in the two bin specification. The fixed-effect and precipitation control scheme from main specification is used. A map of the NOAA Climate Zones is available here:

<https://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php>.

Tables

Table 1: Summary Statistics: 1990-2017 Sample Maintained by IRS

	Returns	Exemptions	Aggregate Income
Time Period:	1990-2017	1990-2017	1990-2017
Annual Mean Outmigration Rate:	0.0504	0.0439	0.0390
Mean Annual Outflow:	10,794	20,221	510,566
Mean Total (Movers and Non-Movers):	202,263	435,226	12,524,130

# Days Each Year	Mean	SD
<30	31.49307	31.93735
30-40	39.16017	23.97559
40-50	53.17385	19.87988
50-60	64.40277	19.35899
60-70	69.78972	15.37675
70-80	71.29063	29.77624
80-90	34.377	37.42157
>90	1.562778	6.686679
Total	365.25	

Notes: Outmigration statistics describe the 1990-2017 sample used in the analyses and maintained by the IRS. Weather statistics describe mean, annual origin-county conditions over this same period. Outmigration statistics represent means of annual outmigration totals by origin county. All annual origin-county means are weighted by the number of destination counties in the data for each origin county. There are 3,055 origin counties represented in the data each year with an average of 52.53 destination counties each.

Table 2: Short-Term Temperature Effects - Three Outcomes

	(1)	(2)	(3)
	Tax Returns	Exemptions	Aggregate Income
Annual Mean Temperature	-0.0000679 (0.000198)	-0.0000860 (0.000191)	-0.0000461 (0.000184)
<i>N</i>	4493236	4493236	4172290

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Mean temperature value is the average of monthly mean temperatures in the origin county. All specifications include origin-state by year fixed effects in addition to the origin-destination-pair fixed effects and precipitation controls of the main specification. Robust standard errors in parenthesis clustered by origin county. Each column is from a separate regression based on the outcome variable in the column headings. Estimates in column 3 are based on fewer observations because data on aggregate income data is only available beginning in 1992, while the samples for returns and exemptions begin in 1990. All samples run through 2017.

Table 3: Main Specification: Three Outcomes

	(1)	(2)	(3)
	Tax Returns	Exemptions	Aggregate Income
$< 30^{\circ}F$	-0.000789** (0.000327)	-0.00100*** (0.000316)	-0.00294*** (0.000346)
30-40 $^{\circ}F$	0.00264*** (0.000375)	0.00239*** (0.000360)	0.000319 (0.000397)
40-50 $^{\circ}F$	0.00133*** (0.000377)	0.00102*** (0.000361)	0.000391 (0.000359)
50-60 $^{\circ}F$	0.00223*** (0.000356)	0.00199*** (0.000344)	0.000618** (0.000313)
70-80 $^{\circ}F$	0.000661** (0.000275)	0.000550** (0.000265)	0.00136*** (0.000269)
80-90 $^{\circ}F$	0.00431*** (0.000337)	0.00412*** (0.000320)	0.00557*** (0.000342)
$> 90^{\circ}F$	0.00960*** (0.00145)	0.00910*** (0.00142)	0.00793*** (0.00126)
N	4493236	4493236	4172290

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Robust standard errors in parenthesis clustered by county-of-origin. All estimates are based on the main, seven-bin specification which includes origin-destination- pair fixed effects and precipitation controls. The outcome variable is the IHS migration ratio of the relevant measure in each column. Estimates in column 3 are based on fewer observations because data on aggregate income data is only available beginning in 1992, while the samples for returns and exemptions begin in 1990. All samples run through 2017. The correlations between the numbers of migrating returns and exemptions is 0.991, while the correlation between migrating returns and aggregate income is 0.923.

Table 4: High-Pass Filtered Estimates

	0-3 Years of Variation	0-6 Years of Variation	0-9 Years of Variation	0-12 Years of Variation	0-15 Years of Variation	Baseline (Unfiltered)
$< 30^\circ F$	-0.00341*** (0.000235)	-0.00829*** (0.000322)	-0.00990*** (0.000378)	-0.00973*** (0.000381)	-0.00911*** (0.000365)	-0.000870*** (0.000334)
$30-40^\circ F$	-0.00234*** (0.000259)	-0.00503*** (0.000345)	-0.00568*** (0.000395)	-0.00527*** (0.000400)	-0.00475*** (0.000389)	0.00268*** (0.000380)
$40-50^\circ F$	-0.000955*** (0.000284)	-0.00351*** (0.000378)	-0.00403*** (0.000443)	-0.00374*** (0.000445)	-0.00344*** (0.000423)	0.00121*** (0.000383)
$50-60^\circ F$	0.000549** (0.000271)	-0.000980*** (0.000266)	-0.00154*** (0.000307)	-0.00148*** (0.000311)	-0.00134*** (0.000302)	0.00217*** (0.000358)
$70-80^\circ F$	0.000298 (0.000209)	0.000596*** (0.000206)	0.000205 (0.000226)	0.0000486 (0.000228)	-0.0000124 (0.000226)	0.000680** (0.000286)
$80-90^\circ F$	0.000216 (0.000177)	0.00257*** (0.000184)	0.00296*** (0.000206)	0.00332*** (0.000215)	0.00366*** (0.000219)	0.00422*** (0.000340)
$> 90^\circ F$	-0.00197*** (0.000704)	0.00324*** (0.000652)	0.00441*** (0.000888)	0.00499*** (0.00104)	0.00536*** (0.00107)	0.00954*** (0.00147)
N	4493020	4493020	4493020	4493020	4493020	4493020

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Each column corresponds to a separate estimation of the main model with the IHS migration rate of tax returns and meteorological variables filtered with a high-pass, time-series filter (Butterworth Filter) using the periodicity indicated in the column titles. Such filtering leaves variation within the indicated periodicity. Robust standard errors in parenthesis clustered by county-of-origin. All estimates are based on the main, seven-bin specification which includes origin-destination- pair fixed effects and precipitation controls. The outcome variable is the IHS migration ratio of tax returns and the analyzed sample runs 1990-2017.

Table 5: Additional Results via Interaction

	(1)	(2)	(3)	(4)	(5)
	Over Time	CC Happening Yale Survy	Weather Variance	Share Over25 Finished HS	Per Capita Income (1000s)
$> 80^{\circ}F$	-1.161*** (0.0777)	-0.0204*** (0.00432)	0.00773*** (0.000631)	0.00627** (0.00294)	-0.000286 (0.000877)
$> 80^{\circ}F$ X Year	0.000582*** (0.0000389)				
$> 80^{\circ}F$ X CC Happening		0.000457*** (0.0000646)			
$> 80^{\circ}F$ X Hot Day MDEV			-0.000402*** (0.0000608)		
$> 80^{\circ}F$ X Share Fin HS				0.000330*** (0.0000897)	
$> 80^{\circ}F$ X income					0.000137*** (0.0000292)
N	4493236	641892	4493236	1123311	4439756

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Robust standard errors in parenthesis clustered by county-of-origin. Reported coefficients are from interaction terms of the regression-specific measure of interest with the $>80^{\circ}F$ temperature bin in the two-bin specification. The $<30^{\circ}F$, $>80^{\circ}F$, and regression-specific variable of interest all enter directly into the regressions, as do all the fixed effects and precipitation controls from the main specification. The outcome variable is the IHS migration ratio of tax returns and the analyzed sample runs 1990-2017.

Table 6: Coefficients by Origin County Types: 1990-2017

	(1) Corn Belt State	(2) Outside Corn Belt	(3) High Ag County	(4) Low Ag County	(5) Low Pop Density	(6) High Pop Density
$< 30^{\circ}F$	-0.00353*** (0.000177)	-0.000433 (0.000297)	-0.00263*** (0.000240)	-0.00190*** (0.000289)	-0.00314*** (0.000256)	-0.00157*** (0.000263)
$> 80^{\circ}F$	0.0103*** (0.000621)	0.00320*** (0.000274)	0.00569*** (0.000335)	0.00281*** (0.000363)	0.00339*** (0.000378)	0.00481*** (0.000388)
N	1145836	3347400	2241875	2250254	2215884	2276824

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Robust standard errors in parenthesis clustered by origin county. Each column is from a separate regression run on a separate sample. The main, two-bin regression specification is used in all cases. The outcome variable in all regressions is the IHS migration ratio of tax returns and the analyzed samples all run 1990-2017. Corn belt states include: Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin (following Feng et al., 2012). High Ag and Density counties are identified as those with above the median level of share-agricultural land and population-per-area values in a given year.

Table 7: Tax Return Migration by Origin Exposures and Destination Norms

		IHS Return Migration Ratio
# of Days in Origin County in Contemporaneous Year	$< 30^{\circ}F$	-0.00216*** (0.000185)
	$> 80^{\circ}F$	0.00410*** (0.000273)
Mean # of Days per Year in Destination County 1983-1989	$< 30^{\circ}F$	0.0768*** (0.0167)
	$> 80^{\circ}F$	-0.160*** (0.0144)
N		4493236

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Robust standard errors in parenthesis clustered by origin county. Regression includes origin-county fixed effects and county-specific time trends for both origin and destination counties. Destination baseline values are mean counts of days $< 30^{\circ}F$ and $> 80^{\circ}F$ in the destination county in the pre-analysis period: 1983-1989. The ratio of migrating to non-migrating Tax Returns is the main outcome in all regression.

A Appendix

A.1 Details of RUM to Gravity Model

We are interested in understanding how push factors – local temperatures in particular – drive internal migration within the United States during the last decades of the twentieth century and first decades of the twenty first. Such internal migration provides a unique opportunity to examine migration decisions which are unconstrained by the policy and administrative barriers which define international migration flows. In our empirical context, households are making a clear decision, weighing the advantages and disadvantages of staying put versus migrating to another location.¹

Using the notation of the Random Utility Model, in each period, t , we model all individuals in county j choosing which county k to live in for the period. Both j and k are from the choice set C , which, for this investigation, can be thought of as all counties in the United States (as we neither model nor analyze international moves). Non-migrators are modeled as selecting $k = j$, with utility in year t of:

$$u_{ijjt} = A_{jt} + \epsilon_{ijjt} \quad (\text{A.1})$$

For all others, $k \neq j$. Such migrators face moving costs, denoted M_{jkt} , and thus utility:²

$$u_{ijkt} = A_{kt} - M_{jkt} + \epsilon_{ijkt} \quad (\text{A.2})$$

In both expressions, A_{kt} represents the utility derived from the amenities of the county of residence for the period. This term is often referred to as the deterministic component of the utility, while ϵ_{ijkt} is the individual-specific stochastic term. Applying McFadden et al. (1973), we assume that ϵ_{ijkt} follows an iid, type-1 extreme value distribution which allows for the expected probability of an individual migrating in year t to be written as (expected value notation omitted for clarity):³

¹Our model and derivations are adapted from those of Beine and Parsons (2015) and Beine et al. (2016).

²Those familiar with dynamic choice models will note the silence of this approach with regard to implications of the present decision on future migration decisions. Consideration of the discounted expected values of optimal future strategies for staying and migrating can be added to this framework (see Beine et al., 2016, section 2.b.ii). We eschew such an extension here because of the requirement for explicit assumptions regarding expectations of future temperature realizations. Given the dynamic availability of local predictions of temperatures under climate change, this is an area for potential future work.

³This distributional assumption on the error term implies the independence from irrelevant alternatives, which buys us the simplicity of an analysis that depends only on the characteristics of the origin and destination. Dropping consideration of so-called “multilateral resistance” (the idea that changes in non-origin, non-destination locations impacts flows between the origin and destination) has long been standard in both the trade and migration literatures (see: Beine and Parsons, 2015). That said, we will be able to address concerns about multilateral resistance to some degree in our empirical estimation strategy.

$$Pr(u_{ijkt} = \max_k u_{ijkt}) = \frac{N_{jkt}}{Pop_{jt}} = \frac{\exp(A_{kt} - M_{jkt})}{\sum_l \exp(A_{lt} - M_{jlt})} \quad (\text{A.3})$$

We now assume that utility from county amenities does not vary by county of origin.⁴ By taking the ratio of stayers and leavers in a county, we can therefore write:

$$\frac{N_{jkt}}{N_{jjt}} = \frac{\exp(A_{kt} - M_{jkt})}{\exp(A_{jt})} \quad (\text{A.4})$$

This ratio will be referred to as the migration rate for the remainder of the paper. Taking natural logarithms of both sides yields:

$$\ln\left(\frac{N_{jkt}}{N_{jjt}}\right) = A_{kt} - A_{jt} - M_{jkt} \quad (\text{A.5})$$

Thus the logged ratio of migrators between two counties and the non-migrators in the origin county in a given period can be written as a linear function of the characteristics in the destination and origin locations and the costs of migrating from the origin to the destination. We now split amenities into weather and non-weather amenities, W_{jt} and N_{jt} , respectively.

$$\ln\left(\frac{N_{jkt}}{N_{jjt}}\right) = W_{kt} + N_{kt} - W_{jt} - N_{jt} - M_{jkt} \quad (\text{A.6})$$

Local weather can be modeled as a function of local climate and a mean-zero stochastic component: $W_{jt} = f(C_j, \varepsilon_{jt})$, where C_j characterizes local climate and $\varepsilon_{jt} \sim (0, \sigma_j^2)$. Following Lemoine (2018) we model local weather as the sum of local climate and a stochastic term: $W_{jt} = C_j + \varepsilon_{jt}$, yielding:

$$\ln\left(\frac{N_{jkt}}{N_{jjt}}\right) = C_k + \varepsilon_{kt} + N_{kt} - C_j - \varepsilon_{jt} - N_{jt} - M_{jkt} \quad (\text{A.7})$$

This equation is evaluated in expectation, and $E(\varepsilon_{kt}) = E(\varepsilon_{jt}) = 0$. Thus, only climate - and not weather deviations from local climate - impacts migration decisions:

$$\ln\left(\frac{N_{jkt}}{N_{jjt}}\right) = C_k + N_{kt} - C_j - N_{jt} - M_{jkt} \quad (\text{A.8})$$

This model assumes that any deviation in weather from long-term norms is idiosyncratic and mean zero. As noted by Lemoine (2018), he is explicitly interested in modeling a

⁴Assuming away any heterogeneity in the deterministic component of utility based on county of origin is the common way forward to a simple empirical specification, though Beine et al. (2016) note that this assumption isn't reflective of reality. A potential approach to working around this assumption is to include origin-time fixed effects in the empirical model, however this approach is not feasible in our investigation as that is the level of variation we are interested in investigating.

climate that is “stationary”. Under climate change however, deviations from long-term weather norms may not in fact be idiosyncratic with mean zero.⁵ We now extend the model to include a term, γ_{jt} , for the contribution of climate change to local weather, which can now be written as $W_{jt} = C_j + \gamma_{jt} + \varepsilon_{jt}$. We remain generally agnostic about the precise character of γ_{jt} , but note if temperature is the dimension of climate and weather of interest (as it is in this paper), one of the central aims of the Paris Climate Accord is to keep the global mean of $\gamma_{jt} < 1.5^\circ C$.

Now, local deviations in realized weather relative to long-term norms (hereafter simply “weather deviations”) can be written as: $\gamma_{jt} + \varepsilon_{jt}$. As the expectation of γ_{jt} is not necessarily zero, it cannot be dropped from the expansion of Equation A.7, leaving us with:

$$\ln \left(\frac{N_{jkt}}{N_{jjt}} \right) = C_k + \gamma_{kt} + N_{kt} - C_j - \gamma_{jt} - N_{jt} - M_{jkt} \quad (\text{A.9})$$

A.2 Data Notes and Sample Construction

Our main analyses rely on annual measures of county-to-county flows of households, individuals, and income which are derived from the Statistics of Income (SOI) data organized and maintained by the Internal Revenue Service (IRS). This data provides migration estimates based on year-to-year address changes of income tax filings and is available from the IRS for the years 1990-2017.⁶

To protect privacy, annual county-to-county flows below fixed thresholds are not reported (this threshold was 10 returns until 2013 and has been 20 since that time).

In addition to the number of returns which move in each direction between each county pair, the SOI data also reports the number of exemptions claimed and aggregate income (beginning in 1992) reported on both the moving and non-moving returns. IRS guidance points to the interpretation of return counts as a proxy for the number of households, while the number of exemptions proxies for the number of individuals that migrate (Gross, 2005). In addition to these proxy measures and the censoring issue noted above, the IRS SOI data is limited in that it only captures households (migrating or not) that file taxes before late-September (until 2011) or the end of December (2012 and onward) for the preceding tax year. As a result, the poor and elderly are likely underrepresented in the IRS SOI data – as members of these populations file tax returns at lower rates – and the extremely wealthy

⁵Given this is a model of decisions, what matters is perceived relationships, thus it is enough for there to be a belief in climate change if such a beliefs leads to the perception (and action based on that perception) that weather deviations include information about the future. In fact, another similar conceptualization, put forward by Hsiang (2016) is that weather deviations have direct effects and “belief effects”

⁶Additional information and raw data is available: <https://www.irs.gov/uac/soi-tax-stats-migration-data>, accessed February 6, 2020.

- which are more likely to have long-delayed filings on the basis of complication (Pierce, 2015).⁷

Weather conditions, temperatures and precipitation, are merged on to the migratory flows based on both origin and destination counties. Weather data is from NOAA’s Global Historical Climatology Network and assigned to counties based on the inverse distance weighted average from all monitors within 300km of the centroid of the relevant county. The temperature measures used in this analysis are annual counts of days with reported mean temperatures falling in eight distinct temperature bins in a given year for a given county. The bins capture counts of days with mean temperatures in 10-degree Fahrenheit ranges from 30°F to 90°F, with the remaining bins counting days <30°F and >90°F. These counts are calculated for each weather monitor in the data and then aggregated to the county level (following Dell et al., 2014). Total precipitation and mean daily temperatures are also calculated for each monitor and aggregated to the county level based on the same inverse distance weighting procedure. Ultimately, only data from origin counties in the continental United States are included in the analytic sample.

Earlier IRS SOI data is available through the University of Michigan ICPSR project.⁸ This data, back to 1983, is used for several analyses looking at changes in the relationships of interest over time, but our main analyses focus on the 1990-2017 sample that is maintained and made available directly by the IRS. Additional county-year data for controls and supporting analyses was taken from the Economic Profile by County data product (CAINC30) provided by the Bureau of Economic Analysis.

A.3 Adding Intertemporal Component

The absence of consideration of the future in the RUM, and resulting gravity models, has been noted previously by (Beine et al., 2016) and (Bertoli et al., 2016), among others. Our approach to introducing expectations regarding the future follows that of (Bertoli et al., 2016), as does our empirical approach to controlling for such expectations as a means of identifying the effects of other independent variables outside the influence of expectations.

We begin by modifying Equations 1 and 2 with terms allowing for consideration of expected future utility flows based on optimal location choices in periods $t + 1$ and after. This gives us:

⁷It is reported that 95% to 98% of filings are received by late September and an even higher share by the end of December. Nevertheless, the unobservability of returns that are filed after these dates likely leads to lower representation of the very rich in the data (Gross, 2005).

⁸Data available: <https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/2937>, accessed July 24, 2019.

$$u_{ijjt} = A_{jt} + \beta * V_{t+1}(j) + \epsilon_{ijjt} \quad (\text{A.10})$$

$$u_{ijkt} = A_{kt} - M_{jkt} + \beta * V_{t+1}(k) + \epsilon_{ijkt} \quad (\text{A.11})$$

Following the steps leading up to Equation A.12 in Appendix Section A.1, we generate the equivalent function for the probability an individual migrates from county i to county j in year t :

$$Pr(u_{ijkt} = \max_k u_{ijkt}) = \frac{N_{jkt}}{Pop_{jt}} = \frac{\exp(A_{kt} - M_{jkt} + \beta * V_{t+1}(k))}{\sum_l \exp(A_{lt} - M_{jlt} + \beta * V_{t+1}(l))} \quad (\text{A.12})$$

Taking the natural log of the ratio of those that migrate to those that don't based on this function leaves us with the analog to Equation A.13 with the terms for expectations regarding the future:

$$\ln\left(\frac{N_{jkt}}{N_{jjt}}\right) = A_{kt} - A_{jt} - M_{jkt} + \beta * [V_{t+1}(k) - V_{t+1}(j)] \quad (\text{A.13})$$

As in the main body of the paper, we can decompose the amenity streams into weather and non-weather, leaving us with Equation 10, which is the conceptual counterpart to our empirical mode.

$$Y_{jkt} = W_{kt} + N_{kt} - W_{jt} - N_{jt} - M_{jkt} + \beta [V_{t+1}(k) - V_{t+1}(j)] \quad (\text{A.14})$$

Our main estimation rolls the terms for future utility flows into the error term of the empirical counterpart to Equation A.14, resulting in a multi-factor error model akin to that presented by Pesaran (2006) for which the common correlated effects (CCE) approach is developed. We estimate as in our main analysis, except for the additional application of the the CCE approach developed by Pesaran (2006) and extended in Chudik and Pesaran (2015). Our implementation of the CCE estimation follows (Bertoli and Moraga, 2013) and (Bertoli et al., 2016) by adding a vector of "auxiliary regressors" to the regression model which is made up of the cross-sectional averages of the independent variable – for us, the migration rate of tax returns – and all independent variables – temperature bin day counts and high and low precipitation indicators – interacted with origin-destination pair indicator variables. Our CCE estimation is therefore:

$$Y_{jkt} = \sum_{b=1}^B \beta_b P_{bjt} + \mathbf{X}_{jkt} \boldsymbol{\mu} + \alpha_{jk} + \delta_{jt} + \mu_{kt} + \tilde{\mathbf{z}}_t \boldsymbol{\lambda}_{jk} + \epsilon_{jkt} \quad (\text{A.15})$$

where $\tilde{\mathbf{z}}_t$ is the vector of cross-sectional averages interacted with dyad-specific indicators.

As discussed by (Bertoli and Moraga, 2013), the CCE specification also removes multi-lateral resistance effects on origin to destination migration flows that arise from the failure of the independence of irrelevant alternatives assumption implicit in the type-1 extreme value distribution applied to the error term in the RUM model and discussed in footnote 6 of the main paper. As a result, the CCE-specification-derived estimates remove both the influence of changes in expectations regarding future temperature (and precipitation) conditions across all counties, but also any other influence of changing conditions in other counties beyond the origin and destination.

Additional Tables

Table A1: Year Fixed Effects

	(1)	(2)	(3)	(4)	(5)
	Baseline	Year FEs	Clim Zone by Yr FEs	State-by-year FEs	O&D State- by-Year FEs
$< 30^{\circ}F$	-0.000789** (0.000327)	-0.000565 (0.000450)	0.000443 (0.000356)	0.0000344 (0.000424)	-0.00000392 (0.000429)
$30-40^{\circ}F$	0.00264*** (0.000375)	-0.00137*** (0.000426)	-0.000554 (0.000349)	-0.000613 (0.000386)	-0.000824** (0.000390)
$40-50^{\circ}F$	0.00133*** (0.000377)	-0.00288*** (0.000519)	-0.00171*** (0.000307)	-0.000708* (0.000363)	-0.000825** (0.000369)
$50-60^{\circ}F$	0.00223*** (0.000356)	-0.00159*** (0.000462)	-0.00130*** (0.000292)	-0.00135*** (0.000336)	-0.00137*** (0.000336)
$70-80^{\circ}F$	0.000661** (0.000275)	-0.00102*** (0.000287)	-0.000458 (0.000320)	0.0000975 (0.000321)	0.000124 (0.000324)
$80-90^{\circ}F$	0.00431*** (0.000337)	0.0000911 (0.000284)	0.000715** (0.000348)	0.000488 (0.000373)	0.000525 (0.000378)
$> 90^{\circ}F$	0.00960*** (0.00145)	-0.00111 (0.00104)	-0.00158* (0.000926)	-0.00157 (0.00100)	-0.00139 (0.000899)
N	4493236	4493236	4493236	4493236	4493236

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Robust standard errors in parenthesis clustered by county-of-origin. Each column reports estimates from a separate regression based on the main specification with the addition of geography-specific year fixed effects at the noted geographic granularity. Column 1 reports estimates from the main specification without year fixed effects for easy comparison. Column 2 adds year fixed effects, column 3 replaces the year fixed effects with NOAA Climate Zone by year fixed effects. Column 4 uses state-by-year fixed effects and column 5 includes origin-county-by-year and destination-county-by-year fixed effects.

Table A2: Tax Returns: Building the Main Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Origin Cnty	Origin & Dest	Origin-to-Dest	Origin & Dest	Origin & Dest	Origin & Dest	Origin & Dest
	Fixed Effects	County FEs	Pair FEs	Time Trends	Quad Time Trends	Cubic Time Trends	Quartic Time Trends
< 30°F	-0.00255*** (0.000424)	-0.00255*** (0.000424)	-0.00254*** (0.000424)	-0.000789** (0.000327)	-0.00668*** (0.000317)	-0.00668*** (0.000318)	-0.00668*** (0.000318)
30-40°F	0.00384*** (0.000422)	0.00384*** (0.000422)	0.00384*** (0.000422)	0.00264*** (0.000375)	-0.00297*** (0.000352)	-0.00297*** (0.000352)	-0.00297*** (0.000353)
40-50°F	0.00164*** (0.000518)	0.00164*** (0.000518)	0.00164*** (0.000518)	0.00133*** (0.000377)	-0.00233*** (0.000349)	-0.00234*** (0.000349)	-0.00234*** (0.000350)
50-60°F	0.00359*** (0.000546)	0.00359*** (0.000546)	0.00359*** (0.000546)	0.00223*** (0.000356)	-0.000585** (0.000278)	-0.000584** (0.000278)	-0.000584** (0.000279)
70-80°F	-0.000417 (0.000350)	-0.000417 (0.000350)	-0.000417 (0.000350)	0.000661** (0.000275)	-0.000464** (0.000228)	-0.000464** (0.000228)	-0.000464** (0.000228)
80-90°F	0.00256*** (0.000397)	0.00256*** (0.000397)	0.00256*** (0.000397)	0.00431*** (0.000337)	0.00429*** (0.000267)	0.00429*** (0.000267)	0.00428*** (0.000267)
> 90°F	0.00779*** (0.00141)	0.00779*** (0.00141)	0.00779*** (0.00141)	0.00960*** (0.00145)	0.00604*** (0.00111)	0.00604*** (0.00111)	0.00604*** (0.00111)
N	4493236	4493236	4493236	4493236	4493236	4493236	4493236

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Robust standard errors in parenthesis clustered by origin county. Each column is from a separate regression. Estimates in the first three columns do in fact differ, but not in the decimal places shown. Column 4 adds county-specific linear time trends (in year) for BOTH origin and destination counties. Column 5 adds county-specific quadratic time trends for BOTH origin and destination counties.

Table A3: Robustness of Main Specification: Tax Returns

	(1)	(2)	(3)	(4)	(5)
	Untransformed Migration Rate	LN((Migratants+1)/ Non-Migrants)	LN(Population) Weights	Population Weights	PPML
main					
< 30°F	-0.146*** (0.0164)	-0.00111*** (0.000285)	-0.00103*** (0.000344)	-0.00277*** (0.000618)	-0.00224*** (0.000128)
30-40°F	0.0149 (0.0160)	0.00178*** (0.000342)	0.00243*** (0.000393)	0.000240 (0.000617)	0.000227** (0.000108)
40-50°F	-0.0446*** (0.0137)	0.00138*** (0.000339)	0.00118*** (0.000397)	-0.0000660 (0.000572)	-0.00115*** (0.000111)
50-60°F	0.0541*** (0.0135)	0.00191*** (0.000300)	0.00200*** (0.000372)	0.0000667 (0.000584)	0.000724*** (0.000125)
70-80°F	0.00372 (0.0153)	0.000601** (0.000271)	0.000538* (0.000289)	-0.0000962 (0.000511)	-0.0000758 (0.0000976)
80-90°F	0.186*** (0.0162)	0.00317*** (0.000324)	0.00403*** (0.000358)	0.00153** (0.000599)	0.00206*** (0.000102)
> 90°F	0.279*** (0.0708)	0.00706*** (0.00108)	0.00849*** (0.00139)	0.00310*** (0.00107)	0.00419*** (0.000438)
<i>N</i>	4493236	4479512	4439756	4439756	4367936

* $p < .1$, ** $p < .05$, *** $p < .01$.

Notes: Robust standard errors in parenthesis clustered by origin county. Each column is from a separate regression. The 1990-2017 sample of tax return flows is used in all cases. In the first column, the dependent variable is this migration rate in levels (i.e.: not transformed by the inverse hyperbolic sine function), and the main regression specification is used. Column 2 again reproduces the main specification but uses the natural log of the ratio of migrant returns (plus 1) to non-migrants as the dependent variable. Columns 3 and 4 exactly reproduce the main baseline specification but add observation weights of the natural log of county population and county population respectively. County populations are BEA estimates based on Census data. Column 5 re-estimates the main specification based on the Poisson Pseudo-Maximum Likelihood estimation procedure rather than OLS.