ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 031/26

Migration, Climate Similarity, and the Consequences of Climate Mismatch

Marguerite Obolensky, Marco Tabellini, Charles A Taylor

www.rfberlin.com JANUARY 2026

Migration, Climate Similarity, and the Consequences of Climate Mismatch

Authors

Marguerite Obolensky, Marco Tabellini, Charles A Taylor

Reference

JEL Codes: J15, J61, N31, N32, Q54, R11

Keywords: Migration, climate, climate matching, economic geography

Recommended Citation: Marguerite Obolensky, Marco Tabellini, Charles A Taylor (2026): Migration, Climate Similarity, and the Consequences of Climate Mismatch. RFBerlin Discussion Paper No. 031/26

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Migration, Climate Similarity, and the Consequences of Climate Mismatch*

Marguerite Obolensky, Marco Tabellini, and Charles A Taylor†

January 2026

Abstract

This paper examines the concept of "climate matching" in migration—the idea that migrants seek out destinations with familiar climates—and studies its implications for the geography of economic activity in the United States. We document that temperature distance between origin and destination predicts the distribution of migrants across U.S. counties, for both internal and international migration in the historical (1850–1940) and modern (1970–2019) periods. These patterns cannot be explained by the spatial correlation of climate or the persistence of ethnic networks, and instead reflect two mechanisms: the transferability of climate-specific skills and climate as an amenity. We then study the economic consequences of climate mismatch during 1880–1920, a period of rapid growth and structural transformation. Using an instrumental variable strategy that interacts origin-country inflow shocks with the timing of county railroad access, we find that mismatch reduced agricultural productivity and accelerated the exit from farming. However, manufacturing output did not rise. Instead, manufacturing productivity declined and population growth was lower in counties with higher climate mismatch. These effects left a lasting imprint: a 1°C increase in 1880–1920 mismatch is associated with 2.5% lower per capita income in 1940.

Keywords: Migration, climate, climate matching, economic geography.
JEL Codes: J15, J61, N31, N32, Q54, R11.

*We thank the editor, Nathan Nunn, six anonymous referees, Ran Abramitzky, David Albouy, Marcella Alsan, Sam Bazzi, Chris Boulos, Tamma Carleton, Paola Conconi, Leonardo D'Amico, Catherine De Vries, Tatyana Deryugina, Christian Dustmann, Silvia Farina, James Feigenbaum, Martin Fiszbein, Nicola Gennaioli, Paola Giuliano, Edward Glaeser, James Hammitt, Gordon Hanson, Joe Henrich, Andrew Hultgren, Peter Huybers, Kelsey Jack, Namrata Kala, William Kerr, Eliana La Ferrara, Ross Mattheis, Carlo Medici, Eoin McGuirk, Tom Nicholas, Elias Papaioannou, Maria Petrova, Jonathan Proctor, Paola Sapienza, Wolfram Schlenker, Ambra Seck, Marlon Seror, Jesse Shapiro, Andrei Shleifer, Clémence Tricaud, Felipe Valencia, Nico Voigtlaender, Matt Weinzierl, Noam Yuchtman, Guglielmo Zappalà, Richard Zeckhauser, and participants at conferences and seminars for useful comments. Giovanni Besana, Michael Cleary, Alessio Corò, Francesco Farsetti, Flavia Ginefra, Marta Leva, Emanuele Licari, Alessandra Nicolò, Marco Panunzi, Monia Tomasella, Francesco Toti, Elena Verzellesi, Hao Yang, and Zhengyang Zhou provided excellent research assistance. All remaining errors are ours. A previous version of the paper circulated with the title "Homeward Bound: How Migrants Seek Out Familiar Climates."

†Obolensky: Kellogg School of Management, Northwestern University (marguerite.obolensky@kellogg.northwestern.edu); Tabellini: Harvard Business School, NBER, CEPR, CReAM, and IZA (mtabellini@hbs.edu); Taylor: Harvard Kennedy School, Harvard University, and NBER (ctaylor@hks.harvard.edu).

"[New England is] more suitable to the nature of our people, who neither finde content in the colder Climates, nor health in the hotter; but (as hearbs and plants) affect their native temperature, and prosper kindly no where else." – Sir Ferdinando Gorges, English founder of the Province of Maine, 1622 (Kupperman, 1984)

1 Introduction

Immigrants tend to cluster geographically within the countries they settle (Altonji and Card, 1991; Card, 2001), leaving persistent and visible spatial imprints—from Norwegians in Minnesota to Vietnamese along the Gulf Coast of Texas to Oklahomans in California after the Dust Bowl. What explains the emergence of these distinctive settlement geographies? Well studied factors include economic opportunities at destination (Borjas, 2001; Cadena and Kovak, 2016), co-ethnic networks (Munshi, 2003; Stuart and Taylor, 2021), geographic frictions such as distance and transportation costs (Schwartz, 1973; Steckel, 1983), and political or cultural sorting (Bishop and Cushing, 2009; Bazzi et al., 2020). These sorting forces can in turn shape long-run economic development, sectoral specialization, spatial inequality, and political ideology.

Despite frequent historical references, another potential driver of migrant sorting has received comparatively little systematic attention: climate similarity between origin and destination. In 1925, U.S. President Calvin Coolidge observed that "the newcomers from Europe commonly sought climatic conditions here [in the United States] like those in which they had been raised. So the Scandinavians are found chiefly in the northern parts of this country" (Coolidge, 1926, p. 255). Similar arguments have been made about the early European settlement of the Americas (Fischer, 1989; Crosby, 2004) and U.S. internal migration during the 19$^{\text{th}}$ century (Steckel, 1983).[1] Yet we lack systematic evidence on how important climate similarity is for migration independent of distance and networks, and whether it matters for economic outcomes.

In this paper, we examine the relevance of climate matching in migration and study its implications for the geography of economic activity in the United States. We analyze migration flows spanning 1850–2019, covering both internal and international movements, to establish the role of climate similarity in shaping migrant settlement. We then assess the economic consequences of climate mismatch focusing on the period 1880–1920, a critical juncture in the structural transformation of the U.S. economy from agriculture to manufacturing. The U.S. is an ideal setting to study the role of climate preferences on migration: its vast range of climate zones

[1] Steckel (1983) noted that a "farmer contemplating a move sought, other things being equal, a location that maximized the return on previous investments in human capital; namely, a place where the climate, soil, and terrain were familiar."

provides migrants with meaningful climatic choice, and its relative lack of internal—and, historically, international—mobility restrictions reduces confounding policy frictions.[2] Moreover, rich historical and modern datasets allow us to measure climate similarity and migrant location choices with high spatial precision, and to link these patterns to subsequent spatial development during a formative period of U.S. economic history.

We begin by documenting a strong positive relationship between climate similarity and migration, which is displayed in Figure 1. Using information from the 1880 full count U.S. Census, the figure shows that the temperature in the county of residence of immigrants correlates with the temperature in their country of origin. That is, immigrants from warmer (colder) countries settled in similarly warmer (colder) parts of the U.S. We formalize this motivating evidence by estimating gravity models common in the trade literature (Anderson and Van Wincoop, 2003), where migration flows depend on bilateral climate distance, physical distance, and destination and origin-specific time-varying factors.

We define climate distance as the absolute difference in average annual temperature and precipitation at origin and destination, and consider international and domestic migration across both historical and modern periods. For international migration, we use full count Census data from the Age of Mass Migration (1880–1920), when over 30 million Europeans moved to the U.S., and modern samples following the 1965 Immigration and Nationality Act, when flows from Latin America and Asia surged. For domestic migration, we rely on linked historical census records from 1850 to 1940—a period marked by the U.S. westward expansion—and on recent administrative tax data capturing county-to-county moves from 2011 to 2019.

In each setting, climate similarity strongly predicts migration: reducing temperature distance by 1°C increases migration flows by 2% to 27%, depending on the context. These magnitudes are comparable to standard estimates of the elasticity of migration with respect to wages.[3] A 5°C reduction in temperature distance—roughly the difference between Chicago and Washington, DC, or between London and Rome—implies the same migration response as reducing physical distance by about 300 to 750 kilometers. These magnitudes suggest that climate acts as a "shadow border" significantly constraining movements between and within countries.

Precipitation similarity also predicts migration flows, but estimates tend to be noisier—consistent with evidence that temperature is a more systematic predictor of human and economic outcomes, whereas precipitation effects tend to be context specific (Carleton and Hsiang, 2016). We thus use overall temperature distance as our preferred measure of climate similarity, while

[2] Although the movement of white individuals within the U.S. was largely unregulated until the 1920s, formal and informal restrictions severely limited the movement of African Americans and Native Americans (Alston and Ferrie, 1999; Nichols, 2014).

[3] Migration–wage elasticity estimates range from 0.5 in cross-country settings (Caliendo et al., 2021) to 1.5 in internal migration contexts such as China (Tombe and Zhu, 2019), and up to 4.5 in Brazil (Morten and Oliveira, 2024).

always controlling for precipitation distance. As a sufficient statistic for the broader climatic environment migrants expect to face, temperature is salient, easy to observe, and strongly correlated with many features relevant for productivity, comfort, and cultural familiarity.

We then go beyond cross-sectional variation and exploit long-run changes in average climate throughout the 20$^{\text{th}}$ century, driven by multi-decadal oceanic cycles (e.g., North Atlantic Oscillation) and anthropogenic climate change, to isolate the impact of changes in climate distance on changes in migration patterns. We ask whether migration increased between two locations that converged climatically over time. This exercise allows us to control for any origin-destination time invariant specific factor, including past networks. This design also helps address concerns that our results might reflect persistent cultural, institutional, or economic similarities between locations correlated with both climate and migration. We find that changes in climate distance from 1900 to 2020 predict changes in both internal and international migration over the same period—with magnitudes close to those from the baseline specification.

The long-difference design is particularly useful in assessing the role of ethnic networks. If early migrants settled in climatically similar locations and later flows simply followed those co-ethnic pioneers, then observed climate matching could reflect persistent enclaves rather than climate-based preferences. The long-difference framework controls for all time-invariant bilateral effects, including the influence of such networks. For these networks to drive our results, their strength would have to co-evolve with climate similarity—rising or falling in tandem over time. While possible in theory, this seems unlikely in practice, especially given additional evidence from two complementary analyses. First, we replicate our baseline specifications while controlling directly for lagged migrant stocks. As in the existing literature (Card, 2001; Munshi, 2003), these past networks exert a positive influence on subsequent flows, yet the estimated effects of temperature distance remain virtually unchanged. Second, we show that the results hold in a setting where ethnic networks were likely absent: newly settled U.S. frontier counties during the 19$^{\text{th}}$ century (Bazzi et al., 2020). Together, these findings suggest that networks alone are unlikely to account for the climate matching patterns we observe.

Beyond migrant networks, we consider two mechanisms that help explain climate matching in migration. The first is climate-specific human capital: migrants may prefer climates akin to where they acquired relevant skills, especially in agriculture and other climate-exposed sectors. Consistent with this idea, we find larger effects for earlier periods (when agriculture dominated employment) and for individuals working in outdoor or climate-intensive occupations. The second mechanism is climate as an amenity. Climate similarity may enhance well-being directly or sustain familiar cultural practices, from food preparation and religious rituals to recreation and social interaction. Supporting this interpretation, climate matching persists among individuals in indoor occupations and remains significant in recent decades, when productivity-related concerns are less relevant. At the same time, we find that the relationship between origin and destination temperatures weakens in hotter U.S. destinations after 1970 when looking at summer temperatures—but not winter temperatures. This is consistent with the diffusion of air conditioning, which mitigates sensitivity to extreme heat, but not cold. Although we cannot isolate the relative contribution of these two forces, the evidence indicates that both climate-specific skills and amenity-based preferences matter—and may reinforce one another.

A natural concern with these patterns is that climate varies smoothly over space, raising the possibility that our estimates may capture residual spatial correlation rather than migration responses to climate similarity per se. This concern is limited for international migration, where long-distance moves break the link between climate and geography: physical distance explains at most 1.6% of the variation in temperature distance. The issue is potentially more salient for U.S. internal migration, where geographic and climatic distances are mechanically correlated. Even in this setting, however, geographic distance explains at most 20% of the variation in temperature distance, leaving substantial independent variation to be exploited. To address this concern directly, our baseline specification controls for geographic distance between origin–destination pairs. Moreover, we subject the analysis to a demanding set of additional exercises designed to absorb local spatial dependence. These include allowing for nonlinear distance effects, separating north–south from east–west movements, controlling directly for latitude and longitude distance, excluding short-range moves and within-state migration, dropping potentially idiosyncratic destinations such as Florida and California, and accounting for time-varying travel frictions linked to railroad expansion.

Across these checks, the evidence for historical internal migration is particularly robust: climate matching remains strong and statistically significant under all spatial specifications we consider. For modern internal migration, the relationship is somewhat smaller but persists across the majority of specifications, except in exercises that isolate vertical movements or allow for nonlinear distance effects. This contrast aligns with our broader evidence that climate matching has weakened over time, as technological change and occupational reallocation have reduced the importance of climate-specific constraints—both in skills and amenities—in shaping migration decisions. Finally, we show that both international and internal migration results are robust to controlling for a wide range of bilateral geographic characteristics correlated with climate (elevation, ruggedness, soils, and coastal access), as well as to using alternative definitions of climate similarity, including extremes and seasonality.

Having established the relevance of climate matching for migrant location choices, we next examine its implications for the economic geography of the U.S. We focus on international immigration between 1880 and 1920—a period of rapid growth and structural transformation (Eckert and Peters, 2025)—and ask whether counties receiving migrants from climatically more distant origins experienced slower economic trajectories. The effects of climate mismatch are *ex ante* ambiguous. On the one hand, mismatch can reduce productivity in agriculture, where climate-specific skills are very important. It can also lower productivity in other sectors, both directly—as some non-agricultural activities are also climate-dependent—and indirectly—through input-output linkages. On the other hand, greater mismatch could accelerate structural change if migrants less suited to local agricultural conditions moved more quickly into manufacturing, the main engine of growth at the time (Eckert and Peters, 2025). The overall effect is thus an empirical question.

To test these ambiguous predictions, we construct an index of climate mismatch measuring the average temperature distance of immigrants in each county-decade. The main empirical challenge is that immigrants' location choices may correlate with unobserved county-level trends. For instance, faster-growing counties may attract migrants regardless of their climate fit—because climate-specific skills are less important there—creating a spurious positive correlation between climate mismatch and subsequent productivity growth. To address this issue, we identify plausibly exogenous variation in climate mismatch by interacting push shocks—aggregate emigration flows from each European country—with pull shocks—the timing of railroad expansion across U.S. counties. The approach follows Sequeira et al. (2020) but exploits the interaction of country-specific shocks over time, in the spirit of Terry et al. (2024). We then estimate 2SLS regressions at the county level between 1880 and 1920 that include county and region-by-decade fixed effects, and account for railroad expansion to mitigate concerns about the exclusion restriction.

We find that climate mismatch led to a marked contraction of the agricultural sector. A 1°C increase in temperature mismatch lowered the value of crops and the land in farms by 16.3% and 4.2%, respectively. This implies that increasing climate mismatch from the 25th to the 75th percentile of the distribution—roughly a 4°C difference, comparable to the gap between average temperatures in Rome and Paris—reduces county-level crop value by 50.9%, land in farms by 16.9%, and farm value per acre by 44.5%. The latter is similar in magnitude to the decline in farmland values experienced by high-erosion counties during the Dust Bowl (Hornbeck, 2012).

As agriculture contracted, the non-agricultural employment share increased. The reallocation away from agriculture implies that the marginal product of labor in manufacturing should decline. Consistent with this prediction, growth in output per worker was 34% lower in a high-mismatch county relative to a low-mismatch county. Yet, we find no evidence that manufacturing output increased; the point estimate is negative (though imprecise), suggesting that mismatch may have depressed productivity in manufacturing both directly (through the sector's climate exposure) and indirectly (through weaker agricultural performance and reduced input–output linkages). We show that a two-sector model with climate-dependent productivity yields predictions that are consistent with these empirical patterns.

The impacts of climate mismatch extend to broader measures of economic development: increasing mismatch from the 25th to the 75th percentile reduces population growth by about 35% from 1880 to 1920. This magnitude is comparable to the difference in population growth between Cook County, IL, one of the industrial centers of the era, and Dakota County, NE, a slow-growing rural county. Evidence from the 1940 full count Census—the first to consistently report income—points to similar long-run outcomes. Residents of counties more exposed to mismatch during 1880–1920 had lower income per capita and hourly earnings in 1940. This relationship holds for both U.S.-born and immigrant residents and is quantitatively meaningful: residents of a county at the 75th percentile of the mismatch distribution were, on average, 10.2% poorer than those in a 25th percentile county. The persistence of these effects into 1940 suggests that climate mismatch is not just a temporary friction, but rather a durable constraint on development.

These results are robust to a wide range of controls and alternative specifications. First, we derive the climate mismatch index by predicting immigrant flows using weather shocks in European sending countries, as in Sequeira et al. (2020). This addresses concerns that migrants from colder (or warmer) origins may have arrived during periods when the railroad network was already expanding into similarly cold (or warm) counties. Second, we estimate long-difference regressions to reduce concerns about bias in two-way fixed effects models with heterogeneous treatment timing or differential trends (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021). Third, we replace region-by-decade with state-by-decade fixed effects to absorb finer spatial heterogeneity. Fourth, we exclude counties with extreme mismatch values. Fifth, we allow for differential trends by interacting decade dummies with the 1880 immigrant share and other county characteristics such as land productivity, frontier exposure, and the initial non-agricultural employment share. Finally, we show that the income results are not driven by farmers, for whom income reporting in the 1940 Census is known to be incomplete. We describe these and additional robustness checks below.

Together, our results provide systematic empirical support for the long-standing idea that historical migration patterns were shaped by climate similarity (Coolidge, 1926; Steckel, 1983; Crosby, 2004). This perspective features prominently in historical accounts of U.S. settlement patterns, notably in Fischer (1989), who argues that enduring regional cultures emerged from the environmental and institutional conditions faced by early settlers. In linking climate matching to economic outcomes, we contribute to a broad literature on the spatial determinants of growth and structural transformation (Eckert and Peters, 2025; Nagy, 2023; Desmet et al., 2025). A large body of work shows that immigration can shape regional development by affecting labor supply, local demand, and sectoral specialization (Sequeira et al., 2020; Peters, 2022;

Terry et al., 2024). We add to this literature by introducing climate similarity as a distinct and economically meaningful dimension of migrant sorting which affected not only where migrants settled but also how local economies evolved.

Our analysis also relates to work emphasizing the role of climate-specific skills and environmental conditions in shaping productivity and development. Michalopoulos (2012) shows that geographic heterogeneity gave rise to location-specific human capital and persistent spatial patterns in ethnolinguistic diversity, while Fenske (2014) links ecological diversity to trade and pre-colonial state formation. Bazzi et al. (2016) show that agroclimatic similarity between origin and destination predicts farmer productivity in Indonesia. Pellegrina and Sotelo (2025) document that Brazil's internal migrants carried origin-specific agricultural knowledge that raised productivity and shifted regional specialization. We show that a similar mechanism operated at scale in the United States, with effects extending to manufacturing. This perspective aligns with a broader literature on environmental factors and agricultural development (Hornbeck, 2010, 2012; Hornbeck and Naidu, 2014) and complements evidence that agricultural conditions shape long-run growth (Fiszbein, 2022). By connecting these strands, we provide evidence that climate-based migrant sorting shaped U.S. regional development.

In addition, our paper speaks to recent work that uses hedonic methods to estimate the amenity value of climate, typically finding that individuals prefer milder temperatures and avoid climatic extremes (Albouy et al., 2016; Sinha et al., 2021). These studies generally assume homogeneous climate preferences across migrants from different origins. A notable exception is Albouy et al. (2021), who document that migrants tend to move to cities with features similar to those of their countries of origin. We complement this work in several ways. First, we provide systematic evidence on climate similarity as a determinant of migration, using mean annual temperature as a transparent and widely applicable metric. Second, we show that this relationship is robust across a range of empirical settings, climate variables, and spatial controls. Third, we explore the mechanisms behind climate matching—highlighting the roles of climate-specific skills and amenities. Fourth, we examine the economic implications of climate mismatch, showing that climate matching had tangible consequences for U.S. development at the turn of the 20$^{\text{th}}$ century.

Our findings can inform the vast literature that, since seminal work by Altonji and Card (1991) and Card (2001), has used shift–share research designs to study the local effects of immigration (Boustan, 2010; Tabellini, 2020; Bazzi et al., 2023). A central challenge in this literature is that the initial migrant shares used as exposure weights are typically treated as given, even though the origins of these initial settlement patterns are not well understood.[4] By documenting the role of climate similarity in early settlement, our results provide a micro-foundation for the

[4] For a formal discussion of shift-share research designs, see Jaeger et al. (2018), Adão et al. (2019), Goldsmith-Pinkham et al. (2020), and Borusyak et al. (2022).

initial shares and offer a basis for instrumenting for enclave formation using climate-based predictors such as temperature distance. This intuition echoes the insight of Hunt (1992), who used climatic variation to predict the settlements of repatriates from Algeria to France in 1962.

Finally, our results are relevant to the growing literature on the global impacts of climate change. Recent papers seek to understand the effects of climate shocks on the global economy through general equilibrium models that assume that the migration costs underlying individuals' location and occupational choices are exogenous to climate change (Cruz and Rossi-Hansberg, 2023; Desmet and Rossi-Hansberg, 2024; Bilal and Rossi-Hansberg, 2025). Our evidence indicates that migrants place a large weight on climate distance—effectively a travel cost endogenous to the spatial distribution of climate change—and that climate mismatch can have persistent negative economic consequences for receiving places. As a result, the welfare losses from climate change may be larger than those implied by models that assume frictionless spatial sorting.

2 Climate Matching and Migration

2.1 Data and Measurement

This section describes the main data sources and variables used in the analysis. Table B.1 provides the complete list.

Climate. We use mean annual temperature (in degrees Celsius) and precipitation (in millimeters) from TerraClimate (Abatzoglou et al., 2018), a global gridded dataset widely used in climate research, to characterize local climate. For U.S. counties, we define climate using long-run normals (1960–2000) to capture the persistent climatic environment migrants expect to face. County-level values are population-weighted averages of temperatures within each county, with gridcell-decade-level population from Goldewijk et al. (2017). Thus, while the underlying climate data are fixed over 1960–2000, the aggregation reflects historical population patterns at each point in time, so any time variation arises from population weights rather than climate change.

For countries outside the United States, we take the 1960–2000 average of annual temperature and precipitation across all gridcells within national borders, weighted by population in each decade (Goldewijk et al., 2017). This approach maximizes spatial precision given the absence of sub-national origin information in historical censuses. We verify robustness to using climate averages over gridcells within a 25 km radius of the country's capital city. For robustness, we also use climate data from the Climatic Research Unit (Harris et al., 2020), which extends back to 1901. This data product is interpolated from weather observations without any underlying atmospheric or hydrological model, which can create spatial or temporal data gaps reflecting weather station availability.

Figure A.1 displays the spatial distribution of average temperature across U.S. counties. Figure 2 shows the temperature distance between each U.S. county and four major immigrant-sending countries, while Figure A.2 plots temperature distances between each U.S. county and selected counties in the Mississippi Delta (Panel A) and the Northeast (Panel B).

International migration. Our analysis covers the two major waves of U.S. immigration (Abramitzky and Boustan, 2017). During the Age of Mass Migration (1850–1920), over 30 million Europeans immigrated to the United States, initially from Northern and Western Europe and later increasingly from Southern and Eastern Europe as transportation costs fell (Hatton and Williamson, 1998; Abramitzky and Boustan, 2017). Immigrants tended to concentrate in distinct regions—Scandinavians and Germans in the Midwest; Italians in the Northeast, Mid-Atlantic, and California. Immigration was overwhelmingly European during this period (Figure A.3, Panel A), and few migrants settled in the U.S. South, with notable exceptions like Mexicans in Texas (Figure A.4, Panel A). Canadians tended to settle in northern border states.

After World War I, restrictive quota laws sharply curtailed inflows, and the immigrant share of the U.S. population fell from 14% in 1920 to 5% by 1960 (Figure A.3, Panel B). The 1965 Immigration and Nationality Act reversed these restrictions and ushered in a new era (King, 2002): by 2010, immigrants came predominantly from Latin America and Asia rather than Europe, and their settlement patterns had shifted toward the South and West (Figure A.4, Panel B).[5]

For the Age of Mass Migration, we use full count U.S. Census data from 1880 to 1920 (Ruggles et al., 2021).[6] For the modern period, we use microdata from the U.S. Census and the American Community Survey (ACS) from 1970 to 2010. In both cases, we aggregate data at the country-of-origin by U.S. county-of-residence by decade. For all census years from 1900 onward, we identify immigrants based on responses to the year-of-arrival question, and restrict the sample to those who arrived in the previous decade. Since this question is not available in the 1880 Census, we include all foreign-born individuals in that year. Accordingly, the 1880 observation reflects a stock measure rather than a decadal arrival flow, and our results are unchanged when excluding 1880.

We follow the harmonization approach of Burchardi et al. (2019), who standardize countries and counties to their 1990 boundaries. For compatibility with modern administrative data, we

[5] See Table A.1 for the top-10 immigrant-sending countries during the two immigration eras.

[6] The 1890 Census was destroyed in a fire. For this reason, we exclude this decade from all analyses.

adapt this procedure to map all counties to their 2010 definitions.[7]

Internal migration. As for international immigration, we consider both historical and modern internal migration. For the historical period, we focus on 1850 to 1940. This period was characterized by the westward expansion of white settlement (Steckel, 1983; Bazzi et al., 2020), the diffusion of railroads (Fogel, 1964; Donaldson and Hornbeck, 2016), and the country's transition from a rural to an urban and industrial economy (Eckert et al., 2023). As a result of this process, the population center of gravity shifted from the East to the Midwest (Figure A.5).[8]

We measure historical internal migration using linked individual Census records from 1850 to 1940 available through the Census Tree Project (Price et al., 2021; Buckles et al., 2025).[9] Harmonized addresses for over 94% of the sample are obtained from the Census Place Project (Berkes et al., 2023). For the modern period, we use IRS Change-of-Address Tables (available for 2011–2022). Our baseline analysis uses 2011–2019 annual county-to-county flows, which draw upon year-to-year address changes reported on individual income tax returns.[10] To avoid any distortions created by the COVID-19 pandemic, we further exclude post-2019 observations (though results are robust to including 2020 and 2021).

2.2 Empirical Strategy

We estimate gravity-style models of migration, building on the trade literature (Anderson and Van Wincoop, 2003). In our framework, origin-destination flows are a function of time-varying origin and destination factors and time-invariant bilateral distances. The central innovation is to include climate distance—measured as the absolute difference in average temperature and precipitation—alongside physical distance.

Because many origin-destination pairs record zero flows, we estimate all regressions using Pois-

[7] A small number of county–year observations (58) display immigrant population shares above one. These cases occur only in the historical 1880–1920 period and almost exclusively in counties with very small populations, likely reflecting enumeration errors or inconsistencies in the boundary–adjustment procedure. We set these observations to missing, although results are virtually unchanged if we retain them—either using the raw values or after winsorizing.

[8] Internal migration patterns changed in the 1930s, due to the Great Depression (Rosenbloom, 2002; Fishback et al., 2006) and environmental shocks like the Dust Bowl (Hornbeck, 2012, 2023).

[9] Since 1890 is not available, we measure migration from 1880 to 1900. Our baseline results include all individuals aged 15 and older in the pre-migration census year in the linked sample, but the estimates are robust to restricting the sample to men. As discussed below, results are also robust to using linked samples from Abramitzky et al. (2020) and to relying on the 1940 full count Census to measure migration between 1935 and 1940.

[10] See IRS Statistics of Income (SOI) Division Tax Stats—Migration Data, `https://www.irs.gov/statistics/soi-tax-stats-migration-data`. The dataset is available for filing years 1991 through 2022, but we focus on migration flows measured after 2011 when the SOI Division assumed responsibility for producing the tabulations and implemented several changes to the underlying construction of the data. These include (i) incorporating all tax returns filed and processed within the full calendar year (rather than only those processed by early fall), (ii) improving year-to-year matching by using the TINs of primary, secondary, and dependent filers, and (iii) revising tabulations and expanding available breakdowns.

son Pseudo-Maximum Likelihood (PPML). Our baseline specification is:

$$M_{odt} = \exp[\alpha_{ot} + \gamma_{dt} + \beta \text{Dist}_{od}^{\text{Climate}} + \theta \text{Dist}_{od}^{\text{Physical}}]\epsilon_{odt} \tag{1}$$

where M_{odt} is the number of migrants from origin o to destination d in period t, and $\text{Dist}_{od}^{Climate} = |\text{Climate}_o - \text{Climate}_d|$ is a vector denoting the absolute difference in climate (temperature and precipitation) between origin o and destination d. We also control for the physical distance between o and d ($\text{Dist}_{od}^{Physical}$) as well as for origin by period fixed effects (α_{ot}) and destination by period fixed effects (γ_{dt}).[11]

We estimate equation (1) for both international and internal migration in the U.S. In the international context, o denotes the migrant's country of origin, while in the internal context o denotes the U.S. county of origin; in all cases, d denotes the U.S. county of destination. Migration flows are aggregated at the decade level, corresponding to census years, except in the case of modern internal migration, where we use annual data. Standard errors are clustered at the country-of-origin by destination-state level for international migration, and at the origin-state by destination-state level for internal migration.

A key identification concern is the spatial correlation of climate.[12] In the international context, transoceanic migration largely breaks the link between climate and geography. Table A.2 regresses temperature distance between country–county pairs on linear, quadratic, and cubic functions of physical distance. The R-squared indicates that geographic distance explains no more than 1.6% of the variation in temperature distance. Columns 4–6 repeat this exercise for county pairs in the contiguous United States. In this case, the R-squared rises to at most 0.199, implying that geographic distance accounts for up to 20% of the variation in temperature distance. While this correlation is stronger than in the international setting, it remains modest, leaving substantial independent variation in climate distance to be exploited.

Figures A.1 and A.2 provide a complementary visual perspective. Figure A.1 shows that, while average temperatures decline with latitude in the United States, geographic features such as mountains and water bodies generate considerable within-country climate heterogeneity. Figure A.2 further illustrates this point by plotting the distribution of climate distance for two example U.S. locations, highlighting meaningful variation in climatic similarity across geographically proximate areas. We systematically address remaining concerns related to spatial correlation in climate in Section 2.5 and in Appendix C.

[11] In all specifications, period refers to the census decade, except for modern internal migration, where it refers to the calendar year.

[12] Climate is no exception to the first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970).

2.3 Main Results

Baseline estimates. We present the results from equation (1) in Table 1, separately for international (columns 1–2) and internal (columns 3–4) migration.[13] In all specifications, the coefficient on temperature distance is negative and statistically significant, confirming the pattern shown in Figure 1: migration flows are higher between origins and destinations with more similar temperatures. The coefficient on geographic distance is also negative and statistically significant across all models, consistent with well-established frictions to migration over space (Schwartz, 1973; Bryan and Morten, 2019). The coefficient on precipitation distance is negative and, except for column 1, statistically significant, suggesting some degree of matching along this dimension of climate. However, it is smaller in size relative to temperature distance; moreover, when probing its robustness across settings it becomes less stable.

One interpretation for the stronger temperature—rather than precipitation—results is that temperature acts as a sufficient statistic for the broader climatic environment migrants expect to face: it is salient, easy to observe, and strongly correlated with many features relevant for productivity, comfort, and cultural familiarity. The weaker and less stable relationship between precipitation distance and migration also aligns with findings in the environmental economics literature, which generally identifies temperature as a more consistent predictor of human and economic outcomes. Temperature affects agricultural productivity, labor supply, and health in direct ways that are harder to mitigate, whereas precipitation effects tend to be more context-specific and mediated by factors such as irrigation, water storage, and soils (e.g., Schlenker et al., 2005; Carleton and Hsiang, 2016; Taylor, 2022; Proctor et al., 2022). Our subsequent analysis thus focuses on temperature distance, while continuing to control for precipitation distance in all specifications.

The implied magnitudes for historical international migration (column 1) as well as for historical and modern internal migration (columns 3 and 4) are similar: lowering temperature distance by 1°C increases migration by 16.1%, 27.0%, and 21.1%, respectively.[14] Results for international modern migration (column 2) are instead an order of magnitude smaller in absolute value, and somewhat less precisely estimated—a pattern to which we return in Section 2.4 below. In this case, reducing temperature distance by 1°C increases migration by 1.9%.

To benchmark these effects, we compare a 5°C reduction in temperature distance—equivalent to the difference in the mean annual temperature of Chicago (9.8°C) and Washington, DC (14.6°C), or London (10.7°C) and Rome (15.5°C)—to geographic distance coefficients. The implied effect is equivalent to reducing geographic distance by approximately 300 to 750 kilometers, depending

[13] Section 2.5 summarizes all robustness checks, which are presented in detail in Appendix C.

[14] The implied magnitude can be calculated as $100 \times (e^{\beta} - 1)$, where β is the estimated coefficient on temperature distance from the PPML regression.

on the specification. The magnitudes are also comparable to standard elasticities in the spatial economics literature. For cross-country European migration, Caliendo et al. (2021) estimate a wage elasticity of roughly 0.5. In internal-migration contexts, estimated wage elasticities range from 1.5 in China (Tombe and Zhu, 2019) to 4.5 in Brazil (Morten and Oliveira, 2024). A 5°C improvement in climate similarity thus generates effects similar to moving several percentage points along typical wage gradients.

Finally, it is informative to compare these magnitudes with those attributed to the "pull" of existing migrant networks (see also Munshi, 2020, for a review). For U.S. internal migration in the early 20th century, Stuart and Taylor (2021) estimate that one additional migrant induces 1.9 (Black) or 0.4 (white) additional movers from the same birthplace. Green (2025) provides complementary evidence, documenting that a one standard deviation increase in exposure to shipmates from other states during World War II increases out-of-state migration over the subsequent five years by 4–5%.[15] For international flows, Beine et al. (2011) find that a 10% increase in the existing diaspora raises new immigration by 1–6%. Our estimated climate-distance effects fall within this range. In the larger historical and internal cases, they approach the upper end of canonical network spillovers.

Exploiting long-run changes in climate. In Table 1, we exploited cross-sectional variation in bilateral climate distances computed from the average climatology of each location over the period 1960–2000. We now leverage variation that arises from *changes* in average climate occurring differentially across space. Specifically, we test whether, over the 20th century, the *change* in climate distance across space predicts the *change* in international and internal migration patterns.[16] We focus on two periods: a historical window in the early 20th century and a modern window in the early 21st century.[17]

Figure 3 visually illustrates how climate distance between each U.S. county and the four countries shown in Figure 2 has changed over time. The maps reveal substantial spatial variation in relative climate convergence and divergence. Faster warming in colder origin countries such as Norway and Germany reduces temperature distance to warmer U.S. regions, generating widespread decreases in climate distance (green shading) across the U.S. Southeast. For Mexico, by contrast, it is the colder northern regions that have become more similar, while the Southeast has grown more dissimilar. Figure A.6 presents analogous maps for two selected U.S. counties—one in the Delta (Panel A) and one in the Northeast (Panel B)—relative to all other

[15] Using Facebook friendship data for the U.S. between 2012 and 2023, Koenen and Johnston (2024) find that one additional social tie increases an individual's probability of migrating to a given location by 0.3 percentage points.

[16] While anthropogenic emissions have driven the overall warming of the climate since about 1980, such spatiotemporal variation is a normal part of the climate system, influenced by multi-decadal oceanic circulation patterns (e.g., North Atlantic Oscillation), solar variability, volcanic activity, and land use change.

[17] The climate over the historical and modern periods are measured over 1901–1930 and 1991–2020, respectively.

U.S. counties.

We replicate the analysis in Table 1 allowing climate distances to vary between the early and the modern periods. In practice, we augment the gravity model in equation (1) by adding origin-by-destination fixed effects to absorb any pair-specific (time-invariant) variation. To allow the effects of the *change* in climate distance on migration to vary depending on geographic distance, we further control for the interaction between geographic distance and period fixed effects. Formally, we estimate:

$$M_{od\tau} = \exp\left[\alpha_{o\tau} + \gamma_{d\tau} + \gamma_{od} + \beta \text{Dist}_{od\tau}^{Climate} + \theta_\tau \text{Dist}_{od}^{Physical}\right] \varepsilon_{od\tau} \quad (2)$$

where everything is as before, but τ refers to the period (historical or modern), γ_{od} are origin-by-destination fixed effects, and $\theta_\tau \text{Dist}_{od}^{Physical}$ are period dummies interacted with geographic distance.

We present results in Table 2. In columns 1 and 2, we examine international migration, considering different time windows to measure migration flows (1900 and 2010; average over 1900–1920 and over 1990–2010, respectively). In columns 3 and 4, we focus on internal migration, using the 1900–1910 linked sample (resp., the 1900 to 1940 averages from the linked sample) and the 2018-2019 IRS migration data (resp., the average over the 2011–2019 period). The coefficient on temperature distance is statistically significant at the 1% level in all specifications. Magnitudes are broadly consistent with those in Table 1—if anything somewhat larger for internal migration. A 1°C increase in temperature distance over time—comparable to the average global temperature increase over the last century—reduces migration by 26.4%–35.4%.

These patterns provide compelling evidence that climate matching is not driven by time-invariant origin-destination characteristics—such as shared language, cultural proximity, economic linkages, or historical migration patterns. By including origin-by-destination fixed effects, the specification absorbs all stable bilateral factors, allowing us to isolate the role of changes in climate distance in driving changes in migration flows. The model in equation (2) also addresses the possibility, discussed in the next section, that the observed effects are simply picking up the persistence of ethnic networks. Since networks tend to be highly persistent over time, any influence they exert that does not vary across periods is captured by the fixed effects.

2.4 Mechanisms

What explains climate matching in migration? We consider three complementary mechanisms: the role of ethnic networks, climate-specific human capital, and climate-as-amenity.

Past networks. Early pioneers may have selected destinations with climates similar to their origins, but subsequent migrants may have simply followed established co-ethnic communities—regardless of climate. The long-difference analysis in Table 2 suggests that this mechanism alone cannot explain our findings. That specification includes origin-by-destination fixed effects, which absorb all time-invariant bilateral characteristics—including persistent ethnic networks. For past networks to still confound the results, their strength would need to evolve over time in tandem with random spatiotemporal changes in climate. While this is possible in theory, it seems unlikely in practice.

We complement this evidence with two additional exercises designed to assess the extent to which network dynamics could account for the observed climate-matching pattern. First, we explicitly control for lagged migration. In Figure 4, we replicate our baseline estimates for both international (Panel A) and internal (Panel B) migration, decade by decade, and include the number of migrants from the same origin-destination pair in the previous period as a control.[18] Across all decades, the results are nearly identical: the gray triangles (with lagged migration controls) are indistinguishable from the black dots (baseline estimates). This pattern suggests that the climate distance effect is not simply a proxy for persistent enclaves.[19] However, this interpretation should be treated with some caution, given that past migration itself is a function of climate similarity.

Second, we consider a historical setting where past networks should have limited scope to influence migration: the U.S. frontier. We restrict destinations to the set of counties that between 1850 and 1890 were on the American frontier, defined by Bazzi et al. (2020) as a set of sparsely populated counties near the moving edge of settlement by whites (who account for the vast majority of the migrants in our sample).[20] Table A.4 reports the results for both international (columns 1–3) and internal (columns 4–8) migration.

For international migration, we focus on the final two frontier decades, 1880 and 1900. Column 1 pools these decades, while columns 2 and 3 consider them separately. Despite the substantial

[18] To facilitate comparison across settings, temperature distance is standardized to have mean zero and unit variance within each sample. For international migration, we use the 1940 stock for 1970 given data availability and the very low migration flows between 1940 and 1960 (Abramitzky and Boustan, 2017). For internal migration, lagged stocks cannot be constructed at the county level because only state of birth is observed; instead, we control for cumulative migration flows between each origin–destination pair using the linked sample. For modern internal migration, we report estimates based on 2018–2019 flows (controlling for 1910–1940 flows from the linked sample); results are similar when averaging over the full 2011–2019 period. Lagged migration controls cannot be constructed for the first observation in each sample (1880 for international migration and 1860 for internal migration) due to data limitations.

[19] Table A.3 replicates the baseline specification (Table 1) while additionally controlling for lagged migrant stocks. Because these stocks are used as regressors, the first observation in each sample—1880 for international migration and 1850–1860 for internal migration—is omitted from the left-hand side in columns 1 and 3. Consistent with the existing literature (Card, 2001; Munshi and Rosenzweig, 2016), prior migrant networks are positively associated with migration flows. The results are virtually unchanged when replacing lagged stocks with stocks measured in the first sample year (e.g., 1880 for international migration and 1850–1860 for internal migration).

[20] Since we cannot compute migration for the 1880–1890 decade, we consider migration between 1880–1900, and use frontier status in 1890 (the last year considered in Bazzi et al., 2020).

reduction in sample size, the estimated effect of temperature distance remains negative, statistically significant, and similar in magnitude to our baseline results. For internal migration, we can exploit a longer window of frontier expansion, 1850–1900. Column 4 pools all decades, while columns 5–8 estimate the relationship decade by decade. Across all specifications, the coefficients remain negative and statistically significant, even though the magnitudes are smaller in absolute value than in the full-sample estimates. One interpretation is that network-related channels partly account for the larger temperature-distance effects documented in the full sample. Another possibility is that internal migration to the frontier reflects a different composition of origin counties—specifically, origins that were never on the frontier—so that selection dynamics differ from those in the full sample. We cannot distinguish between these explanations, but the main take-away is that the negative relationship between temperature distance and migration remains strong even in this setting.

Climate-specific skills vs. Climate-as-amenity. We then turn to two complementary explanations: climate-specific skills and climate-as-amenity. Figure 4 discussed above documented that the elasticity of migration with respect to climate distance declines over time, particularly for international migration. In other words, while climate matching remains a relevant factor, its role in shaping location choices becomes less pronounced. This temporal pattern is consistent with the idea that migrants partly value climate similarity because of climate-specific human capital. Between 1850 and 1940, the share of U.S. employment in agriculture fell from 55% to 17% (Lebergott, 1966). Since climate-specific skills—such as knowledge of crop suitability, seasonal cycles, and pest management—are especially important for agricultural workers, this decline suggests a diminishing role for climate-based productivity as the economy transitioned toward industry and services.

We provide further evidence for the role of climate-specific skills in Figure 5, which shows that the negative effect of temperature distance is larger for individuals employed in agriculture—a pattern that holds for both international (Panel A) and internal (Panel B) migrants. Figure A.7 corroborates this idea by classifying occupations according to whether they are performed primarily indoors or outdoors and whether they are climate intensive.[21] Consistent with the climate-specific human capital mechanism, climate distance has stronger effects for workers in outdoor and climate-intensive occupations.[22]

[21] Industries (resp., occupations) are defined using the Census *IND1950* (resp., *OCC1950*) codes. Following Eckert and Peters (2025), we classify agriculture as codes 105 (agriculture), 116 (forestry), and 126 (fishery). For internal migrants, industry and occupation are measured at baseline; for international migrants, this is not possible, since industry and occupation are observed only in the destination census year. To facilitate comparison, temperature distance is standardized to mean zero and unit variance within each relevant subsample.

[22] A complementary explanation may be that farmers are better at observing and inferring climatic similarity across locations. Then, even if climatic sorting were driven primarily by climate-as-amenity rather than by climate-specific skills, we would still expect stronger sorting among farmers. Our data do not allow us to disentangle these two interpretations, and both may operate simultaneously.

However, climate distance also predicts migration flows for individuals working in sectors where climate-specific skills should be less relevant—such as services and other indoor occupations. Moreover, even in recent decades, climate similarity continues to influence location choices, particularly for internal migrants (Figure 4, Panel B). These findings suggest that migrants also value climate for its amenity value. For instance, if climate conditions influence cultural practices—from cooking traditions and recreational activities to religious observances—then climate similarity may contribute to a sense of familiarity and comfort in destination regions.[23]

The role of air conditioning. In principle, technologies can insulate migrants from some of the discomfort associated with climate mismatch. While this protection is incomplete—many activities still take place outdoors—such innovations should weaken the relationship between migration and climate similarity. Table A.5 explores this idea by examining how the spread of air conditioning (AC) affected migrants' sensitivity to temperature. The table replicates our international migration analysis separately for the historical period (1880–1920, columns 1–3) and the modern period (1970–2010, columns 4–6). We distinguish between January and July temperature distances and interact each with a dummy equal to one if the destination county's average July temperature lies above the national median ("South"). This allows us to test whether migrants' avoidance of hot destinations weakened after the 1960s when AC was rapidly adopted in the U.S. (Biddle, 2008).[24]

Consistent with our previous results, the effect of temperature distance is negative in all specifications, though smaller in magnitude in the modern era, indicating that the overall importance of climate matching has declined over time. Columns 2 and 5 separate winter and summer temperatures. In the historical period, January temperature distance has a stronger effect, likely reflecting the fact that many migrants at the time came from colder European countries. In the modern period, the July coefficient becomes larger and statistically significant, a pattern consistent with the changing composition of migrants, who increasingly originate from warmer regions.

Columns 3 and 6 introduce interactions between temperature distance and the "South" dummy. In the historical period (column 3), the coefficient on the interaction between the South dummy and July temperature distance is negative and statistically significant, implying that migrants historically avoided hotter destinations that were less similar to their climates of origin. In the modern period (column 6), this interaction becomes positive and statistically significant, showing that migrants are no longer deterred by heat to the same extent. The weakening of this

[23] This interpretation resonates with findings in Albouy et al. (2021), who document a relationship between the climatic and geographic features of immigrants' countries of origin and the U.S. cities where they settle.

[24] Ideally, we would replicate this analysis for internal migration as well. However, comparable data on county-to-county migration flows are not available systematically for the 1970–2000 period, which is critical for assessing the impact of air conditioning on migration.

summer-temperature effect in hot regions is precisely what we would expect if AC reduced the discomfort associated with climatic mismatch. While July temperature distance continues to exert an overall negative influence on migration, its magnitude is much smaller in the post-1970 era, indicating that technological adaptation has mitigated the amenity cost of heat.

If this mechanism operates through amenities rather than productivity, its effects should be concentrated among non-farmers, whose thermal exposure is largely mediated by indoor environments. Columns 7 and 8 test this implication by estimating the specification in column 6 separately for farmers and non-farmers. Consistent with the amenity-based interpretation, the interaction between July temperature distance and the South dummy is positive and statistically significant for non-farmers, but negative and imprecisely estimated for farmers. Given the limited relevance of AC for agricultural work, these patterns indicate that the moderating effect of AC operates mainly through amenities rather than through climate-sensitive production.

Taken together, these results reinforce the view that part of the climate-matching effect operates through the amenity value of climate, and that this channel is sensitive to technologies—such as air conditioning—that alter how individuals experience the weather. We emphasize that the evidence presented here is suggestive and does not allow us to separately identify the climate-specific skills and amenity channels. Nevertheless, our results point to the relevance of both mechanisms—and possibly to their complementarity. For example, food preferences are shaped by culinary practices in the origin country, which in turn depend on what can be grown, preserved, and prepared in the local climate. In such cases, productivity-based and amenity-based motives for climate matching may reinforce each other.

2.5 Summary of Robustness Checks

The climate matching patterns documented in Section 2.3 are robust to a wide range of specifications and checks. We present these in Appendix C and summarize them here. In Appendix C.1, we first assess the sensitivity of the results to the sequential inclusion of additional controls and fixed effects (Tables C.1–C.4), excluding continents sending relatively few migrants to the U.S. or decades and measuring climate with capital-city temperatures (Table C.5), and clustering standard errors at larger levels (Table C.6).

In Appendix C.2, we then address the concern that climate distance may proxy for geographic distance, particularly in the context of internal migration. We replicate the internal migration analysis under a series of increasingly demanding spatial specifications: (i) estimating regressions separately by direction of movement (Table C.7); (ii) excluding short-range and within-state moves, as well as potentially idiosyncratic destinations such as California and Florida

(Table C.8); and (iii) allowing for nonlinear effects of geographic distance (Table C.9). Across these exercises, the historical internal migration results remain highly robust. For modern internal migration, the estimated effects are generally smaller but continue to hold, except when focusing on vertical movements or when allowing for flexible polynomial controls for geographic distance. The stronger and more stable patterns for historical migration are consistent with our earlier evidence that climate matching has weakened over time.

In Appendix C.3, we further show that results are robust to controlling for origin–destination differences in a wide range of geographic characteristics—including elevation, coastal access, ruggedness, latitude and longitude, and soil quality (Tables C.10–C.17). Focusing on historical internal migration, we also verify that the findings are unchanged when we control for a rich set of bilateral economic and demographic characteristics (Figure C.1) and when we use time-varying travel-cost proxies (Table C.18). In Appendix C.4, we assess robustness to alternative climate definitions, such as measures of seasonal variability (Tables C.19–C.20) and alternative temperature statistics, including minimum and maximum temperatures and growing-season indices (Tables C.21–C.22). In Appendix C.5, we show that the historical internal migration results are robust to alternative sample restrictions and to using the linked samples from Abramitzky et al. (2020) as well as the 1940 full count U.S. Census (Table C.23).

3 The Economic Consequences of Climate Mismatch

The evidence thus far indicates that migrants systematically sort into places with climates similar to their homelands. We now examine what happens when this matching margin fails—when migrants end up in places whose climates differ substantially from those they know. We focus on international immigration to the U.S. between 1880 and 1920. This setting offers two advantages. First, the staggered expansion of the railroad network—combined with variation in the timing of national-level inflows from different origin countries—generates plausibly exogenous variation in climate mismatch. Second, agriculture was a major sector, climate-specific skills were central to production, and local economies were closely tied to the land. At the same time, the U.S. was undergoing a rapid structural transformation from agriculture to manufacturing (Eckert and Peters, 2025), making this a period in which shifts in sectoral productivity and labor allocation could have consequential effects for local development.

The economic effects of climate mismatch are *ex ante* ambiguous. On the one hand, when migrants' climate-specific skills do not transfer well to local conditions, agricultural productivity should fall, lowering the value of output and reducing incentives to cultivate marginal land. Because agriculture and non-agriculture were tightly linked at the time, these productivity losses could spill over into other sectors, slowing output growth and the pace of structural transformation. Moreover, sectors such as construction and manufacturing were themselves partly climate dependent, so mismatch could reduce productivity there directly. On the other hand, climate mismatch could promote structural change if counties with poorly matched migrants shifted more rapidly into manufacturing, while better-matched areas remained concentrated in agriculture. Appendix E develops a simple two-sector model that formalizes these mechanisms and provides a framework for interpreting the empirical results that follow.

3.1 Empirical Strategy

Baseline estimating equation. To study the effects of climate mismatch on economic development, we consider a panel of U.S. counties observed between 1880 and 1920.[25] We restrict attention to counties with positive population in 1880, and estimate regressions of the form:

$$y_{dt} = \alpha_d + \alpha_{rt} + \beta \text{Mismatch}_{dt} + X'_{dt}\gamma + \varepsilon_{dt} \tag{3}$$

where y_{dt} denotes an economic outcome for county d in decade t; α_d and α_{rt} are county fixed effects and census region-by-decade fixed effects; and, X_{dt} is a vector of control variables introduced below. We cluster standard errors at the county level.

The index of climate mismatch, Mismatch_{dt}, is defined as the weighted average of the temperature distance between each immigrant's country of origin and the destination county:

$$\text{Mismatch}_{dt} = \sum_o \omega_{odt} \times \text{TempDist}_{od}. \tag{4}$$

TempDist_{od} is the absolute difference in average annual temperature between origin country o and destination county d, expressed in degree Celsius, and the weights are given by

$$\omega_{odt} = \frac{M_{odt}}{M_{dt}} \tag{5}$$

with M_{odt} denoting the number of immigrants from origin o to county d arriving between decades $t-1$ and t, and $M_{dt} = \sum_o M_{odt}$ denoting total immigration into county d during the same period.

Figure A.8 shows the spatial distribution of climate mismatch in 1880 (Panel A) and 1920 (Panel B). The 1880 map already displays substantial geographic variation, with higher mismatch concentrated in the South (notably the Gulf Coast and Florida) and scattered pockets in the

[25] Because climate mismatch cannot be constructed for 1890, we exclude this decade from the analysis.

West, while mismatch remains relatively low in much of the Northeast and Upper Midwest. By 1920, mismatch intensifies across the Southeast, Mid-Atlantic, and Rocky Mountain West, while remaining comparatively low in much of the Northeast and Upper Midwest.

Instrument for climate mismatch. The inclusion of county fixed effects in equation (3) absorbs all time-invariant factors correlated with both climate mismatch and economic development, while census region-by-decade fixed effects capture shocks common to counties within the same region and decade. Identification therefore comes from within-county variation over time—that is, from comparing changes in economic outcomes across counties within the same census region and decade. Even with this set of fixed effects, changes in climate mismatch could correlate with other time-varying county characteristics that also influence economic growth. For example, in faster-growing counties climate may be less relevant to migrants' location decisions, and such counties may attract migrants from a wide range of origins regardless of climate fit, generating a spurious relationship between mismatch and economic development. To address such concerns, we construct an instrument for climate mismatch that exploits plausibly exogenous push and pull forces shaping the distribution of immigrants across counties over time.

Our strategy builds on Sequeira et al. (2020) and exploits the roll-out of the U.S. railroad network as a key determinant of immigrant settlements: counties received more immigrants when they become newly connected to the railroad at a time of high immigrant inflows. Following Terry et al. (2024), we extend this framework to account for the staggered arrival of different national groups. Specifically, if a county becomes connected to the railroad during a period of intense Norwegian immigration, it will attract relatively more Norwegians; if it is connected when German inflows peaked, it will receive more Germans, and so on. The interaction between the timing of railroad access and origin-specific immigration waves generates plausibly exogenous variation in the composition of immigrant inflows—and therefore in the degree of climate mismatch across U.S. counties and decades.

We restrict attention to European immigrants—who accounted for more than 85% of total arrivals during this period—because consistent decadal data on aggregate inflows by country of origin are available only for European countries (see Section 3.2 for details). We augment the baseline PPML model from equation (1) to estimate:

$$M_{odt} = \exp\left[\alpha_{ot} + \gamma_{dt} + \beta_1 \text{Dist}_{od}^{Climate} + \beta_2 \text{Dist}_{od}^{Physical} + \beta_3 \log(M_{ot}) \times R_{dt}\right] \varepsilon_{odt} \tag{6}$$

As in equation (1), M_{odt} denotes the number of immigrants from origin o who arrived in the previous decade and were living in county d in decade t, $\text{Dist}_{od}^{Climate}$ and $\text{Dist}_{od}^{Physical}$ refer to physical and climate distances, and α_{ot} and γ_{dt} are country-of-origin by decade and county-of-destination by decade fixed effects.[26] The key innovation of this specification is the interaction term $\log(M_{ot}) \times R_{dt}$. This captures two sources of variation: the (log) number of migrants from origin o arriving between decades $t-1$ and t, which proxies for origin-specific push shocks, and an indicator for whether county d was first connected to the railroad between decades $t-1$ and t (i.e., the arrival window used to construct M_{odt}), which proxies for local pull shocks. The interaction $\log(M_{ot}) \times R_{dt}$ therefore measures how much a U.S. county–origin pair responds when a country experiences a large emigration wave *and* the U.S. county becomes newly connected to the railroad.

Column 1 of Table A.6 reports the estimates from equation (6). The coefficient on temperature distance is negative and statistically significant, consistent with our baseline results in Table 1.[27] The interaction between $\log(M_{ot})$ and R_{dt} is positive and statistically significant: immigration from origin o into county d is higher in decades when that origin is sending more migrants to the United States and the county gains railroad access. In column 2, we perform a placebo exercise by replicating the specification in column 1 but interacting the 10-year leads of both the railroad indicator and the (log) country-specific migration push shock. The estimated coefficient is close to zero and statistically insignificant.

We then use the fitted values from equation (6) to construct a predicted index of climate mismatch. Let $\widehat{M}_{odt}$ denote the predicted number of immigrants from origin o to county d between decades $t-1$ and t, and define predicted shares $\widehat{\omega}_{odt} = \widehat{M}_{odt} / \sum_{o'} \widehat{M}_{o'dt}$. Substituting these shares into equation (4) yields the predicted mismatch index, $\widehat{\text{Mismatch}}_{dt}$. Because this is a generated regressor, we compute standard errors in the second stage using a block bootstrap: we resample counties with replacement, re-estimate equation (6) within each draw, reconstruct the predicted mismatch measure, and re-estimate equation (3) on each bootstrap sample. The reported standard errors are based on the distribution of the resulting estimates of the treatment effect.

Panels C and D of Figure A.8 display the predicted mismatch index for 1880 and 1920. Although the broad patterns resemble those in the actual mismatch maps (Panels A and B), the predicted surface is noticeably smoother, reflecting the fact that the instrument filters out much of the idiosyncratic variation present in the data.

Instrument validity. Predicted mismatch depends on the composition of predicted migration inflows—not their level. When we normalize predicted flows to obtain shares, any term in equation (6) that is constant across origins for a given county-decade (i.e. γ_{dt}) cancels out of

[26] As in Section 2, we include all foreign-born individuals in the 1880 sample, cluster standard errors at the country-of-origin by U.S. state-of-destination level.

[27] The coefficient on precipitation distance is similar to that in column 1 of Table 1, while the coefficient on geographic distance becomes small and imprecisely estimated. The number of observations is lower than in the baseline specification (Table 1, column 1) because data on aggregate migration flows are missing for some countries.

$\widehat{\omega}_{odt}$ by construction. Variation in $\widehat{\text{Mismatch}}_{dt}$ therefore comes from three sources: (i) bilateral geography, $\text{Dist}_{od}^{Climate}$ and $\text{Dist}_{od}^{Physical}$; (ii) origin–decade fixed effects, α_{ot}, capturing push forces driven by conditions in sending countries; and, (iii) the interaction $\log(M_{ot}) \times R_{dt}$, capturing how national origin inflows are differentially transmitted across U.S. counties depending on when they are connected to the railroad network. Since the bilateral distance terms (physical and climate) are predetermined and time-invariant, any induced level differences in local outcomes are absorbed by the county fixed effects in equation (3). Conditional on these fixed effects and on region-by-decade shocks, the timing of M_{ot} and R_{dt} is plausibly orthogonal to county-specific shocks to economic development, so the instrument isolates variation in mismatch arising from exogenous push–pull forces rather than from endogenous shifts in migrant location choice.

A remaining concern is that railroad expansion could affect local economic development directly, violating the exclusion restriction. To address this, all regressions control for the main effect of railroad connectivity. The identifying assumption could also be violated if railroad expansion in certain counties (within a census region) during a given decade systematically attracted immigrant groups that were a particularly good climate match for those counties. To address this possibility, we replicate the analysis constructing the instrument using immigrant flows predicted from exogenous weather shocks in origin countries following Sequeira et al. (2020) and Medici (2024). We present this and further robustness checks in Section 3.4, including long-difference regressions that alleviate concerns about bias in two-way fixed-effects settings with heterogeneous treatment (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021).

3.2 Data and Measurement

We complement the migration and climate data described in Section 2.1 with additional sources. First, to construct predicted climate mismatch, we use information on the year each county was first connected to the railroad network from Atack (2016). These data provide a decade-by-decade record of railroad expansion across U.S. counties. We combine them with newly digitized annual records on the number of immigrants from each European country entering the United States between 1860 and 1920, compiled from Willcox (1929). Together, these datasets allow us to construct the push and pull shocks underlying our instrument.

Next, to track population and employment, we use county-level data from the full count decennial U.S. Census of Population between 1880 and 1920 (Ruggles et al., 2021). From these, we obtain total population and the number of men of working age (15–64) by employment sector and nativity. To measure economic outcomes in agriculture and manufacturing, we draw on the Census of Agriculture (Haines and Inter-university Consortium for Political and Social

Research, 2000) and the Census of Manufactures (Haines and Inter-university Consortium for Political and Social Research, 2010; Haines et al., 2018). These sources provide information on agricultural output, land use, capital, and farm value, as well as manufacturing output, establishment counts, and wages.[28] Finally, we draw on the 1940 full count Census, which for the first time reports income systematically, to measure earnings and construct a proxy for hourly wages.[29]

3.3 Results

The agricultural sector. We begin by examining how climate mismatch affected the agricultural sector. Table 3 reports OLS and 2SLS estimates (Panels A and B), as well as the first stage (Panel C). Across all outcomes, OLS and IV coefficients are similar in sign but differ substantially in magnitude; we discuss this pattern in detail at the end of the section.[30] Column 1 shows that climate mismatch reduced the value of crops produced. The negative and statistically significant 2SLS coefficient implies that a 1°C increase in mismatch lowers the value of county-level agricultural production by 16.3%, or $83,474 lower crop value relative to the 1880 mean. For reference, the difference in mismatch between the 25th and 75th percentiles of the distribution is about 4°C—comparable to the gap between Boston and Raleigh, Chicago and Richmond, or Denver and Dallas. This implies that crop values were roughly 50.9% lower in a high (75th percentile) than in a low (25th percentile) mismatch county. Column 2 shows that mismatch also reduced the share of county land devoted to agriculture: increasing mismatch from the 25th to the 75th percentile reduces the farmland share by 10.8 percentage points (16.9% of the mean).

Columns 3 to 5 of Table 3 examine the intensive margin of agricultural performance conditional on land use, using crop revenue per acre, farm value per acre, and equipment value per acre as dependent variables.[31] The coefficients are negative and statistically significant. This indicates that mismatch weakened agricultural performance not only by shrinking the sector but also by reducing productivity on cultivated acres. The magnitudes are large: increasing mismatch from the 25th to the 75th percentile (4°C) reduces per acre crop, farm, and equipment values by 33.0%, 44.5%, and 34.6%, respectively. Our estimates are in line with those by Hornbeck

[28] We use data from the Census of Agriculture for each decade from 1880 to 1920 and from the Census of Manufactures for decades 1880–1920, except for 1910, which is unavailable. For the Census of Agriculture, we restrict the sample to counties with at least 100 farms in each year to limit concerns that counties with very few farms enter or exit the sample in ways that could influence the results.

[29] The 1940 Census reports only wage and salary income and provides no continuous measure of farm profits. As a result, total income is understated for farm households.

[30] Panel C indicates a strong first stage, as reflected in the large Kleibergen–Paap F-statistic reported in Panel B.

[31] Crop revenue per acre provides a direct measure of land productivity, while farm value per acre and equipment value per acre reflect the capitalized value of agricultural profits and capital intensity, respectively.

(2012), who finds that high-erosion counties during the Dust Bowl experienced a 30% decline in farmland value per acre. These intensive-margin estimates align with evidence on the limits of agricultural adaptation to climate change. We find that a 1°C increase in climate mismatch reduces crop revenue per acre by roughly 9.5%, a magnitude comparable to the estimated revenue impact of 1°C of long-term warming (Burke and Emerick, 2016).[32]

Columns 6 and 7 show that mismatch reduced the number of farms and increased farm size, although the latter effect is small and imprecisely estimated. This is consistent with lower profitability and farmer exit, as in the Dust Bowl (Hornbeck, 2012).

In Table A.7, we turn to crop-specific outcomes. Corn, the dominant staple crop in many regions at the time (Olmstead and Rhode, 2008), exhibits clear negative effects: climate mismatch reduces both acreage (column 1) and output (column 2). Since the decline in acreage is larger than the decline in output, there is a mechanical increase in yield, which is defined as crop production per acre (column 3). However, this reflects a contraction along the extensive margin rather than a genuine productivity gain. Wheat behaves differently. Columns 4–6 show that mismatch increases wheat acreage and output by similar proportions, leaving wheat yields essentially unchanged. A possible interpretation is that mismatch induced farmers to substitute away from corn—a crop whose cultivation was particularly climate-sensitive and relied more heavily on accumulated local experience—toward wheat, which was generally more robust across environments and required less climate-specific knowledge for basic cultivation (Olmstead and Rhode, 2002; Raz, 2025).[33]

As noted above, OLS coefficients are consistent in sign with 2SLS estimates but are an order of magnitude smaller (in absolute value). Classical measurement error in mismatch attenuates OLS coefficients toward zero, but is unlikely to explain the magnitude of the difference. A more plausible explanation is that high-performing agricultural counties attracted a more heterogeneous—and more climatically mismatched—set of migrants, since climate-specific skills were less binding in these places. This selection pattern would generate a positive correlation between agricultural performance and mismatch, biasing OLS coefficients toward zero. A complementary explanation is that 2SLS identifies a local average treatment effect (LATE) for migrants whose location choices respond to the instrument. These compliers place less

[32] This comparison requires mapping changes in mean temperature to degree days (Schlenker and Roberts, 2009). A uniform 1 °C increase in mean temperature over the growing season generates approximately 180 additional degree days. For the average U.S. county, we assume that roughly one-third of these (about 60 degree days) fall above the harmful threshold, with the remaining two-thirds (about 120 degree days) being beneficial. Applying these shifts to the growing degree day coefficients reported in columns 2 and 4 of Table 3 in Burke and Emerick (2016) implies a predicted revenue decline of 10–22%.

[33] Corn is sensitive to photoperiodism (latitude-dependent day length), necessitating intensive selective breeding and local adaptation to ensure maturity within specific seasonal windows. While wheat has its own vulnerabilities such as stem rust, the dissemination of the Red Fife variety across much of the U.S. in the mid-to-late 19th century provided a hearty reliable grain for settlers. Wheat's broader environmental tolerance offered a more standardized risk profile than corn, particularly in contexts with limited local climate knowledge.

weight on climate similarity and co-ethnic networks, and have fewer climate-relevant skills as well as weaker access to local knowledge. Because they tend to settle in newly connected and less diversified counties—where adaptation failures are costlier—their outcomes yield larger IV effects.

Finally, although all regressions control for railroad connectivity, one may still worry that the instrument captures direct economic effects of railroad expansion, violating the exclusion restriction. Any such direct effects would bias the 2SLS estimates upwards, because railroads stimulated agricultural activity and market access (Donaldson and Hornbeck, 2016). Since our IV estimates are substantially more negative than OLS ones, the sign pattern goes in the opposite direction of what such a violation would generate. This makes it unlikely that direct railroad effects are driving our results. We examine additional threats to identification and present a wide range of robustness checks in Section 3.4.

The non-agricultural sector. The contraction of the agricultural sector documented in Table 3 raises the question of whether climate mismatch accelerated the transition into non-agricultural activities. Between 1880 and 1920, structural transformation in the U.S. was rapid and economic growth was driven largely by manufacturing (Eckert and Peters, 2025; Hornbeck and Rotemberg, 2024). If mismatch pushed workers out of agriculture and into other sectors, it could in principle have stimulated local industrial expansion. Table 4 examines this possibility.

In column 1, the dependent variable is the non-agricultural employment share. The positive and statistically significant coefficient indicates that mismatch favored the reallocation of labor out of agriculture and into the non-agricultural sector. The 2SLS estimates imply that the non-agricultural employment share rose by 3.7 percentage points per 1°C mismatch, or 14.8 percentage points more in a high-mismatch county than in a low-mismatch county (38% relative to the mean).[34] If mismatch pushes workers out of agriculture and into non-agricultural activities, diminishing returns imply that the marginal product of labor outside agriculture should fall. Although we lack comprehensive data for all non-agricultural sectors, we can observe manufacturing, the dominant non-agricultural sector at the time (Eckert and Peters, 2025; Hornbeck and Rotemberg, 2024). Column 2 shows that moving from the 25th to the 75th percentile of climate mismatch lowers manufacturing output per worker by roughly 30.0%. Column 3 reports a negative but small and imprecisely estimated coefficient for average wages, which may reflect downward wage rigidity, though we view this as suggestive.

The effect of mismatch on manufacturing output, however, is *ex ante* ambiguous. On the one hand, an influx of labor should increase total manufacturing output. On the other hand, climate

[34] This is in line with results from Colmer (2021), who estimates that increasing temperature by 1°C decreases agricultural employment share by 7 percentage points in contemporary India.

mismatch may depress manufacturing productivity—directly, because parts of manufacturing remained climate exposed, and indirectly, because weaker agricultural performance reduced demand for manufactured inputs and disrupted input–output linkages.[35] Column 4 shows a negative and sizable, though imprecisely estimated, effect of mismatch on manufacturing output. This suggests that productivity losses outweigh any scale effects, but the estimate is not precise enough for definitive conclusions. Finally, column 5 shows that mismatch reduces the number of manufacturing establishments—a pattern that mirrors agriculture and is consistent with lower profitability and the exit of marginal producers.

Taken together, these results indicate that climate mismatch did not stimulate non-agricultural growth. Although mismatch increased non-agricultural employment, it reduced productivity in the sector and did not raise manufacturing output. Combined with the negative effects on agriculture, mismatch acted as a drag on overall local economic activity. Appendix E develops a simple two-sector model that formalizes these mechanisms and shows how mismatch can simultaneously accelerate labor reallocation while lowering productivity in both sectors.

Population growth and income per capita. In Table 5, we examine the implications of climate mismatch for broader economic development. Column 1 documents that mismatch slowed population growth. The 2SLS coefficient implies that each 1°C in mismatch is associated with 10.2% lower population growth, whereby increasing mismatch from the 25th to the 75th percentile county lowers population growth by roughly 35%. This is a large effect—comparable to the difference between fast-growing industrial counties such as Cook County, IL, and slow-growing isolated counties such as Dakota County, NE.

We next turn to income per capita using the 1940 Census of Population when income was first systematically measured. Columns 2–7 of Table 5 report cross-sectional regressions relating income in 1940 to the average mismatch a county experienced between 1880 and 1920, controlling for census-region fixed effects and the year in which the railroad first reached the county.[36] While these estimates should be viewed as suggestive, they provide a useful complement to the panel results above. In column 2, the dependent variable is log average income. The negative and statistically significant coefficient indicates that counties with higher historical mismatch had lower income per capita in 1940. Comparing a high (75th percentile) to a low (25th percentile) mismatch county, the former is about 10.2% poorer. Column 3 uses (log) average hourly earnings—income divided by total annual hours. The 2SLS coefficient is again negative and statistically significant, and is nearly twice as large as the coefficient on total income, indicating that mismatch substantially reduced labor productivity.

[35] According to the 1900 Census of Manufactures, nearly half of U.S. manufacturing gross output was linked to agriculture through food processing (including grist mills and meatpacking), textiles (such as cotton and leather goods), wood products, tobacco, and the production of agricultural equipment and inputs.

[36] We measure income for men aged 25–64, but results are unchanged when using alternative age restrictions.

Columns 4–5 and 6–7 replicate columns 2–3, for U.S.-born and immigrant men separately. The patterns are similar across the two groups. Although coefficients are somewhat larger (in absolute value) for immigrants, they are negative, statistically significant, and large for U.S.-born men. This suggests that the adverse effect of mismatch was not confined to migrants who were directly exposed to unfamiliar climates, but spilled over to the U.S.-born population.

Taken together, these results corroborate the earlier evidence that climate mismatch hindered local development. Beyond shrinking agricultural activity and reducing farm productivity, mismatch reduced output per worker in manufacturing, slowed down population growth, and left a persistent imprint on county-level income. These effects were not limited to immigrants, but extended also to U.S.-born residents.

3.4 Robustness Checks

The results presented in Section 3.3 are robust to a wide set of inference and specification checks. All relevant tables can be found in Appendix D.

Panel regressions. As discussed in Section 3.1, one concern with the instrument is that migrants from colder (or warmer) origins may have arrived during periods when the railroad network was already expanding into similarly cold (or warm) counties. To address this possibility, we reconstruct the instrument using migration flows predicted purely from exogenous weather shocks in European origin countries following Sequeira et al. (2020). Appendix D.1 describes the implementation in detail. Results appear in Panel B of Table D.1 for agricultural outcomes and Table D.4 for manufacturing outcomes and population growth. We next address concerns regarding two-way fixed-effects (TWFE) estimators in the presence of heterogeneous treatment effects (De Chaisemartin and d'Haultfoeuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021). Panel C of Tables D.1 and D.4 reports long-difference regressions that compare 1880 to 1920 directly, thereby avoiding the TWFE structure. Panels D and E then replace region-by-decade fixed effects with state-by-decade fixed effects—absorbing all time-varying shocks at the state level—and trim the top and bottom 1% of the mismatch distribution to mitigate the influence of potential outliers.

Tables D.2 and D.5 then assess whether the instrument might be correlated with baseline or time-invariant characteristics that differentially shaped local development. First, we control for the 1880 immigrant population share. Second, when analyzing agriculture, we include crop-suitability measures based on climate and soil constraints (Ramankutty et al., 2002); when analyzing manufacturing or population, we instead control for the 1880 non-agricultural employment share to capture differential pre-trends in industrialization (Eckert and Peters, 2025). Third, we account for exposure to the frontier using the measure from Bazzi et al. (2020). Results for these checks are presented in Panels B–D, with Panel E including all controls jointly. Next, Tables D.3 and D.6 extend the baseline specification (reported in Panel A) by adding controls for predicted average precipitation distance (Panel B) and predicted geographic distance (Panel C), constructed analogously to temperature distance. Panel D includes all predicted distance measures simultaneously.

Taken together, the estimates remain qualitatively stable, supporting the validity of the empirical strategy. In some cases, coefficient magnitudes and precision vary across specifications, but these fluctuations do not display a systematic pattern across outcomes or checks.

Cross-sectional specification. Finally, we examine the robustness of the cross-sectional income results (Table 5, columns 2–7) in Tables D.7, D.8, and D.9. In all tables, Panel A presents the baseline IV estimates. In Table D.7, we use the weather-predicted migration instrument (Panel B), replace census-region fixed effects with state fixed effects (Panel C), trim counties with mismatch above the 99th or below the 1st percentile (Panel D), and exclude farmers to ensure that incomplete reporting of farm income does not drive the results (Panel E). In Table D.8, we add controls for (i) the 1880 immigrant share (Panel B), (ii) the 1880 non-agricultural employment share (Panel C), (iii) frontier exposure from Bazzi et al. (2020) (Panel D), and (iv) all three jointly (Panel E). In Table D.9, we augment the baseline specification controlling for 1880–1920 average predicted precipitation distance (Panel B), 1880–1920 average predicted geographic distance (Panel C), and the two jointly (Panel D). Across specifications, the magnitude and statistical significance of the coefficients remain remarkably stable.

4 Conclusion

In this paper, we provide systematic evidence that individuals tend to settle in places with climates similar to where they came from. These patterns hold across periods, geography, and migrant groups, and are not explained by the spatial correlation of climate or by other well-known determinants of migration—including pre-existing migrant networks and geographic distance. Across our main specifications, a 1°C reduction in temperature distance increases migration by 2 to 27%, underscoring that climate similarity is a first-order determinant of migrant settlement. Exploring the mechanisms, we show that both climate-specific skills and climate-as-amenity likely contribute to climate matching, suggesting that climate distance acts as a "shadow border" that constrains the spatial allocation of human capital.

We then examine the implications of climate matching for the economic geography of the United States. Focusing on international immigration between 1880 and 1920—a period of rapid growth and structural transformation (Eckert and Peters, 2025)—we construct a measure of a county's climate mismatch as the average temperature distance of the immigrants who settled there in each decade. To address endogeneity concerns, we develop an instrument that exploits the interaction between the timing of national immigration waves by country of origin and the staggered roll-out of the U.S. railroad network.

We document that climate mismatch slowed the growth of agricultural output and productivity. Although mismatch accelerated the reallocation of labor out of agriculture, it did not translate into higher manufacturing output, indicating slower productivity growth in more mismatched counties. Rather than generating a compensating industrial boom, mismatch acted as a friction to structural transformation—pushing labor out of farming while depressing productivity in the absorbing manufacturing sector. These effects left an imprint: areas with high climate mismatch exhibited lower population growth and had lower income per capita in 1940. The income effects are similar for immigrants and U.S.-born men, indicating that the initial climatic alignment of a population can shape the long-run economic prosperity of receiving areas. The persistence of these effects underscores the limits of human adaptation to environmental differences.

Our findings open several avenues for future research. First, while we document that migrants match on climate, the precise reasons why temperature plays such a central role remain incompletely understood—specifically the degree to which it captures expectations about individual productivity versus cultural, psychological, or physiological channels. Second, it would be valuable to test whether climate matching operates outside the United States, in settings with different economic structures, mobility constraints, and institutional environments, where the welfare and distributional consequences of mismatch may differ. Third, our climate-similarity measure may help refine or complement shift–share strategies that rely on historical settlement patterns (Altonji and Card, 1991; Card, 2001), offering a new way to predict migration flows and to study the economic, political, and cultural impacts of immigration. Finally, our results suggest that climate distance functions as an endogenous migration cost, implying that climate change may reshape mobility frictions themselves, with important implications for spatial adjustment and long-run inequality

References

Abatzoglou, J. T., S. Z. Dobrowski, S. A. Parks, and K. C. Hegewisch (2018): "TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015," *Scientific data*, 5, 1–12.

Abramitzky, R. and L. Boustan (2017): "Immigration in American economic history," *Journal of economic literature*, 55, 1311–1345.

Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. Pérez (2021): "Automated linking of historical data," *Journal of Economic Literature*, 59, 865–918.

Abramitzky, R., L. Boustan, K. Eriksson, S. Perez, and M. Rashid (2020): "Census Linking Project: Version 2.0 [dataset]." *Data retrieved from, https://censuslinkingproject. org.*

Adão, R., M. Kolesár, and E. Morales (2019): "Shift-Share Designs: Theory and Inference," *Quarterly Journal of Economics*, 134, 1949–2010.

Albouy, D., H. Cho, and M. Shappo (2021): "Immigration and the pursuit of amenities," *Journal of Regional Science*, 61, 5–29.

Albouy, D., W. Graf, R. Kellogg, and H. Wolff (2016): "Climate amenities, climate change, and American quality of life," *Journal of the Association of Environmental and Resource Economists*, 3, 205–246.

Alston, L. J. and J. P. Ferrie (1999): *Southern Paternalism and the American Welfare State: Economics, Politics, and Institutions in the South, 1865-1965*, Cambridge University Press New York.

Altonji, J. G. and D. Card (1991): "The effects of immigration on the labor market outcomes of less-skilled natives," in *Immigration, trade, and the labor market*, University of Chicago Press, 201–234.

Anderson, J. E. and E. Van Wincoop (2003): "Gravity with gravitas: A solution to the border puzzle," *American economic review*, 93, 170–192.

Atack, J. (2016): "Historical Geographic Information Systems (GIS) Database of U.S. Railroads for 1826–1911," May 2016; revised 2023.

Atack, J. and R. A. Margo (2011): "The impact of access to rail transportation on agricultural improvement: The American Midwest as a test case, 1850–1860," *Journal of Transport and Land Use*, 4, 5–18.

Bazzi, S., A. Ferrara, M. Fiszbein, T. Pearson, and P. A. Testa (2023): "The Other Great Migration: Southern Whites and the New Right," *The Quarterly Journal of Economics*, 138, 1577–1647.

Bazzi, S., M. Fiszbein, and M. Gebresilasse (2020): "Frontier culture: The roots and persistence of "rugged individualism" in the United States," *Econometrica*, 88, 2329–2368.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2016): "Skill transferability, migration, and development: Evidence from population resettlement in Indonesia," *American Economic Review*, 106, 2658–2698.

Beine, M., F. Docquier, and Ç. Özden (2011): "Diasporas," *Journal of Development Economics*, 95, 30–41.

Berkes, E., E. Karger, and P. Nencka (2023): "The census place project: A method for geolocating unstructured place names," *Explorations in Economic History*, 87, 101477.

Biddle, J. (2008): "Explaining the spread of residential air conditioning, 1955–1980," *Explorations in Economic History*, 45, 402–423.

Bilal, A. and E. Rossi-Hansberg (2025): "Anticipating Climate Change Across the United States," Working Paper 31323, National Bureau of Economic Research.

Bishop, B. and R. G. Cushing (2009): *The big sort: Why the clustering of like-minded America is tearing us apart*, Houghton Mifflin Harcourt.

Borjas, G. J. (2001): "Does immigration grease the wheels of the labor market?" *Brookings papers on economic activity*, 2001, 69–133.

Borusyak, K., P. Hull, and X. Jaravel (2022): "Quasi-experimental shift-share research designs," *The Review of Economic Studies*, 89, 181–213.

Boustan, L. P. (2010): "Was postwar suburbanization "white flight"? Evidence from the black migration," *The Quarterly Journal of Economics*, 125, 417–443.

Bryan, G. and M. Morten (2019): "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 127, 2229–2268.

Buckles, K., A. Haws, J. Price, and H. E. Wilbert (2025): "Breakthroughs in historical record linking using genealogy data: The census tree project," *Explorations in Economic History*, 101717.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019): "Migrants, ancestors, and foreign investments," *The Review of Economic Studies*, 86, 1448–1486.

Burke, M. and K. Emerick (2016): "Adaptation to climate change: Evidence from US agriculture," *American Economic Journal: Economic Policy*, 8, 106–140.

Cadena, B. C. and B. K. Kovak (2016): "Immigrants equilibrate local labor markets: Evidence from the Great Recession," *American Economic Journal: Applied Economics*, 8, 257–290.

Caliendo, L., L. D. Opromolla, F. Parro, and A. Sforza (2021): "Goods and factor market integration: A quantitative assessment of the EU enlargement," *Journal of Political Economy*, 129, 3491–3545.

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

Card, D. (2001): "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 19, 22–64.

Carleton, T. A. and S. M. Hsiang (2016): "Social and economic impacts of climate," *Science*, 353, aad9837.

Carroll, M. L., J. R. Townshend, C. M. DiMiceli, P. Noojipady, and R. A. Sohlberg (2009): "A new global raster water mask at 250 m resolution," *International Journal of Digital Earth*, 2, 291–308.

Colmer, J. (2021): "Temperature, labor reallocation, and industrial production: Evidence from India," *American Economic Journal: Applied Economics*, 13, 101–124.

Coolidge, C. (1926): *Foundations of the Republic; Speeches and Addresses*, New York, Scribner.

Crosby, A. W. (2004): *Ecological imperialism: the biological expansion of Europe, 900-1900*, Cambridge University Press.

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

Danielson, J. J. and D. B. Gesch (2011): "Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010)," Open-File Report 2011-1073, U.S. Geological Survey.

De Chaisemartin, C. and X. d'Haultfoeuille (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–2996.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2025): "Human Capital Accumulation Across Space," Tech. rep., National Bureau of Economic Research.

Desmet, K. and E. Rossi-Hansberg (2024): "Climate change economics over time and space," *Annual Review of Economics*, 16.

Donaldson, D. and R. Hornbeck (2016): "Railroads and American economic growth: A "market access" approach," *The Quarterly Journal of Economics*, 131, 799–858.

Eckert, F., A. Gvirtz, J. Liang, and M. Peters (2020): "A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790," Tech. rep., National Bureau of Economic Research.

Eckert, F., J. Juneau, and M. Peters (2023): "Sprouting Cities: How Rural America Industrialized," in *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 113, 87–92.

Eckert, F. and M. Peters (2025): "Spatial structural change," Tech. rep., National Bureau of Economic Research.

Fenske, J. (2014): "Ecology, trade, and states in pre-colonial Africa," *Journal of the European Economic Association*, 12, 612–640.

Fischer, D. H. (1989): *Albion's seed: Four British folkways in America*, vol. 1, America: A Cultural History.

Fishback, P. V., W. C. Horrace, and S. Kantor (2006): "The impact of New Deal expenditures on mobility during the Great Depression," *Explorations in Economic History*, 43, 179–222.

Fiszbein, M. (2022): "Agricultural diversity, structural change, and long-run development: Evidence from the United States," *American Economic Journal: Macroeconomics*, 14, 1–43.

Fogel, R. W. (1964): *Railroads and American economic growth*, Johns Hopkins Press Baltimore.

Glaeser, E. L. and K. Tobio (2007): *The Rise of the Sunbelt*, National Bureau of Economic Research.

Goldewijk, K. K., A. Beusen, J. Doelman, and E. Stehfest (2017): "Anthropogenic land use estimates for the Holocene–HYDE 3.2," *Earth System Science Data*, 9, 927–953.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020): "Bartik instruments: What, when, why, and how," *American Economic Review*, 110, 2586–2624.

Goodman-Bacon, A. (2021): "Difference-in-differences with variation in treatment timing," *Journal* *of Econometrics*, 225, 254–277.

Green, A. (2025): "Networks and geographic mobility: Evidence from world war II navy ships," Tech. rep., Working Paper.

Haines, M. R., P. Fishback, and P. Rhode (2018): "United States Agriculture Data, 1840 - 2012," Inter-university Consortium for Political and Social Research [distributor].

Haines, M. R. and Inter-university Consortium for Political and Social Research (2000): "Historical, Demographic, Economic, and Social Data: The United States, 1790–2000, Census of Agriculture Files," ICPSR dataset.

——— (2010): "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002," Inter-university Consortium for Political and Social Research [distributor].

Harris, I., T. J. Osborn, P. Jones, and D. Lister (2020): "Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset," *Scientific Data*, 7.

Hatton, T. J. and J. G. Williamson (1998): *The age of mass migration: Causes and economic impact*, Oxford University Press on Demand.

Hengl, T. (2018a): "Soil bulk density (fine earth) 10 x kg / m-cubic at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (v0.2)," .

——— (2018b): "Soil pH in H2O at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (v0.2)," .

Hengl, T. and S. Gupta (2019): "Soil water content (volumetric %) for 33kPa and 1500kPa suctions predicted at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (v0.1)," .

Hengl, T. and I. Wheeler (2018): "Soil organic carbon content in x 5 g / kg at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (v0.2)," .

Hornbeck, R. (2010): "Barbed wire: Property rights and agricultural development," *The Quarterly Journal of Economics*, 125, 767–810.

——— (2012): "The enduring impact of the American Dust Bowl: Short-and long-run adjustments to environmental catastrophe," *American Economic Review*, 102, 1477–1507.

——— (2023): "Dust Bowl Migrants: Environmental Refugees and Economic Adaptation," *The Journal of Economic History*, 83, 645–675.

Hornbeck, R. and S. Naidu (2014): "When the levee breaks: black migration and economic development in the American South," *American Economic Review*, 104, 963–990.

Hornbeck, R. and M. Rotemberg (2024): "Growth off the rails: Aggregate productivity growth in distorted economies," *Journal of Political Economy*, 132, 3547–3602.

Hunt, J. (1992): "The Impact of the 1962 Repatriates from Algeria on the French Labor Market," *Industrial and Labor Relations Review*, 45, 556–572.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-share instruments and the impact of immigration," Tech. rep., National Bureau of Economic Research.

King, D. (2002): *Making Americans: Immigration, race, and the origins of the diverse democracy*, Harvard University Press.

Koenen, M. and D. Johnston (2024): "Social ties and residential choice: Micro evidence and equilibrium implications," Tech. rep., Technical report, Harvard University.

Kupperman, K. O. (1984): "Fear of Hot Climates in the Anglo-American Colonial Experience," *The William and Mary Quarterly*, 41, 213–240.

Lebergott, S. (1966): "Labor force and employment, 1800–1960," in *Output, Employment, and Productivity in the United States after 1800*, NBER, 117–204.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner (2004): "European seasonal and annual temperature variability, trends, and extremes since 1500," *Science*, 303, 1499–1503.

Medici, C. (2024): "Closing ranks: Organized labor and immigration," Tech. rep., CESifo.

Michalopoulos, S. (2012): "The origins of ethnolinguistic diversity," *American Economic Review*, 102, 1508–1539.

Morten, M. and J. Oliveira (2024): "The effects of roads on trade and migration: Evidence from a planned capital city," *American Economic Journal: Applied Economics*, 16, 389–421.

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the US Labor Market," *Quarterly Journal of Economics*, 118, 549–599.

——— (2020): "Social networks and migration," *Annual Review of Economics*, 12, 503–524.

Munshi, K. and M. Rosenzweig (2016): "Networks and misallocation: Insurance, migration, and the rural-urban wage gap," *American Economic Review*, 106, 46–98.

Nagy, D. K. (2023): "Hinterlands, city formation and growth: Evidence from the US westward expansion," *Review of Economic Studies*, 90, 3238–3281.

Nichols, R. L. (2014): *American Indians in US history*, vol. 248, University of Oklahoma Press.

Olmstead, A. L. and P. W. Rhode (2002): "The red queen and the hard reds: Productivity growth in American wheat, 1800–1940," *The Journal of Economic History*, 62, 929–966.

——— (2008): *Creating Abundance: Biological Innovation and American Agricultural Development*, Cambridge: Cambridge University Press.

Pauling, A., J. Luterbacher, C. Casty, and H. Wanner (2006): "Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation," *Climate dynamics*, 26, 387–405.

Pellegrina, H. S. and S. Sotelo (2025): "Migration, Specialization, and Trade: Evidence from Brazil's March to the West," *Journal of Political Economy*, 133, 3993–4049.

Peters, M. (2022): "Market size and spatial growth—evidence from Germany's post-war population expulsions," *Econometrica*, 90, 2357–2396.

Price, J., K. Buckles, J. Van Leeuwen, and I. Riley (2021): "Combining family history and machine learning to link historical records: The Census Tree data set," *Explorations in Economic History*, 80, 101391.

Proctor, J., A. Rigden, D. Chan, and P. Huybers (2022): "More accurate specification of water supply shows its importance for global crop production," *Nature Food*, 3, 753–763.

Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney (2002): "The global distribution of cultivable lands: current patterns and sensitivity to possible climate change," *Global Ecology and*

Biogeography, 11, 377–392.

Raz, I. T. (2025): “Soil Heterogeneity, Social Learning, and the Formation of Close-Knit Communities,” *Journal of Political Economy*, 133, 2643–2691.

Rosenbloom, J. L. (2002): *Looking for work, searching for workers: American labor markets during industrialization*, Cambridge University Press.

Ruggles, S., C. A. Fitch, G. Ronald, H. J. David, M. A. Nelson, E. Roberts, M. Schouweiler, and M. Sobek (2021): “IPUMS Ancestry Full Count Data: Version 3.0 [dataset],” *Minneapolis, MN: IPUMS*.

Schlenker, W., W. Michael Hanemann, and A. C. Fisher (2005): “Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach,” *American Economic Review*, 95, 395–406.

Schlenker, W. and M. J. Roberts (2009): “Nonlinear temperature effects indicate severe damages to US crop yields under climate change,” *Proceedings of the National Academy of sciences*, 106, 15594–15598.

Schwartz, A. (1973): “Interpreting the effect of distance on migration,” *Journal of Political Economy*, 81, 1153–1169.

Sequeira, S., N. Nunn, and N. Qian (2020): “Immigrants and the Making of America,” *Review of Economic Studies*, 87, 382–419.

Sinha, P., M. Caulkins, and M. Cropper (2021): “The value of climate amenities: A comparison of hedonic and discrete choice approaches,” *Journal of Urban Economics*, 126, 103371.

Steckel, R. H. (1983): “The economic foundations of East-West migration during the 19th century,” *Explorations in Economic History*, 20, 14–36.

Stuart, B. A. and E. J. Taylor (2021): “Migration networks and location decisions: Evidence from US mass migration,” *American Economic Journal: Applied Economics*, 13, 134–175.

Tabellini, M. (2020): “Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration,” *The Review of Economic Studies*, 87, 454–486.

Taylor, C. A. (2022): “Irrigation and Climate Change: Long-run Adaptation and its Externalities,” *Working Paper*.

Terry, S. J., T. Chaney, K. B. Burchardi, L. Tarquinio, and T. A. Hassan (2024): “Immigration, innovation, and growth,” *American Economic Review, Forthcoming*.

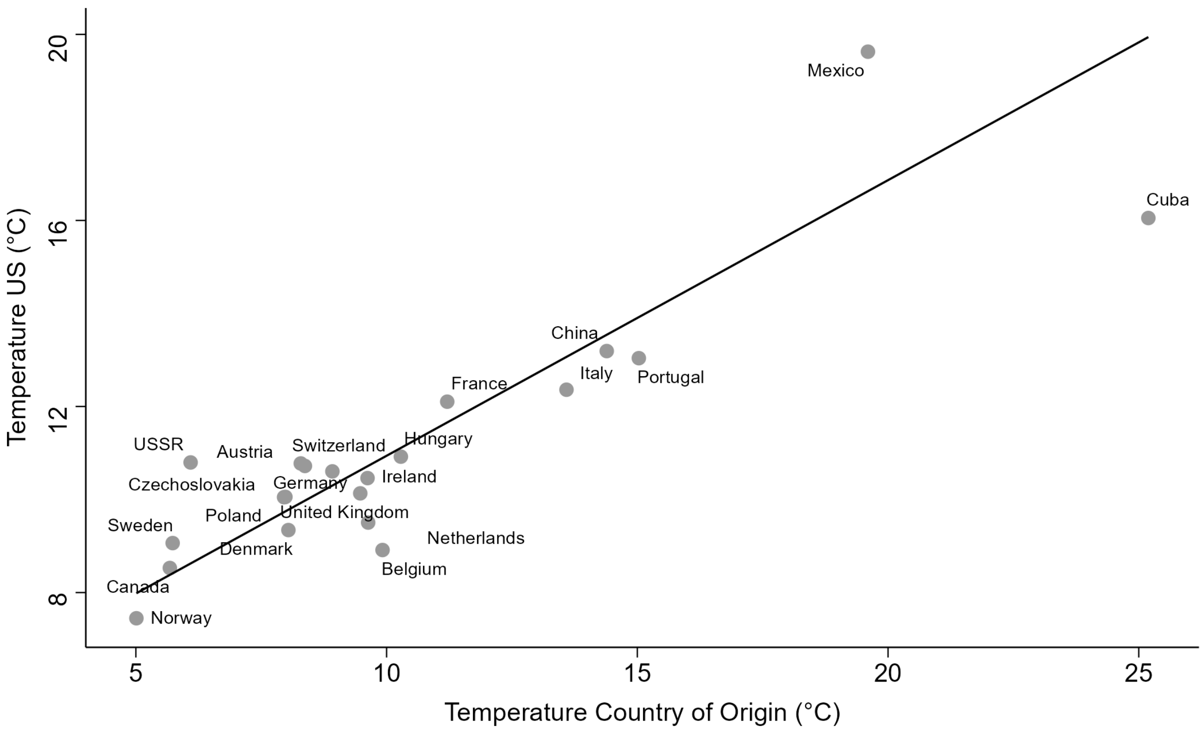
Tobler, W. R. (1970): “A computer movie simulating urban growth in the Detroit region,” *Economic geography*, 46, 234–240.

Tombe, T. and X. Zhu (2019): “Trade, migration, and productivity: A quantitative analysis of China,” *American Economic Review*, 109, 1843–1872.

Willcox, W. F. (1929): *Statistics of Migrations, National Tables, United States*, NBER, 372–498.

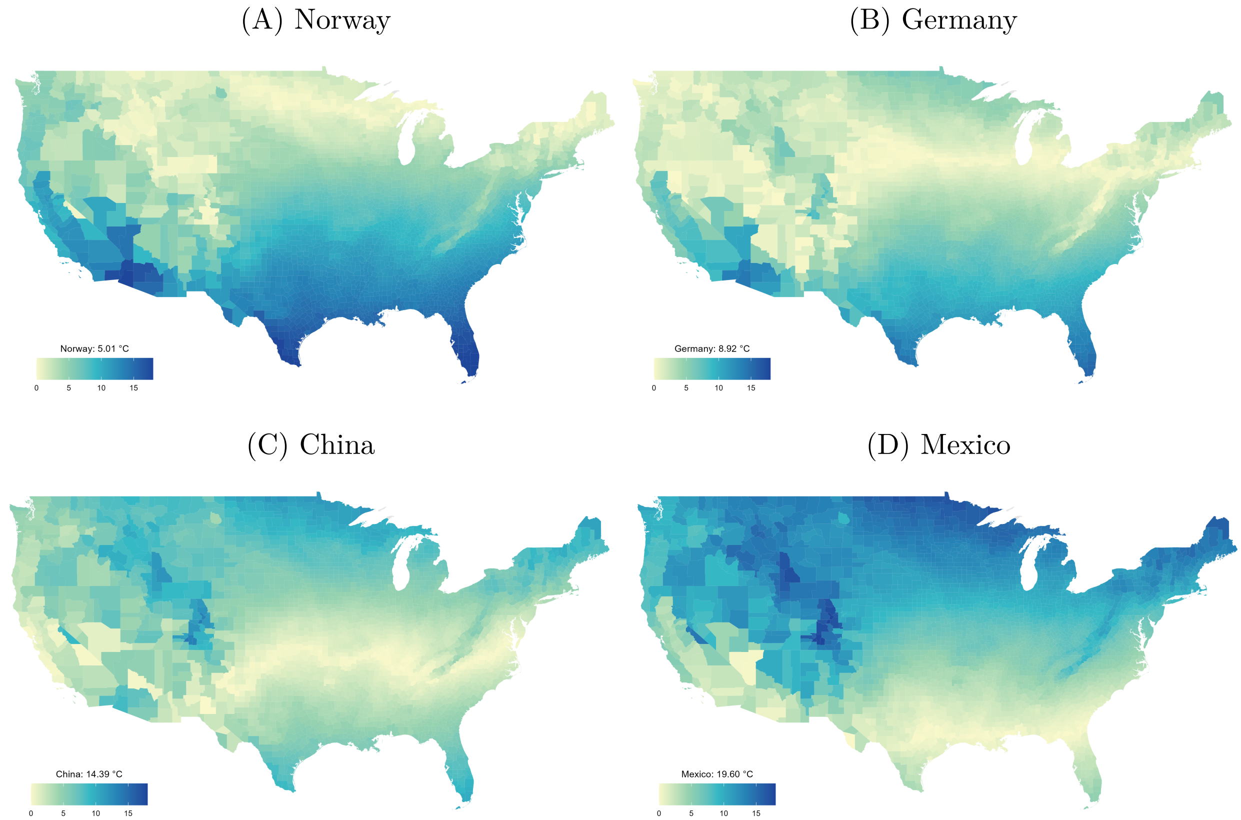
Figures and Tables

Figure 1. Temperature Matching for Immigrants in the U.S. (1880)



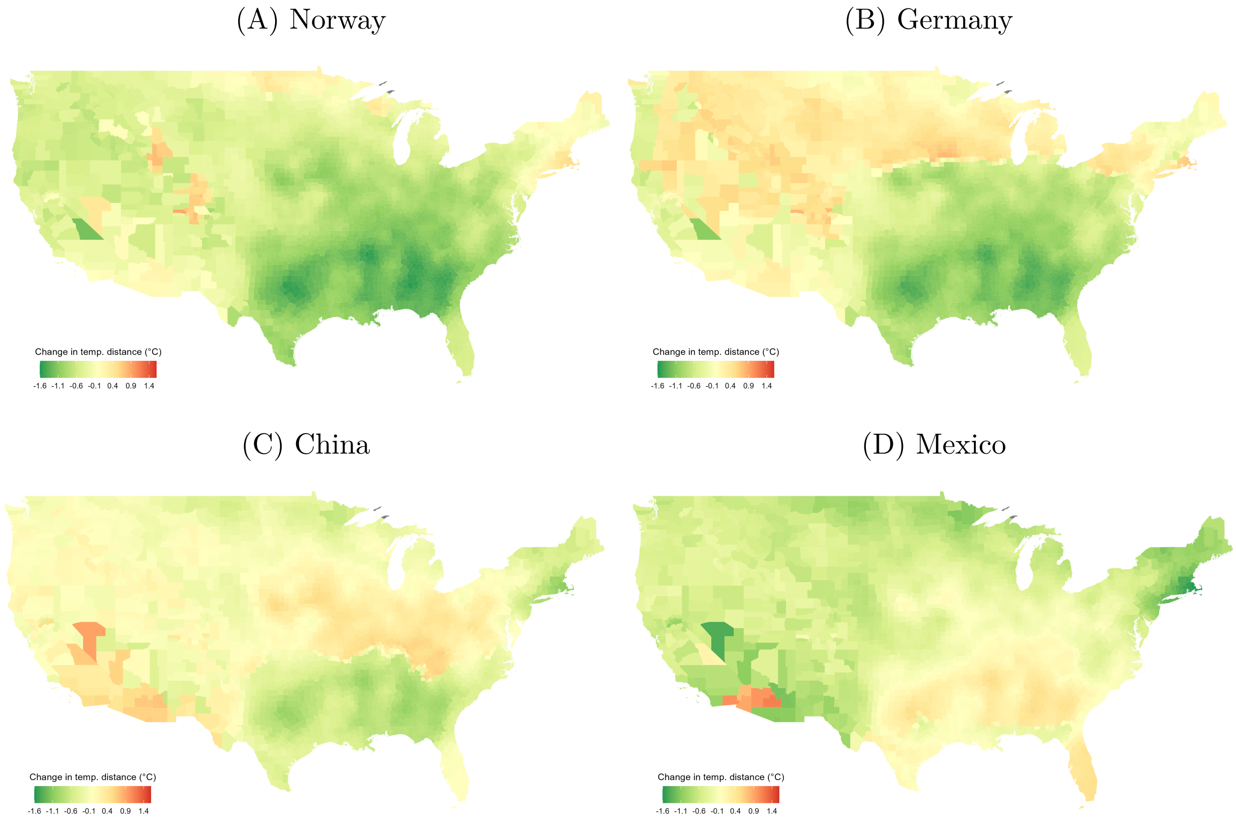
Notes: The figure plots the relationship between temperature in immigrants' countries of origin (x-axis) and the weighted average temperature across U.S. counties in which immigrants from each origin resided in 1880 (y-axis), with weights given by each county's share of immigrants from that origin. The sample includes immigrants living in the contiguous United States in 1880 from countries accounting for at least 0.1% of the foreign-born population at the time. Country boundaries are harmonized to their 1990 definitions following Burchardi et al. (2019). Temperature, in degrees Celsius, is measured as mean annual temperature over the 1960–2000 period using TerraClimate data (Abatzoglou et al., 2018). Regressions are weighted by the number of individuals from each origin country residing in the United States. The estimated regression coefficient is 0.593, with a robust standard error of 0.087.

Figure 2. Temperature Distance Between Selected Countries and U.S. Counties



Notes: The figure shows the absolute difference in temperature between each U.S. county and four example countries (Norway, Germany, China, Mexico). Temperature is the mean annual temperature in degrees Celsius averaged over the period 1960–2000 using TerraClimate data (Abatzoglou et al., 2018).

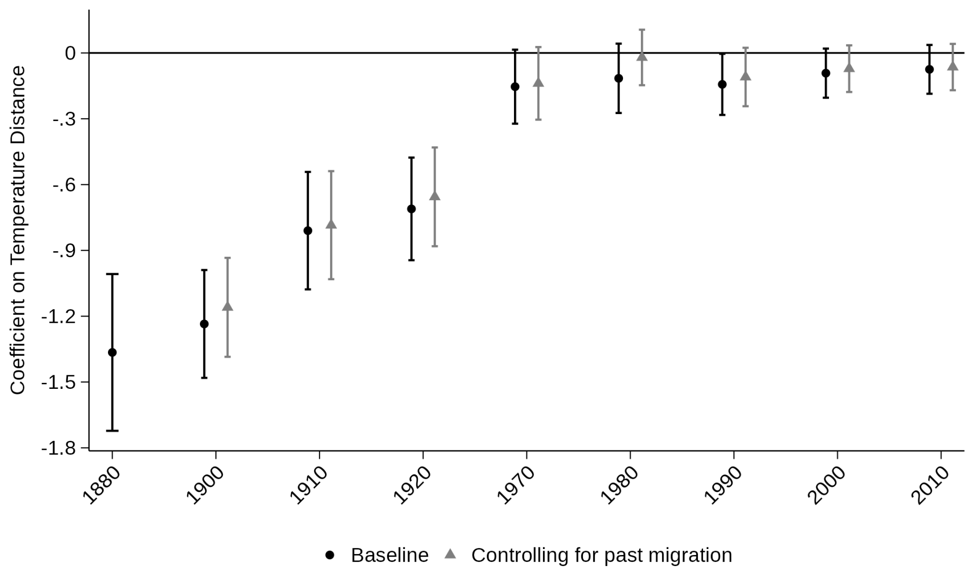
Figure 3. Change in Temperature Distance Between Selected Countries and U.S. Counties



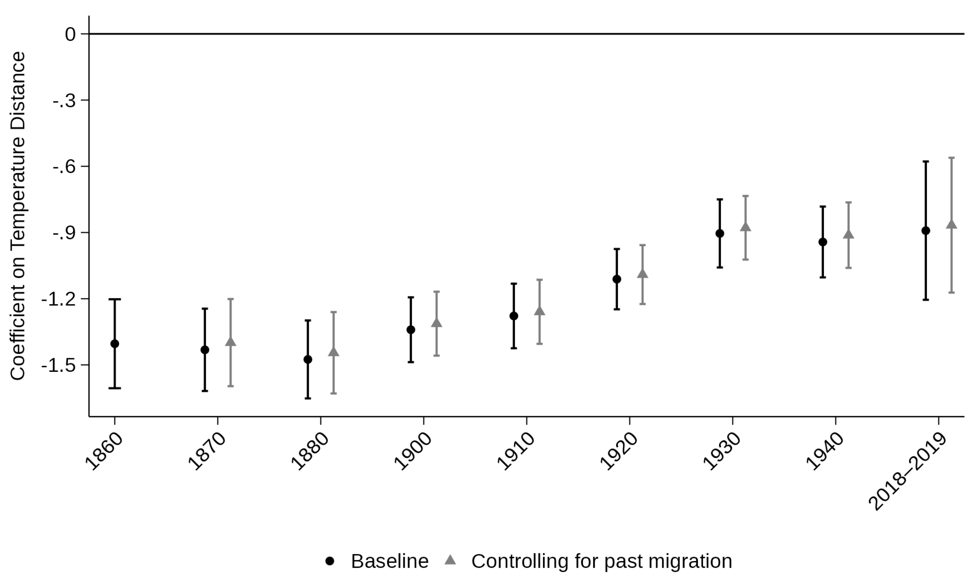
Notes: The figure plots the change in the absolute difference in temperature between each U.S. county and four example countries (Norway, Germany, China, Mexico). Temperature is the mean annual temperature in degrees Celsius. These averages are taken over the periods 1901–1930 and 1991–2020 using CRU climate data (Harris et al., 2020). Green areas represent climate convergence over time, and red areas climate divergence.

Figure 4. Climate Distance and Migration, by Period and Sample

(A) International Migration



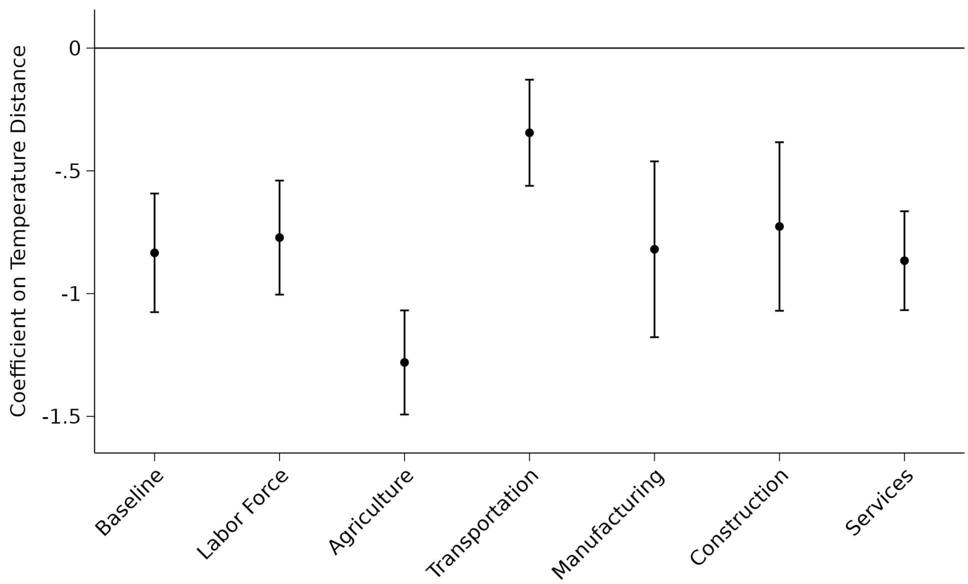
(B) U.S. Internal Migration



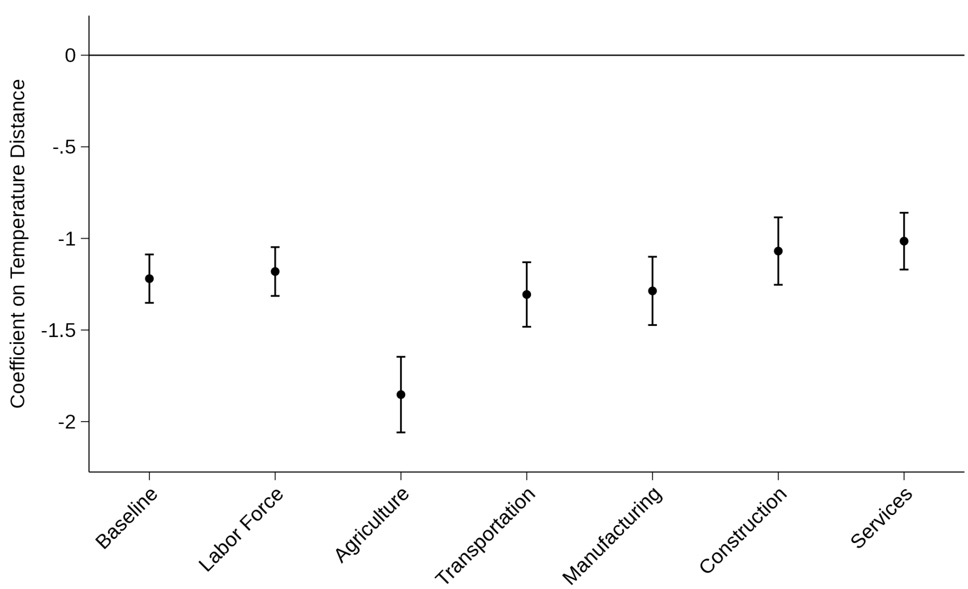
Notes: The figure plots the coefficient, with corresponding 95% confidence intervals, on the absolute difference in temperature between (A) international country of birth and U.S. destination county, and (B) U.S. origin and destination counties for internal movers. Dots refer to the baseline specification, described in the notes of Table 1. Triangles refer to specifications that further control for lagged number of migrants from a given origin to a given destination county. The dependent variable is the number of immigrants over the corresponding period. Temperature distances are standardized within the relevant sample to have zero mean and standard deviation equal to one. Standard errors are clustered at the country (resp., U.S. state) of origin by U.S. state of destination level.

Figure 5. Climate Distance and Migration: Heterogeneity by Sector

(A) International Migration



(B) U.S. Internal Migration



Notes: The figure plots the coefficient, with corresponding 95% confidence intervals, on the absolute difference in temperature between (A) international country of birth and U.S. destination county, and (B) U.S. origin and destination counties for internal movers. The first dot in Panel A (resp., Panel B) shows the baseline results of column 1 (resp., column 3) of Table 1. From the second dot onward, the number of migrants refers to migrant men in the sector reported on the x-axis (defined according to 3-digit IND1950 code from IPUMS). For internal migrants, sector is measured in the baseline decade (before the move). Temperature distances are standardized within the relevant sample to have zero mean and standard deviation equal to one. Standard errors are clustered at the country (resp., U.S. state) of origin by U.S. state of destination level.

Table 1. Climate Distance and Migration: Main Results

Dep. Variable:	Number of Migrants			
	Intl-US 1880–1920	Intl-US 1970–2010	Internal 1850–1940	Internal 2011–2019
	(1)	(2)	(3)	(4)
Temperature Distance	-0.175***	-0.019*	-0.315***	-0.237***
	(0.021)	(0.011)	(0.019)	(0.041)
Precipitation Distance	-0.002	-0.008***	-0.017***	-0.030***
	(0.004)	(0.002)	(0.002)	(0.006)
Distance (100 km)	-0.064***	-0.051***	-0.194***	-0.204***
	(0.013)	(0.004)	(0.014)	(0.022)
Observations	1,939,195	2,475,100	31,092,699	25,088,936
Pseudo R-squared	0.916	0.893	0.728	0.776
Mean Temp. Dist.	9.622	9.621	4.923	4.989
SD Temp. Dist.	5.754	5.685	3.578	3.632
Mean Precip. Dist.	52.53	53.96	31.01	31.06
SD Precip. Dist.	42.07	43.90	23.76	24.01
Origin×Decade/Year FE	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes

Notes: The samples in columns 1–2 include U.S. county-country pairs for each decade over the 1880–1920 (except for 1890) and 1970–2010 periods, respectively. The definition of countries of origin is fixed to 1990 boundaries and harmonized following the procedure in Burchardi et al. (2019). The samples in columns 3–4 include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each calendar year from 2011–2012 to 2018–2019, respectively. In columns 1–2, Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade (except for 1880, when information on year of arrival is not available). In column 3, Number of Migrants is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. In column 4, Number of Migrants is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. In columns 1–2 (resp., columns 3–4), Origin o refers to the origin country (resp., the origin U.S. county). Destination d refers to a U.S. county in all columns. All regressions control for origin by decade (columns 1–3) or year (column 4) fixed effects and destination by decade (columns 1–3) or year (column 4) fixed effects. Standard errors, reported in parentheses, are clustered at the country (resp., U.S. state) of origin by U.S. state of destination in columns 1–2 (resp., columns 3–4) level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2. Climate Distance and Migration in the Long Run

Dep. Variable:	Number of Migrants			
	Intl-US	Intl-US	Internal	Internal
	(1)	(2)	(3)	(4)
Temperature Distance	-0.326***	-0.307***	-0.437***	-0.310***
	(0.124)	(0.111)	(0.102)	(0.075)
Precipitation Distance	-0.021	-0.016	-0.040***	-0.033***
	(0.014)	(0.015)	(0.009)	(0.008)
Observations	814,336	915,866	899,850	2,257,910
Pseudo R-squared	0.991	0.990	0.966	0.966
Mean Temp. Dist.	9.388	9.552	3.174	3.748
SD Temp. Dist.	5.831	5.855	2.859	3.193
Mean Precip. Dist.	50.63	51.77	21.83	24.58
SD Precip. Dist.	40.16	41.10	20.71	21.66
Origin×Period FE	Yes	Yes	Yes	Yes
Destination×Period FE	Yes	Yes	Yes	Yes
Origin×Destination FE	Yes	Yes	Yes	Yes
Distance×Period FE	Yes	Yes	Yes	Yes
Migration Historical Period	1900	Avg. 1900–1920	1900–1910	Avg. 1910–1940
Migration Modern Period	2010	Avg. 1990–2010	2018–2019	Avg. 2011–2019

Notes: The samples in columns 1–2 include U.S. county-country pairs for the two periods defined at the bottom of the table. The definition of countries of origin is fixed to 1990 boundaries and harmonized following the procedure in Burchardi et al. (2019). The samples in columns 3–4 include county-pairs in the contiguous U.S. for the two periods defined at the bottom of the table. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation), measured in degrees Celsius (resp., millimeters), between the origin and the destination in the historical (1901-1930) and modern (1991-2020) periods. In column 1–2 (resp., columns 3–4), Origin o refers to the origin country (resp., the origin U.S. county). Destination d refers to a U.S. county in all columns. All regressions include the absolute difference in precipitation between the origin and the destination in the historical and modern periods, origin by period fixed effects, destination by period fixed effects, origin by destination fixed effects, and the interaction between the modern period dummy and geographic distance. Standard errors, reported in parentheses, are clustered at the country (resp., U.S. state) of origin by U.S. state of destination level in columns 1–2 (resp., columns 3–4). Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 3. The Effects of Climate Mismatch on Agriculture

Dep. Variable:	Log Value of Crops	Share Farmland	Log Value per Acre: Crop	Log Value per Acre: Farm	Log Value per Acre: Equipment	Log Number of Farms	Log Avg Farm Size
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS							
Temperature Mismatch	-0.036***	-0.006***	-0.021***	-0.032***	-0.036***	-0.024***	0.004
	(0.006)	(0.001)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
R-squared	0.944	0.924	0.948	0.958	0.956	0.923	0.858
Panel B: 2SLS							
Temperature Mismatch	-0.178***	-0.027***	-0.100***	-0.147***	-0.106***	-0.139***	0.024
	(0.029)	(0.007)	(0.022)	(0.022)	(0.019)	(0.018)	(0.017)
R-squared	0.945	0.925	0.948	0.958	0.956	0.924	0.858
KP F-stat	610.08	610.08	610.08	610.08	610.08	610.08	610.08
Panel C: First Stage							
Predicted Mismatch	0.649***	0.649***	0.649***	0.649***	0.649***	0.649***	0.649***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
R-squared	0.908	0.908	0.908	0.908	0.908	0.908	0.908
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
1880 Dep. Var. Mean	512,108	0.638	2.121	18.59	0.750	1,774	145.9
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 and at least 100 farms in each decade from 1880 to 1920. The year 1890 is excluded, because it is not possible to construct the measure of climate mismatch. Panels A and B report OLS and 2SLS estimates, respectively. Panel C reports the first stage. Temperature Mismatch is the weighted average of the absolute temperature difference (in degrees Celsius) between the county of destination and the country of origin of the migrants, with weights equal to the share of migrants from each origin arriving in the county in each decade, relative to all migrants arriving in the county in that decade. See Section 3.1 and equation (4). Predicted Mismatch (Panel C) is the corresponding measure of predicted average temperature distance, constructed as described in Section 3.1. The dependent variables are the log value of crops (column 1), the share of land in farms (column 2), the log per-acre value of crops (column 3), farm value (column 4), and equipment value (column 5), the log number of farms (column 6), and the log average farm size (column 7). All regressions include county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and, in Panels B and C, are further adjusted using the block bootstrap procedure described in Section 3.1. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 4. The Effects of Climate Mismatch on Manufacturing

Dep. Variable:	Non-Ag Emp Share	Log Manufacturing: Output per Worker	Log Manufacturing: Avg Wages	Log Manufacturing: Output	Log Manufacturing: # Establishments
	(1)	(2)	(3)	(4)	(5)
Panel A: OLS					
Temperature Mismatch	0.005***	-0.030***	-0.001	-0.073***	-0.046*
	(0.002)	(0.008)	(0.005)	(0.027)	(0.025)
R-squared	0.857	0.759	0.906	0.794	0.703
Panel B: 2SLS					
Temperature Mismatch	0.037***	-0.089***	-0.011	-0.141	-0.144*
	(0.007)	(0.029)	(0.015)	(0.096)	(0.083)
R-squared	0.858	0.759	0.906	0.794	0.703
KP F-stat	590.40	415.57	415.57	415.82	415.57
Panel C: First Stage					
Predicted Mismatch	0.651***	0.661***	0.661***	0.661***	0.661***
	(0.036)	(0.041)	(0.041)	(0.041)	(0.041)
R-squared	0.907	0.919	0.919	0.919	0.919
Observations	9,196	6,895	6,895	6,896	6,895
1880 Dep. Var. Mean	0.389	9,589	265.5	2,253,570	107.2
County FE	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880, for the period 1880–1920. The year 1890 is excluded, because it is not possible to construct the measure of climate mismatch. In columns 2–5, 1910 is also excluded, because data from the Census of Manufactures are missing for that year. Panels A and B report OLS and 2SLS estimates, respectively. Panel C reports the first stage. Temperature Mismatch is the weighted average of the absolute temperature difference (in degrees Celsius) between the county and the country of origin of the migrants, with weights equal to the share of migrants from each origin arriving in the county in each decade, relative to all migrants arriving in the county in that decade. See Section 3.1 and equation (4). Predicted Mismatch (Panel C) is the corresponding measure of predicted average temperature distance, constructed as described in Section 3.1. In column 1, the dependent variable is the number of men 15-64 employed in sectors other than agriculture, relative to all men 15-64 in the labor force. In columns 2–5, the dependent variables are the log: manufacturing output per worker (column 2); average manufacturing wages (column 3); manufacturing output (column 4); number of establishments (column 5). All regressions include county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and, in Panels B and C, are further adjusted using the block bootstrap procedure described in Section 3.1. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Climate Mismatch, Population Growth, and 1940 Income per Capita

Dep. Variable:	Log Population	All men: Log Average Income	All men: Log Hourly Earnings	U.S.-born men: Log Average Income	U.S.-born men: Log Hourly Earnings	Immigrant men: Log Average Income	Immigrant men: Log Hourly Earnings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS							
Temperature Mismatch	-0.044***	-0.034***	-0.050***	-0.036***	-0.051***	0.008	-0.023***
	(0.007)	(0.003)	(0.004)	(0.003)	(0.004)	(0.006)	(0.006)
R-squared	0.895	0.338	0.407	0.351	0.416	0.176	0.148
Panel B: 2SLS							
Temperature Mismatch	-0.108***	-0.027***	-0.053***	-0.024***	-0.050***	-0.036***	-0.061***
	(0.031)	(0.005)	(0.005)	(0.005)	(0.005)	(0.010)	(0.009)
R-squared	0.894	0.317	0.378	0.325	0.381	0.182	0.160
KP F-stat	590.41	644.98	646.36	644.98	646.36	569.77	560.23
Panel C: First Stage							
Predicted Mismatch	0.651***	0.622***	0.623***	0.622***	0.623***	0.608***	0.608***
	(0.038)	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)
R-squared	0.907	0.747	0.750	0.747	0.750	0.736	0.738
Observations	9,197	2,890	2,867	2,890	2,867	2,752	2,707
County FE	Yes						
Region×Decade FE	Yes						
Region FE		Yes	Yes	Yes	Yes	Yes	Yes
Railroad Arrival Year		Yes	Yes	Yes	Yes	Yes	Yes

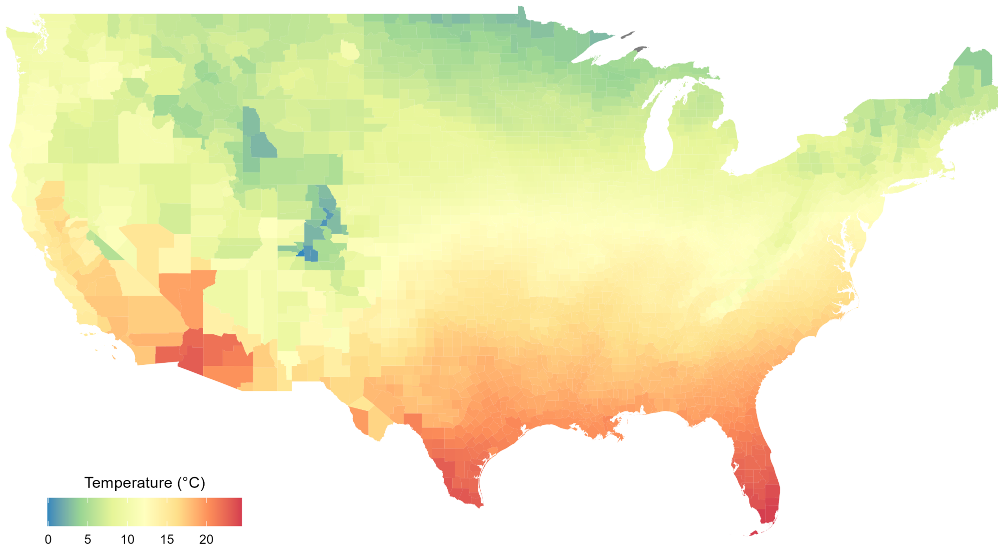
Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 for the period 1880–1920 (column 1) and for 1940 (columns 2–7). In column 1, the year 1890 is excluded, because it is not possible to construct the measure of climate mismatch. Panels A and B report OLS and 2SLS estimates, respectively. Panel C reports the first stage. In column 1, Temperature Mismatch is the weighted average of the absolute temperature difference (in degrees Celsius) between the county and the country of origin of the migrants, with weights equal to the share of migrants from each origin arriving in the county in each decade, relative to all migrants arriving in the county in that decade. See Section 3.1 and equation (4). In columns 2–7, Temperature Mismatch is averaged over the 1880–1920 period. Predicted Mismatch (Panel C) is the corresponding measure of predicted average temperature distance, constructed as described in Section 3.1. In column 1, the dependent variable is log county population. In columns 2–3, the dependent variables are log income per capita and log hourly earnings for men aged 25–64; columns 4–5 (resp., 6–7) report the same outcomes for U.S.-born men (resp., immigrant men). Column 1 includes county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. Columns 2–7 include census region fixed effects and the year in which a county was first connected to the railroad. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and, in Panels B and C, are further adjusted using the block bootstrap procedure described in Section 3.1. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Appendix

Table of Contents

A Additional Table and Figures

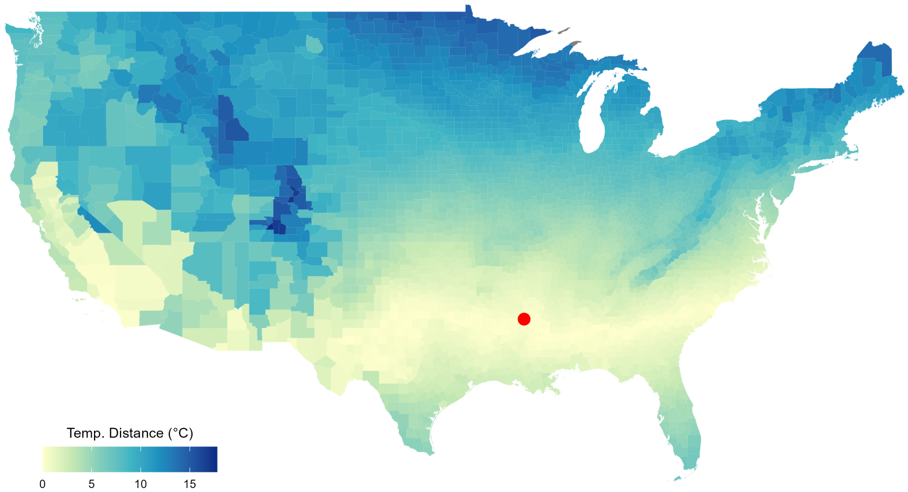
Figure A.1. U.S. County Temperature (1960–2000)



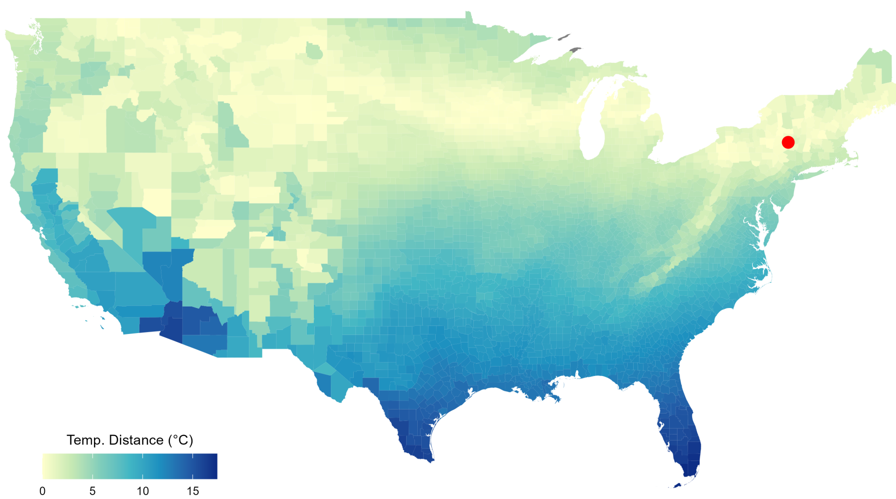
Notes: The figure displays mean annual temperature in each U.S. county in degrees Celsius averaged over the period 1960–2000 using TerraClimate data (Abatzoglou et al., 2018).

Figure A.2. Temperature Distances Between U.S. Counties

(A) Relative to the Delta Region



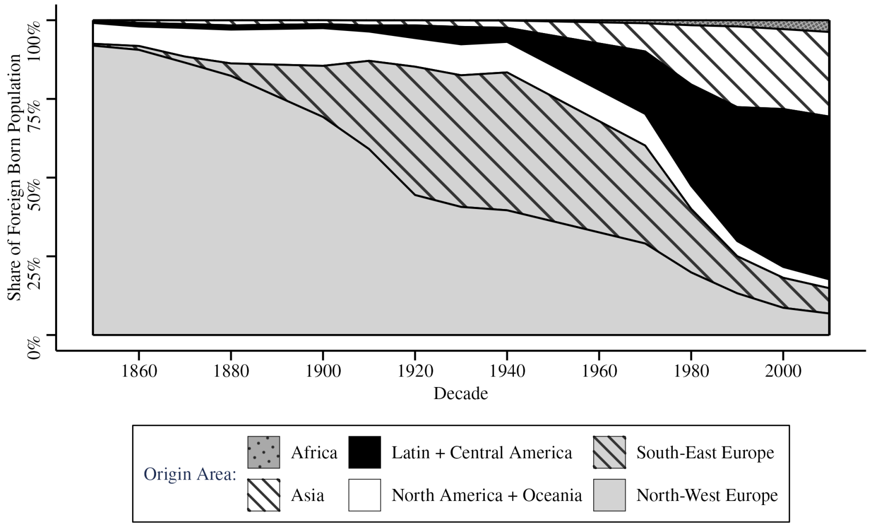
(B) Relative to the U.S. Northeast



Notes: The figure plots the absolute temperature difference between each U.S. county and the U.S. county indicated by the red dot. Temperature is defined as the annual mean in degrees Celsius averaged over the period 1960–2000 using TerraClimate data (Abatzoglou et al., 2018). For reference, temperature distance between the Delta region and New York City is 12° C.

Figure A.3. U.S. Foreign Born Population, 1850 - 2010

(A) U.S. Foreign Born Population by Origin

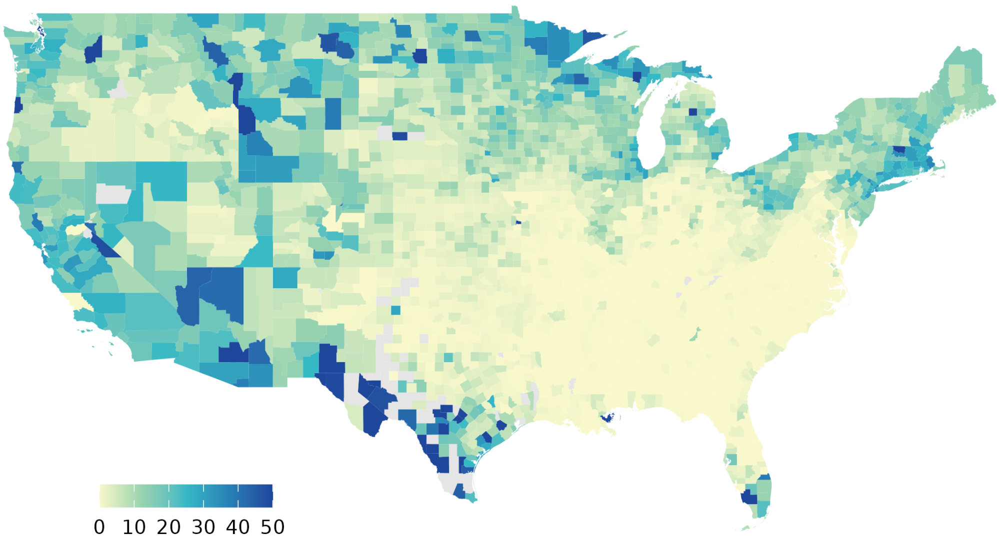


(B) U.S. Foreign Born Population Share

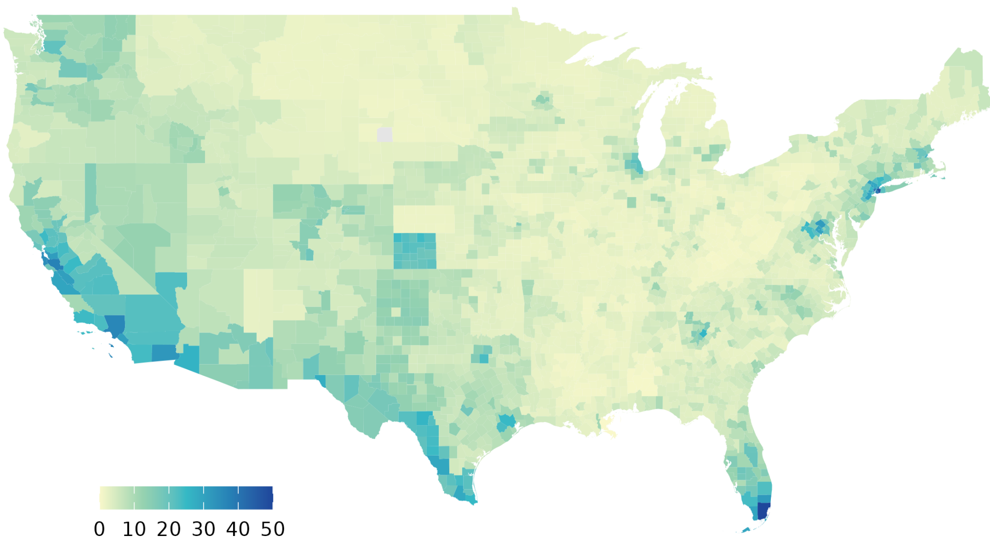
Notes: Panel A displays the U.S. foreign-born population share by origin region from 1850 to 2010. Panel B displays the immigrant population share from 1850-2010. Data for 1850 to 1940 come from full count U.S. Censuses (Ruggles et al., 2021). Data for 1970–2010 come from IPUMS micro-samples and from the American Community Survey (ACS).

Figure A.4. Immigrant Population Share

(A) 1920



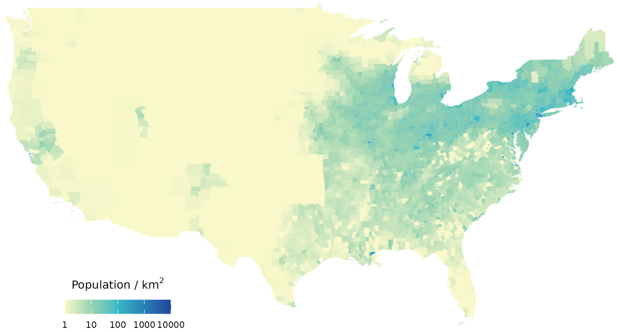
(B) 2010



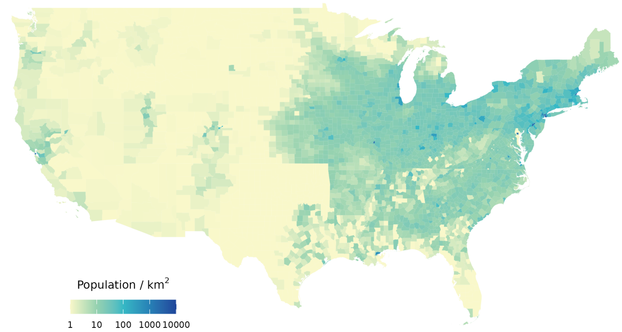
Notes: The figure displays the foreign-born population share (%) across U.S. counties in 1920 (Panel A) and 2010 (Panel B), using data from the full count U.S. Census and the American Community Survey (ACS), respectively.

Figure A.5. Distribution of U.S. Population, by Decade

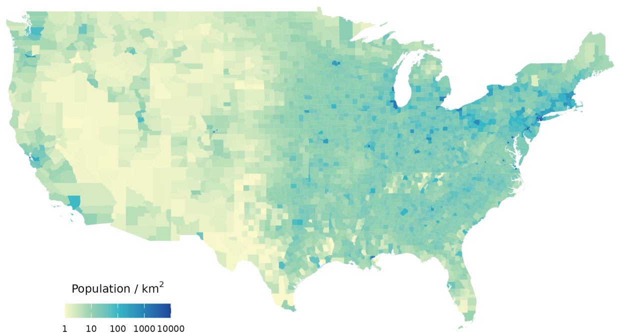
(A) 1860



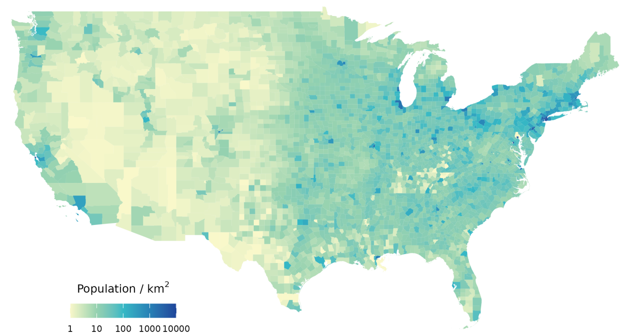
(B) 1880



(C) 1920



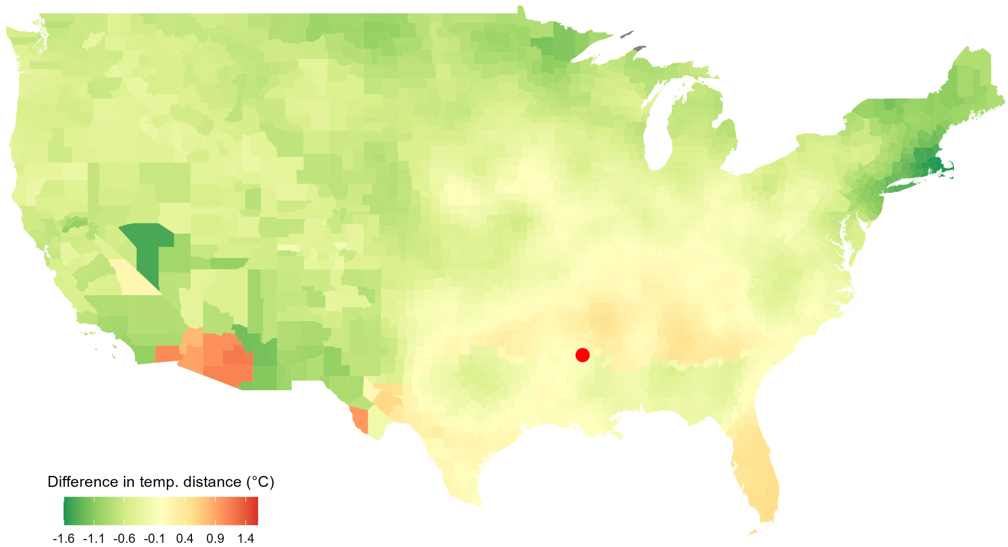
(D) 1940



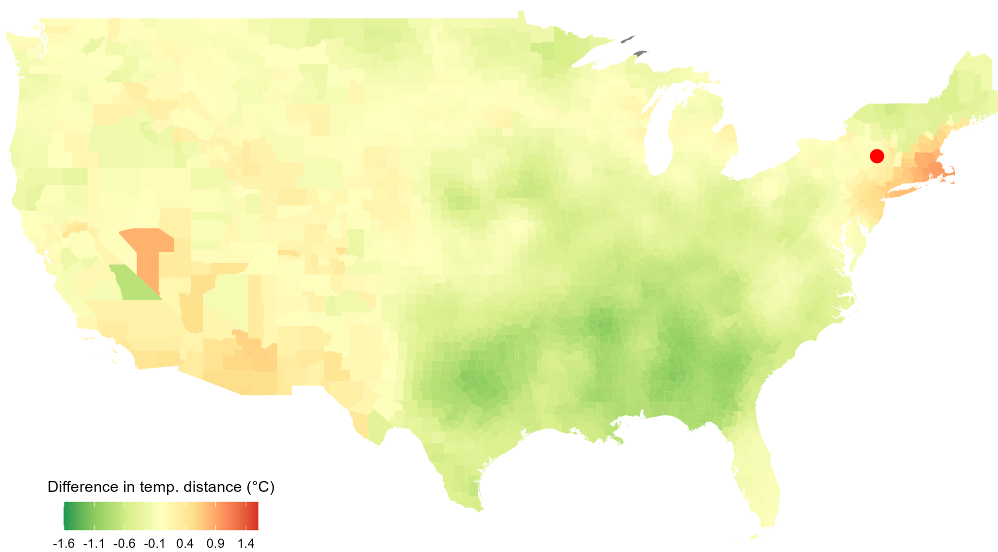
Notes: Each map plots county population density (individuals per square km) in a given decade, using data from full count U.S. Censuses (Ruggles et al., 2021).

Figure A.6. Historical Change in Temperature Distance Between U.S. Counties

(A) Relative to the Delta Region



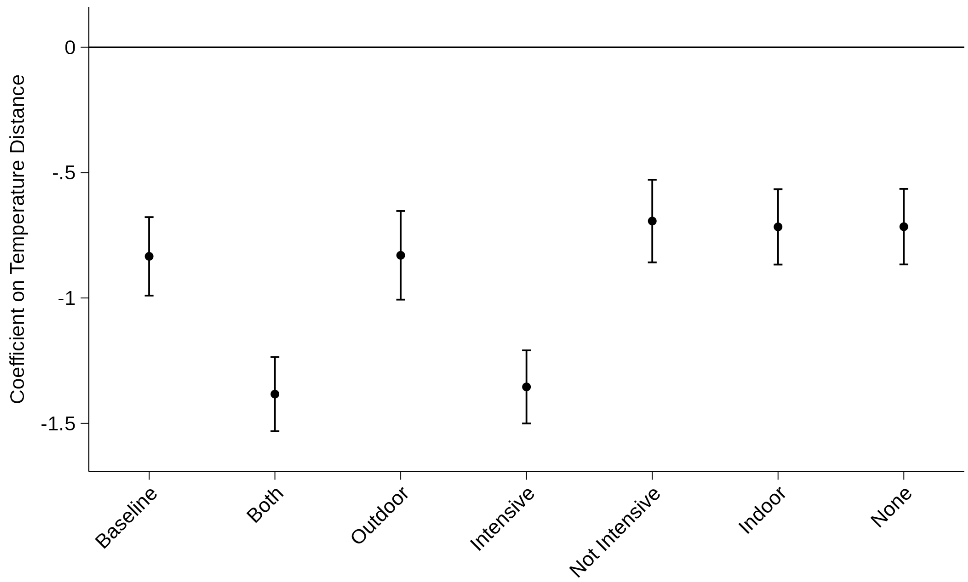
(B) Relative to the U.S. Northeast



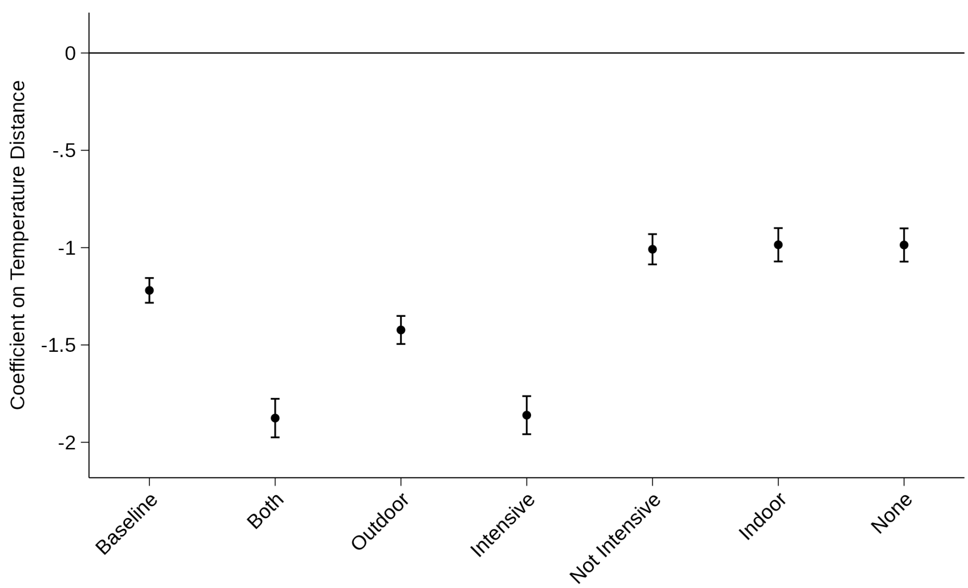
Notes: The figure plots the change in the absolute difference in temperature between each U.S. county and the U.S. county indicated by the red dot. Temperature is defined as the annual mean in degrees Celsius. These averages are taken over the period 1901–1930 and 1991–2020 using CRU data. To interpret the figure, note that in Panel A, the *distance* in temperature between the Delta region and New York City decreased by 1° C: (from 7.2° C to 6.2° C); New York City today remains colder on average than the Delta region but the difference is smaller than it was 100 years ago. As a result New York City appears bright green on the map in Panel A, representing climate convergence.

Figure A.7. Climate Distance and Migration: Heterogeneity by Occupation

(A) International Migration

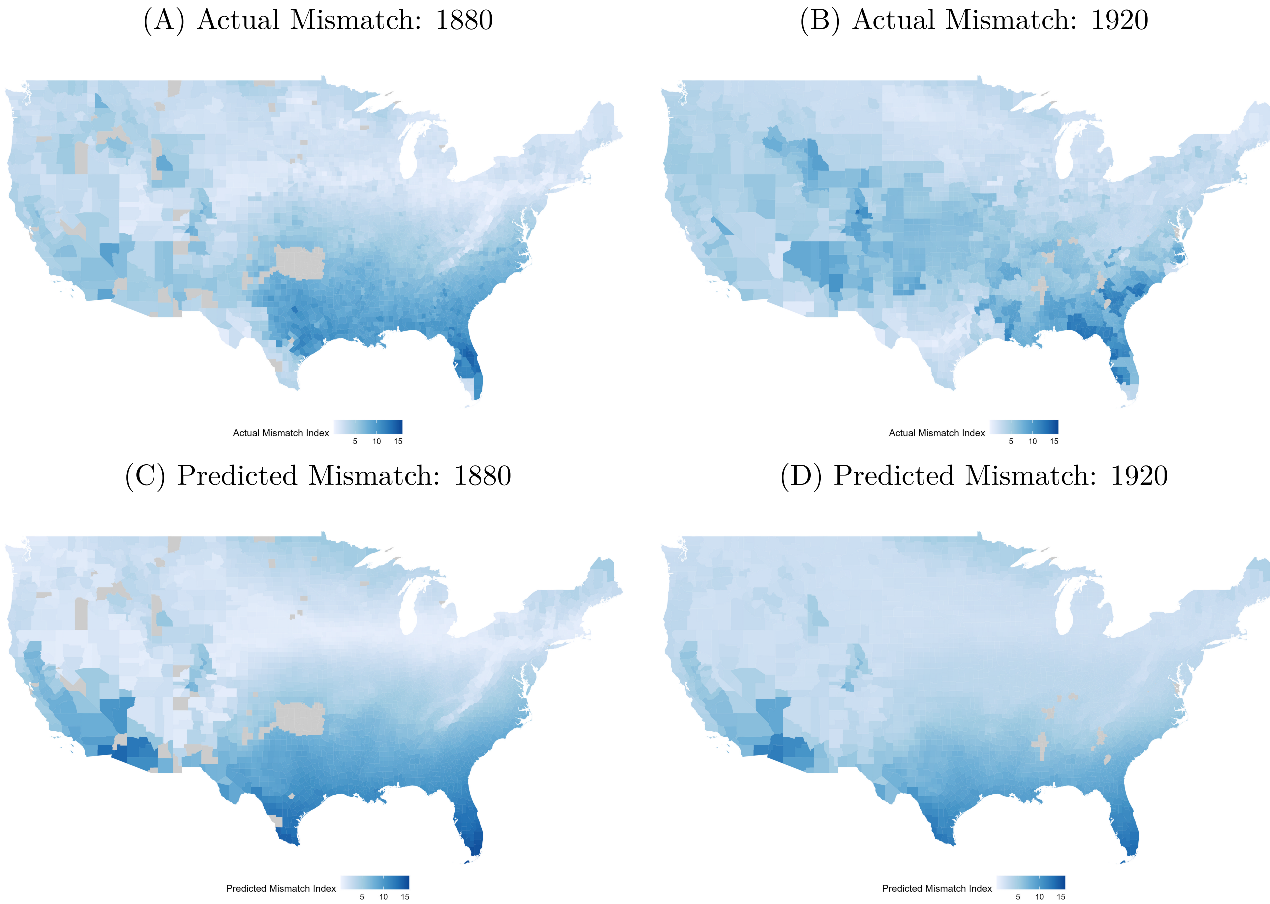


(B) U.S. Internal Migration



Notes: The figure plots the coefficient, with corresponding 95% confidence intervals, on the absolute difference in temperature between (A) international country of birth and U.S. destination county, and (B) U.S. origin and destination counties for internal movers. The first dot in Panel A (resp., Panel B) shows the baseline results of column 1 (resp., column 3) of Table 1. From the second dot onward, the number of migrants refers to migrant men in the sector reported on the x-axis (defined according to 3-digit OCC1950 code from IPUMS). For internal migrants, occupation is measured in the baseline decade (before the move). Temperature distances are standardized within the relevant sample to have zero mean and standard deviation equal to one. Standard errors are clustered at the country (resp., U.S. state) of origin by U.S. state of destination level.

Figure A.8. Temperature Mismatch Maps: Actual and Predicted



Notes: The figure plots actual (Panels A and B) and predicted (Panels C and D) temperature mismatch across U.S. counties in 1880 and 1920. Actual mismatch is defined in equation (4), while predicted mismatch is constructed from the fitted values of equation (6), as explained in Section 3.1. Higher values correspond to a larger temperature distance between origin countries and destination counties. Modern day Oklahoma is missing from the 1880 maps, as it was part of Indian Territory.

Table A.1. Top-10 Immigrant Sending Countries, 1920 and 2010

Year	1920		2010	
Rank	Country	Immigrant Share (%)	Country	Immigrant Share (%)
1	Germany	11.77	Mexico	28.93
2	Italy	11.49	China	5.213
3	USSR	11.46	Philippines	4.304
4	Canada	9.017	India	3.250
5	United Kingdom	8.231	Germany	2.949
6	Poland	7.652	Vietnam	2.931
7	Ireland	7.391	El Salvador	2.807
8	Austria	4.777	USSR	2.698
9	Sweden	4.461	Cuba	2.576
10	Mexico	3.525	Canada	2.331

Notes: The table reports the share of immigrants from each sending country, relative to all foreign-born individuals living in the U.S. in 1920 and 2010. Data come from the full-count U.S. Census and the ACS 5-year sample for 1920 and 2010, respectively. Country definitions are harmonized to 1990 as in Burchardi et al. (2019).

Table A.2. Spatial Correlation of Climate: Flexibly Predicting Temperature Distance

Dep. Variable:	Temperature Distance					
	International Migration			U.S. Internal Migration		
	(1)	(2)	(3)	(4)	(5)	(6)
Distance (in 100km)	-37.55	-37.55	-37.55	3,649***	3,650***	3,650***
	(47.97)	(49.84)	(50.27)	(228.2)	(180.4)	(180.2)
Distance (in 100km)^2		473.9***	473.9***		-3,505***	-3,505***
		(51.09)	(50.90)		(199.6)	(197.0)
Distance (in 100km)^3			-264.1***			530.6***
			(54.49)			(170.3)
Observations	553,046	553,046	553,046	9,650,342	9,650,342	9,650,342
R-squared	0.000	0.013	0.016	0.102	0.197	0.199

Notes: The sample includes U.S. county-country pairs in columns 1–3 and county pairs in the contiguous U.S. in columns 4–6. Temperature Distance is the absolute difference in mean annual temperature between the origin and the destination, measured in degrees Celsius. Distance is the physical distance between the origin and the destination, expressed in 100 km. In columns 1–3 (resp., columns 4–6), Origin o refers to the origin country (resp., the origin U.S. county). Destination d refers to a U.S. county in all columns. Standard errors, reported in parentheses, are clustered at the country (resp., U.S. state) of origin by U.S. state of destination level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.3. Climate and Migration: Controlling for Lagged Migrants

Dep. Variable:	Number of Migrants			
	Intl-US 1900–1920	Intl-US 1970–2010	Internal 1860–1940	Internal 2011–2019
	(1)	(2)	(3)	(4)
Temperature Distance	-0.153***	-0.016	-0.307***	-0.231***
	(0.019)	(0.010)	(0.018)	(0.040)
Past Migrant Stock	1.179***	0.056*	1.424**	6.729***
	(0.260)	(0.033)	(0.577)	(2.065)
Observations	1,442,644	2,475,100	28,823,701	25,088,936
Pseudo R-squared	0.916	0.894	0.731	0.778
Mean Temp. Dist.	9.597	9.621	4.956	4.989
SD Temp. Dist.	5.747	5.685	3.600	3.632
Origin×Decade/Year FE	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes

Notes: The table replicates the baseline specification from Table 1, additionally controlling for pre-existing migrant stocks (expressed per 100,000 individuals). In columns 1–2, Past Migrant Stock is the immigrant stock in the destination county in the previous decade; for 1970, we use the 1940 stock. In column 3, Past Migrant Stock is measured as the cumulated lagged flows from the linked sample up to the previous decade. In column 4, Past Migrant Stock is measured as the average origin–destination flows from the linked sample over 1910–1940. Lagged migration controls cannot be constructed for the first observation in each historical sample (1880 for international migration and 1860 for internal migration), which is therefore omitted from the corresponding regressions. All regressions include precipitation and geographic distances, origin-by-decade (or, origin-by-year) fixed effects and destination-by-decade (or, destination-by-year) fixed effects. Standard errors, reported in parentheses, are clustered at the country-of-origin by U.S. state-of-destination level in columns 1–2 and at the U.S. state-of-origin by U.S. state-of-destination level in columns 3–4. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.4. Climate Distance and Migration: Evidence from the U.S. Frontier

Dep. Variable:	Number of Migrants							
	International Migration			U.S. Internal Migration				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature Distance	-0.227***	-0.254***	-0.204***	-0.199***	-0.268***	-0.144***	-0.177***	-0.175***
	(0.026)	(0.030)	(0.031)	(0.022)	(0.032)	(0.028)	(0.031)	(0.023)
Observations	62,274	36,000	26,274	612,554	150,711	147,853	165,332	148,658
Pseudo R-squared	0.904	0.890	0.917	0.472	0.478	0.456	0.488	0.441
Mean Temp. Dist.	10.79	11.35	10.01	4.401	4.670	4.083	4.347	4.506
SD Temp. Dist.	6.726	6.773	6.582	3.258	3.432	3.062	3.150	3.354
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period	1880–1900	1880	1900	1850–1900	1850–1860	1860–1870	1870–1880	1880–1900

Notes: The table replicates the baseline specifications for historical international and internal migration (Table 1, columns 1 and 3) restricting attention to destination counties on the U.S. frontier (as defined in Bazzi et al., 2020) in each decade. Columns 1–3 report estimates for international immigration; columns 4–8 report estimates for internal migration, restricting attention to origin counties that were never on the frontier. Column 1 pools 1880 and 1900; columns 2–3 report each decade separately; column 4 pools all decades from 1850 to 1900; columns 5–8 report each decade separately. The dependent variable is the number of migrants from the origin to the destination. Temperature Distance is the absolute difference in temperature between the origin and the destination. All regressions include the absolute difference in precipitation, physical distance, origin by decade fixed effects, and destination by decade fixed effects. Standard errors, reported in parentheses, are clustered at the country-of-origin (resp., U.S. state of origin) by U.S. state-of-destination level in columns 1–3 (resp., in columns 4–8). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.5. Climate Distance and Migration: The Role of Air Conditioning

Dep. Variable:	Number of Migrants							
	Historical: 1880–1920			Modern: 1970–2010				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Annual Temp. Dist.	-0.175***			-0.019*				
	(0.021)			(0.011)				
January Temp. Dist.		-0.053***	-0.042**		-0.013**	-0.008	-0.016	-0.008
		(0.014)	(0.017)		(0.006)	(0.006)	(0.010)	(0.006)
July Temp. Dist.		-0.030*	0.001		-0.042**	-0.080***	-0.074***	-0.081***
		(0.017)	(0.020)		(0.019)	(0.022)	(0.022)	(0.022)
South × Jan. Temp. Dist.			-0.062***			-0.026***	-0.052***	-0.019**
			(0.023)			(0.009)	(0.017)	(0.008)
South × July Temp. Dist.			-0.052*			0.059***	-0.011	0.046**
			(0.028)			(0.020)	(0.028)	(0.019)
Observations	1,939,195	1,939,195	1,939,195	2,475,100	2,475,100	2,475,100	1,956,231	2,470,464
Pseudo R-squared	0.916	0.913	0.914	0.893	0.892	0.893	0.918	0.895
Temp. Dist. Mean	9.622			9.621				
Jan. Temp. Dist. Mean		17.93	17.93		18.08	18.08	18.14	18.05
July Temp. Dist. Mean		5.258	5.258		5.187	5.187	5.162	5.190
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes U.S. county-country pairs for each decade over the 1880–1920 period (except for 1890) in columns 1–3, and over the 1970–2010 period in columns 4 to 6. The definition of countries of origin is fixed to 1990 boundaries and harmonized following the procedure in Burchardi et al. (2019). Columns 1 and 4 replicate the baseline specification of Table 1, columns 1 and 2. Columns 2 and 5 replace the absolute difference in mean annual temperature with temperature measured in January and July. Columns 3 and 6 interact the absolute difference in January and July temperature with an indicator equal to one if the destination county's average July temperature is above the national median. Columns 7 and 8 replicate column 6 separately for farmers and non-farmers, respectively. All regressions include the absolute difference in precipitation, geographic distance, country of origin by decade fixed effects, and destination county by decade fixed effects. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.6. Predicting Migration Flows Using Push-Pull Shocks

Dep. Variable:	Number of Migrants	
		Placebo
	(1)	(2)
$\log(M_{ot}) \times R_{dt}$	0.296***	
	(0.037)	
$\log(M_{ot+1}) \times R_{dt+1}$		0.018
		(0.017)
Temperature distance	-0.142***	-0.182***
	(0.030)	(0.015)
Precipitation distance	-0.006	0.0003
	(0.005)	(0.003)
Distance	-0.001	-0.087***
	(0.036)	(0.008)
Observations	281,819	245,852
Pseudo R-squared	0.899	0.962
Destination×Decade FE	Yes	Yes
Origin×Decade FE	Yes	Yes

Notes: The sample includes U.S. county-country pairs for each decade over the 1880–1920 (except for 1890). The definition of countries of origin is fixed to 1990 boundaries and harmonized following the procedure in Burchardi et al. (2019). Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade (except for 1880, when information on year of arrival is not available). $\log(M_{ot}) \times R_{dt}$ is the interaction between the log number of migrants arriving from origin o between decade $t-1$ and t and an indicator for whether a county d was first connected to the rail during the same period. $\log(M_{ot+1}) \times R_{dt+1}$ is the same interaction measured in the decade spanning t and $t+1$. Temperature (resp., Precipitation) Distance is the absolute difference in temperature (resp., precipitation) between the origin and destination. Distance is geographic distance, expressed in 100 km, between the origin and the destination. Origin o refers to the origin country. Destination d refers to a U.S. county. All regressions also control for precipitation and geographic distances, origin by decade fixed effects and destination by decade fixed effects. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.7. Climate Mismatch and Crop Yields

Dep. Variable:	Corn			Wheat		
	Log Acres	Log Output	Log Yield	Log Acres	Log Output	Log Yield
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Temperature Mismatch	-0.023***	-0.035***	-0.012***	0.153***	0.147***	-0.005
	(0.009)	(0.009)	(0.003)	(0.042)	(0.042)	(0.005)
R-squared	0.923	0.918	0.879	0.856	0.864	0.729
Panel B: 2SLS						
Temperature Mismatch	-0.170***	-0.125**	0.045***	0.181**	0.180**	-0.001
	(0.047)	(0.051)	(0.012)	(0.083)	(0.086)	(0.013)
R-squared	0.923	0.918	0.879	0.856	0.864	0.729
KP F-stat	609.67	609.67	609.67	598.67	598.61	598.67
Panel C: First Stage						
Predicted Mismatch	0.649***	0.649***	0.649***	0.648***	0.648***	0.648***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
R-squared	0.908	0.908	0.908	0.906	0.906	0.906
Observations	8,463	8,463	8,463	8,273	8,279	8,273
1880 Dep. Var. Mean	27,902	805,425	24.52	16,626	217,309	10.85
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 and at least 100 farms in each decade from 1880 to 1920. The year 1890 is excluded, because it is not possible to construct the measure of climate mismatch. Panels A and B report OLS and 2SLS estimates, respectively. Panel C reports the first stage. Temperature Mismatch is the weighted average of the absolute temperature difference (in degrees Celsius) between the county and the country of origin of the migrants, with weights equal to the share of migrants from each origin arriving in the county in each decade, relative to all migrants arriving in the county in that decade. See Section 3.1 and equation (4). Predicted Mismatch (Panel C) is the corresponding measure of predicted average temperature distance, constructed as described in Section 3.1. The dependent variables are the log corn acres (column 1), the log corn output (column 2), the log corn yield (column 3), the log wheat acres (column 4), the log wheat output (column 5), and the log wheat yield (column 6). All regressions include county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and, in Panels B and C, are further adjusted using the block bootstrap procedure described in Section 3.1. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

B Data Appendix

Table B.1. Summary of Data Sources

Source	Description	Area	Spatial Resolution	Period	Time Resolution
Climate					
TerraClimate (Abatzoglou et al., 2018)	Mean annual temperature (°C) and precipitation (mm) from monthly gridded dataset; (i) For U.S. counties, county-level values are population-weighted averages of gridcells within county borders. For countries outside the U.S., values are population-weighted averages over gridcells within national borders; (ii) Robustness using averages within a 25 km radius around national capital cities.	Global	4.6 km × 4.6 km (at the equator)	1960-2000	Averaged over period
Climatic Research Unit (CRU) (Harris et al., 2020)	Mean temperature (°C) and precipitation (mm/month) derived from a monthly gridded observational product (no underlying climate model).	Global	55.7 km × 55.7 km (at the equator)	1901-1930; 1991-2020	Averaged over period
Immigration					
U.S. Census of Population (Ruggles et al., 2021)	International immigration to the U.S.; individual characteristics of the U.S. population. Full-count data for 1880-1920; 1% Form 1 Metro sample for 1970; 5% State sample for 1980 and 1990; 5% sample of 2000 Census; and ACS 5-Year sample of the 2010 Census.	U.S.	Individual	1880-1920 (excluding 1890); 1970-2010	Decade
European Immigration to the U.S. (Willcox, 1929)	Number of immigrants from each European country entering the U.S. in a given decade.	U.S.	Country	1870-1920	Decade
Census Tree Project (Price et al., 2021; Buckles et al., 2025)	Linked samples of individuals between censuses. Combined with the Census Place Project (Berkes et al., 2023) to construct historical internal migration matrices. Restricted to individuals 15 or older in the baseline decade. Because the 1890 Census is unavailable, we link individuals observed in 1880 to 1900.	U.S.	Individual	1850-1940	Decade
Census Place Project (Berkes et al., 2023)	Individual harmonized addresses. Combined with the Census Tree Project to construct historical internal migration matrices. Because the 1890 Census is unavailable, we link individuals observed in 1880 directly to 1900.	U.S.	Individual	1850-1940	Decade

IRS Statistics of Income Division (https://www.irs.gov/statistics/soi-tax-stats-migration-data)	Modern internal migration matrices. Constructed from IRS Change-of-Address Tables for 2011–2022, which draw upon year-to-year address changes reported on individual income tax returns.	U.S.	County-to-county	2011-2019	Calendar year
Economic Indicators					
Census of Agriculture (Haines and Inter-university Consortium for Political and Social Research, 2000)	County-level agricultural characteristics. Restricted to counties with at least 100 farms in each census year.	U.S.	County	1880-1920	Decade
Census of Manufactures (Haines and Inter-university Consortium for Political and Social Research, 2010; Haines et al., 2018)	County-level manufacturing characteristics.	U.S.	County	1880-1920 (excluding 1910)	Decade
U.S. Census of Population (Ruggles et al., 2021)	Individual income and hours worked. Restricted to men 25–64.	U.S.	Individual	1940	Decade
Miscellaneous					
Geographic Crosswalks (Eckert et al., 2020; Berkes et al., 2023)	Crosswalks between historical and modern county boundaries.	U.S.	County-to-county	1850-2000	Decade
Replication files from Burchardi et al. (2019)	Crosswalks between historical and modern country boundaries; crosswalks between U.S. county-groups and U.S. counties.	U.S., Global	Country-to-country; County-to-county	1880-2010	Decade
U.S. Frontier (Bazzi et al., 2020)	Total frontier experience; dummy for being on the frontier in a decade.	U.S.	County	1850-1890	Decade
Connection to Railroads (Atack and Margo, 2011)	Railroads connection data.	U.S.	County	1850-1930	Decade
Replication package from Donaldson and Hornbeck (2016)	Market access; Lowest-cost county-to-county freight routes resulting in time-varying transportation costs.	U.S.	County	1850-1940	Decade

Geographic Features	Elevation (Danielson and Gesch, 2011); Ruggedness, calculated as the spatial standard deviation of elevation values within a 4.5 km square neighborhood; Distance to coast (Carroll et al., 2009). Countries and U.S. counties values are processed in Google Earth Engine and constructed as spatial averages of gridded values within U.S. county borders and national borders, respectively.	Global	11.1 $\times$ 11.1 km (at the equator)	1960-2000	Average
Soil Characteristics	Bulk density (Hengl, 2018a), organic matter (Hengl and Wheeler, 2018), soil pH (Hengl, 2018b), water content (Hengl and Gupta, 2019).	Global	0.25 $\times$ 0.25 km (at the equator)	1960-2000	Average

C Robustness Checks — Evidence of Climate Matching

C.1 Assessing Sensitivity to Controls and Alternate Specifications

Gradual inclusion of controls. In Tables C.1–C.4, we replicate the baseline results from Table 1, introducing controls and fixed effects sequentially. Tables C.1 and C.2 focus on historical and modern immigration to the U.S., respectively. Column 1 estimates a parsimonious specification that includes only temperature and precipitation distance. Column 2 adds geographic distance (measured analogously to climate distance, as the distance between the county centroid and the origin-country capital). Column 3 further includes year dummies. In all cases, the coefficient on temperature distance remains negative and statistically significant. Columns 4 and 5 sequentially introduce continent-of-origin and U.S. state-of-destination fixed effects, and then country-of-origin and U.S. county-of-destination fixed effects. Finally, column 6 reports the baseline specification from Table 1, columns 1 and 2, which interacts year dummies with country and county fixed effects. Across the two tables, the number of observations declines as additional controls and fixed effects are included, reflecting the increasingly demanding sources of identifying variation.

Tables C.3 and C.4 present the internal migration results for both the historical linked sample (1850–1940) and the modern IRS data (2011–2019). Column 1 reports a very parsimonious specification; column 2 adds geographic distance, which reduces the coefficient on temperature distance from -0.448 to -0.325 (historical migration) and from -0.393 to -0.243 (modern migration) but leaves it highly statistically significant. Column 3 introduces county-of-origin, county-of-destination, and period fixed effects (decade for the historical sample; calendar year for the modern sample). Column 4 interacts period dummies with state-of-origin and state-of-destination fixed effects. Column 5 presents the baseline specification, interacting county-of-origin and county-of-destination fixed effects with period fixed effects. For the modern IRS data, our baseline excludes moves after 2019 to avoid potential distortions from the COVID-19 pandemic, but column 6 of Table C.4 shows that including moves from 2019–2020 and 2020–2021 yields virtually identical results. As with the international analysis—and even more so here—the number of observations falls sharply as fixed effects become more demanding, reflecting the increasingly stringent identifying variation.

International immigration: sample and climate measurement. Table C.5 combines the historical and modern international immigration results. Columns 1–4 report estimates for the Age of Mass Migration. Column 1 presents the baseline historical specification. In column 2, we restrict the sample to European immigrants—who accounted for over 85% of arrivals during this period (Abramitzky and Boustan, 2017). Despite the resulting reduction in sample size and climate variation, the estimates remain closely aligned with the baseline. Column 3 excludes immigrants arriving in 1880, addressing concerns that the absence of year-of-arrival information in that census year implies a stock rather than flow measure of migration; the results are unchanged.[37] Column 4 replaces population-weighted origin-country climate with capital-city temperatures, yielding estimates similar in magnitude and statistical significance to the baseline.

Columns 5–7 turn to modern international immigration. Column 5 reports the baseline modern specification. Column 6 excludes European immigrants, showing that the results are not driven by relatively small sending regions in the contemporary period. Column 7 again measures origin-country climate using capital-city temperatures. While the point estimate remains negative, it is estimated less precisely than in the historical sample and is no longer statistically significant—consistent with greater within-country climatic heterogeneity in the modern period.

Clustering structure. In the baseline specifications, standard errors are clustered at the country of origin by U.S. state of destination level (for international migration) or the U.S. state of origin by U.S. state of destination level (for internal migration). Table C.6 shows that the coefficients remain statistically significant when we instead cluster at the country of origin by U.S. Census region level (international immigration) or the U.S. Census region of origin by U.S. Census region of destination level (internal migration).

[37] In 1880, year of arrival is not reported, so immigrant counts reflect stocks rather than flows; excluding 1880 therefore alleviates concerns about mixing stock and flow measures or allowing the results to be driven by this difference in measurement.

Table C.1. International Immigration (1880-1920): Gradual Inclusion of Controls

Dep. variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.349***	-0.384***	-0.384***	-0.245***	-0.163***	-0.175***
	(0.021)	(0.023)	(0.023)	(0.023)	(0.020)	(0.021)
Precipitation Distance	-0.029***	-0.026***	-0.026***	-0.013***	0.001	-0.002
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Distance (100 km)		-0.028***	-0.028***	-0.048***	-0.064***	-0.064***
		(0.002)	(0.002)	(0.008)	(0.014)	(0.013)
Observations	2,202,216	2,202,216	2,202,216	2,190,415	2,041,380	1,939,195
Pseudo R-squared	0.210	0.279	0.282	0.482	0.850	0.916
Mean Temp. Dist.	9.775	9.775	9.775	9.786	9.677	9.622
SD Temp. Dist.	5.694	5.694	5.694	5.698	5.742	5.754
Mean Precip. Dist.	55.81	55.81	55.81	55.85	52.54	52.53
SD Precip. Dist.	45.89	45.89	45.89	45.96	41.77	42.07
Decade FE			Yes	Yes	Yes	
Origin Continent FE				Yes		
Destination State FE				Yes		
Origin FE					Yes	
Destination FE					Yes	
Origin×Decade FE						Yes
Destination×Decade FE						Yes

Notes: This table replicates the baseline results from Table 1 for historical immigration to the U.S., introducing controls and fixed effects sequentially across columns. The sample includes U.S. county-country pairs for each decade over the 1880–1920 (except for 1890). Column 1 includes only Temperature Distance and Precipitation Distance, defined as the absolute temperature and precipitation differences between the country of origin and the county of destination, measured in degrees Celsius and millimeters respectively. Column 2 adds Distance, defined as the physical distance between the origin country and the destination county, expressed in 100 km. Column 3 includes census year fixed effects. Column 4 adds continent-of-origin and U.S. state-of-destination fixed effects, and column 5 adds country-of-origin and U.S. county-of-destination fixed effects. Column 6 reports the baseline specification from Table 1, column 1, which interacts census year dummies with country-of-origin and county-of-destination fixed effects. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination level. Significance levels: *** p< 0.01, ** p< 0.05, * p< 0.10.

Table C.2. International Immigration (1970-2010): Gradual Inclusion of Controls

Dep. variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.112***	-0.128***	-0.128***	-0.128***	-0.020*	-0.019*
	(0.025)	(0.030)	(0.030)	(0.026)	(0.010)	(0.011)
Precipitation Distance	-0.003	-0.001	-0.001	-0.005**	-0.008***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Distance (100 km)		-0.022***	-0.022***	-0.034***	-0.051***	-0.051***
		(0.006)	(0.006)	(0.009)	(0.004)	(0.004)
Observations	2,765,230	2,765,230	2,765,230	2,750,410	2,749,695	2,475,100
Pseudo R-squared	0.035	0.090	0.113	0.367	0.871	0.893
Mean Temp. Dist.	9.791	9.791	9.791	9.798	9.770	9.621
SD Temp. Dist.	5.697	5.697	5.697	5.703	5.697	5.685
Mean Precip. Dist.	55.67	55.67	55.67	55.71	55.14	53.96
SD Precip. Dist.	45.71	45.71	45.71	45.78	45.25	43.90
Decade FE			Yes	Yes	Yes	
Origin Continent FE				Yes		
Destination State FE				Yes		
Origin FE					Yes	
Destination FE					Yes	
Origin×Decade FE						Yes
Destination×Decade FE						Yes

Notes: This table replicates the baseline results from Table 1 for modern immigration to the U.S., introducing controls and fixed effects sequentially across columns. The sample includes U.S. county-country pairs for each decade over the 1970–2010 period. Column 1 includes only Temperature Distance and Precipitation Distance, defined as the absolute temperature and precipitation differences between the country of origin and the county of destination, measured in degrees Celsius and millimeters respectively. Column 2 adds Distance, defined as the physical distance between the origin country and the destination county, expressed in 100 km. Column 3 includes census year fixed effects. Column 4 adds continent-of-origin and U.S. state-of-destination fixed effects, and column 5 adds country-of-origin and U.S. county-of-destination fixed effects. Column 6 reports the baseline specification from Table 1, column 2, which interacts census year dummies with country-of-origin and county-of-destination fixed effects. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination level. Significance levels: *** p< 0.01, ** p< 0.05, * p< 0.10.

Table C.3. U.S. Internal Migration (1850-1940): Gradual Inclusion of Controls

Dep. Variable:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Temperature Distance	-0.448***	-0.325***	-0.318***	-0.316***	-0.315***
	(0.037)	(0.023)	(0.019)	(0.019)	(0.019)
Precipitation Distance	-0.057***	-0.027***	-0.018***	-0.017***	-0.017***
	(0.006)	(0.004)	(0.002)	(0.002)	(0.002)
Distance (100 km)		-0.196***	-0.193***	-0.193***	-0.194***
		(0.036)	(0.014)	(0.014)	(0.014)
Observations	38,428,111	38,428,111	38,031,643	36,367,197	31,092,699
Pseudo R-squared	0.213	0.268	0.700	0.712	0.728
Mean Temp. Dist.	5.062	5.062	5.052	5.029	4.923
SD Temp. Dist.	3.667	3.667	3.660	3.649	3.578
Mean Precip. Dist.	32.85	32.85	32.81	32.49	31.01
SD Precip. Dist.	24.43	24.43	24.42	24.29	23.76
Decade FE			Yes	Yes	
Origin FE			Yes	Yes	
Destination FE			Yes	Yes	
Origin State×Decade FE				Yes	
Destination State×Decade FE				Yes	
Origin×Decade FE					Yes
Destination×Decade FE					Yes

Notes: The table replicates the baseline results from Table 1 for historical internal-migration, introducing controls and fixed effects sequentially across columns. The sample includes county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940. Number of Migrants is the number of people aged 15+ in the baseline year who moved from the origin to the destination county over the subsequent decade. Column 1 includes only Temperature Distance and Precipitation Distance, defined as the absolute temperature and precipitation differences between the origin and the destination county, measured in degrees Celsius and millimeters respectively. Column 2 adds Distance, defined as the physical distance between counties, expressed in 100 km. Column 3 introduces census year, county-of-origin, and county-of-destination fixed effects. Column 4 adds the interaction of census year dummies with U.S. state-of-origin and U.S. state-of-destination fixed effects. Column 5 reports the baseline specification from Table 1, column 3, which interacts census year dummies with county-of-origin and county-of-destination fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C.4. U.S. Internal Migration (2011-2019): Gradual Inclusion of Controls

Dep. Variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.404***	-0.223***	-0.237***	-0.237***	-0.237***	-0.235***
	(0.066)	(0.052)	(0.041)	(0.041)	(0.041)	(0.041)
Precipitation Distance	-0.059***	-0.023***	-0.030***	-0.030***	-0.030***	-0.030***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Distance (100 km)		-0.272***	-0.203***	-0.204***	-0.204***	-0.204***
		(0.056)	(0.022)	(0.022)	(0.022)	(0.022)
Observations	38,601,368	38,601,368	31,823,048	31,823,048	25,088,936	31,045,614
Pseudo R-squared	0.147	0.206	0.783	0.783	0.776	0.777
Mean Temp. Dist.	5.052	5.052	5.016	5.016	4.989	4.987
SD Temp. Dist.	3.662	3.662	3.644	3.644	3.632	3.632
Mean Precip. Dist.	32.87	32.87	32.00	32.00	31.06	31.01
SD Precip. Dist.	24.43	24.43	24.24	24.24	24.01	23.99
Year FE			Yes	Yes		
Origin FE			Yes	Yes		
Destination FE			Yes	Yes		
Origin State×Year FE				Yes		
Destination State×Year FE				Yes		
Origin×Year FE					Yes	Yes
Destination×Year FE					Yes	Yes

Notes: The table replicates the baseline results from Table 1 for modern internal-migration, introducing controls and fixed effects sequentially across columns. The sample includes county-pairs in the contiguous U.S. for each calendar year from 2011–2012 to 2018–2019. Number of Migrants is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Column 1 includes only Temperature Distance and Precipitation Distance, defined as the absolute temperature and precipitation differences between the origin and the destination county, measured in degrees Celsius and millimeters respectively. Column 2 adds Distance, defined as the physical distance between counties, expressed in 100 km. Column 3 introduces year, county-of-origin, and county-of-destination fixed effects. Column 4 adds the interaction of year dummies with U.S. state-of-origin and U.S. state-of-destination fixed effects. Column 5 reports the baseline specification from Table 1, column 4, which interacts calendar year dummies with county-of-origin and county-of-destination fixed effects. Column 6 replicate column 5 including moves from 2019–2020 and 2020–2021. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p< 0.01$, ** $p< 0.05$, * $p< 0.10$.

Table C.5. International Immigration: Additional Robustness

Dep. Variable:	Number of Migrants						
	Historical: 1880–1920				Modern: 1970–2010		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Temperature Distance	-0.175***	-0.144***	-0.162***	-0.142***	-0.019*	-0.034*	-0.017
	(0.021)	(0.030)	(0.021)	(0.021)	(0.011)	(0.018)	(0.013)
Precipitation Distance	-0.002	-0.006	0.003	0.002	-0.008***	-0.010***	-0.007***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.002)	(0.002)	(0.002)
Distance (100 km)	-0.064***	-0.001	-0.060***	-0.083***	-0.051***	-0.054***	-0.051***
	(0.013)	(0.036)	(0.013)	(0.014)	(0.004)	(0.004)	(0.004)
Observations	1,939,195	376,782	1,450,765	1,914,683	2,475,100	2,003,462	2,445,142
Pseudo R-squared	0.916	0.906	0.915	0.915	0.893	0.903	0.893
Mean Temp. Dist.	9.622	5.070	9.605	9.462	9.621	10.68	9.446
SD Temp. Dist.	5.754	3.701	5.754	5.844	5.685	5.549	5.772
Mean Precip. Dist.	52.53	33.47	52.54	54.09	53.96	58.71	55.28
SD Precip. Dist.	42.07	29.41	42.13	47.91	43.90	45.37	48.71
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intl Mig Period	1880–1920	1880–1920	1900–1920	1880–1920	1970–2010	1970–2010	1970–2010
Origin Climate	Weighted Avg	Weighted Avg	Weighted Avg	Capital City	Weighted Avg	Weighted Avg	Capital City
Countries	Any	Europe only	Any	Any	Any	Non-Europe only	Any

Notes: The table reports robustness checks for historical (columns 1–4) and modern (columns 5–7) immigration to the U.S. with respect to alternative climate definitions and sample restrictions. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance refers to the physical distance between the origin country and the destination county (expressed in 100 km). Columns 1 replicates the historical baseline specification from Table 1, column 1. Column 2 restricts the sample to European origin countries only. Column 3 drops 1880 from the sample. Column 4 replaces the origin-country climate with capital-city TerraClimate temperature and precipitation. Column 5 replicates the modern baseline specification from Table 1, column 2. Column 6 drops European origin countries. Column 7 replaces the origin-country climate with capital-city TerraClimate temperature and precipitation. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p< 0.01$, ** $p< 0.05$, * $p< 0.10$.

Table C.6. Climate Distance and Migration: Clustering by U.S. Census Region

Dep. Variable:	Number of Migrants			
	Intl-US 1880–1920	Intl-US 1970–2010	Internal 1850–1940	Internal 2011–2019
	(1)	(2)	(3)	(4)
Temperature Distance	-0.175***	-0.019*	-0.315***	-0.237***
	(0.030)	(0.011)	(0.035)	(0.061)
Precipitation Distance	-0.002	-0.008***	-0.017***	-0.030***
	(0.005)	(0.002)	(0.004)	(0.009)
Distance (100 km)	-0.064***	-0.051***	-0.194***	-0.204***
	(0.021)	(0.005)	(0.026)	(0.038)
Observations	1,939,195	2,475,100	31,092,699	25,088,936
Pseudo R-squared	0.916	0.893	0.728	0.776
Mean Temp. Dist.	9.622	9.621	4.923	4.989
SD Temp. Dist.	5.754	5.685	3.578	3.632
Mean Precip. Dist.	52.53	53.96	31.01	31.06
SD Precip. Dist.	42.07	43.90	23.76	24.01
Origin×Decade/Year FE	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes

Notes: The table replicates all columns of Table 1. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. Census region of destination level (columns 1-2) and the U.S. Census region of origin by U.S. Census region of destination level (columns 3-4). Significance levels: *** p $<$ 0.01, ** p $<$ 0.05, * p $<$ 0.1.

C.2 Addressing the Spatial Correlation of Climate

Because climate is spatially correlated, a concern is that our estimates may reflect migrants' aversion to traveling long distances rather than responses to climate distance. As discussed in the main text (Section 2.2), physical distance explains at most 1.6% of the variation in temperature distance between countries of origin and U.S. counties (Table A.2, columns 1–3). In the internal migration context, physical distance explains up to 20% of the variation in temperature distance (Table A.2, columns 4–6). This is higher than in the international context, but remains modest, leaving substantial independent variation in climate distance beyond what can be attributed to geographic distance alone.

While we find these patterns reassuring, we seek to address any remaining concerns about the spatial correlation of climate in different ways. In the main specification, we control for the geographic distance between county of origin and county of destination—accounting for the fact that most people make short-distance moves to places that mechanically have similar climates. However, since controlling linearly for physical distance may not adequately address all concerns related to the spatial correlation of climate, we perform a series of additional exercises.

In Table C.7, we replicate the baseline specification (reported in column 1) separately by direction of movement. For historical internal migration (Panel A), the coefficient on temperature distance remains negative and statistically significant for most directions, with the exception of North–South moves (column 7). For modern internal migration (Panel B), the estimates remain quantitatively and qualitatively similar to the baseline for horizontal moves (columns 2–4). By contrast, coefficients are less precisely estimated for vertical migration (columns 5–7), and the coefficient for South–North moves (column 6) is positive and statistically significant. Overall, the results indicate that the relationship between temperature distance and migration is stronger for horizontal than for vertical moves, consistent with historical accounts emphasizing the role of east–west migration in shaping U.S. settlement patterns (Steckel, 1983). Notably, this horizontal gradient persists into the modern period.

In Table C.8, we assess the robustness of the results to excluding various subsets of moves: adjacent counties (column 2), next-to-adjacent counties (column 3), moves within 100 km (column 4), within-state migration (column 5), moves across adjacent states (column 6), and states with potentially distinctive migration patterns, such as California (column 7) and Florida (column 8). Despite the substantial reduction in sample size, the coefficients remain negative and statistically significant in both the historical (Panel A) and modern (Panel B) migration samples. In both cases, the magnitude of the estimates declines when short-distance moves are excluded—a pattern that is more pronounced for modern migration—but the coefficients remain precisely estimated.

In Table C.9, we subject the baseline specification to a particularly demanding robustness test by flexibly controlling for geographic distance using quadratic, cubic, and quartic polynomials. These controls absorb a large share of the spatial variation that could potentially confound the relationship between climate distance and migration. In the historical internal migration sample (Panel A), the coefficient on temperature distance remains negative and statistically significant, though somewhat attenuated in magnitude, indicating that the relationship is not driven by simple nonlinearities in geographic distance. In the modern sample (Panel B), the estimates become smaller and less stable, consistent with the weaker climate–migration relationship documented elsewhere in the paper. Taken together, these patterns suggest that while this specification places substantial demands on the data, the climate-distance effect remains robust in the historical context—when climate-specific skills were likely more important—whereas it is understandably weaker in the modern period, when adaptation technologies such as air conditioning may have reduced the relevance of climate matching.

Table C.7. Climate Distance and Migration, by Direction of Move

Dep. Variable:	Number of Migrants						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: U.S. Internal Migration, 1850–1940							
Temperature Distance	-0.315***	-0.254***	-0.074***	-0.105***	-0.176***	-0.133*	-0.030
	(0.019)	(0.019)	(0.024)	(0.022)	(0.031)	(0.078)	(0.040)
Observations	31,092,709	20,437,696	11,152,402	9,096,593	10,458,754	4,175,383	6,189,623
Pseudo R-squared	0.728	0.713	0.768	0.769	0.816	0.856	0.848
Panel B: U.S. Internal Migration, 2011–2019							
Temperature Distance	-0.237***	-0.234***	-0.284***	-0.137**	-0.041	0.171**	0.017
	(0.041)	(0.038)	(0.074)	(0.058)	(0.065)	(0.071)	(0.084)
Observations	25,088,941	12,364,382	3,551,971	4,403,901	6,454,843	2,625,742	1,779,170
Pseudo R-squared	0.776	0.773	0.847	0.816	0.834	0.872	0.914
Direction	All	Horizontal	East–West	West–East	Vertical	South–North	North–South
Origin×Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The samples in Panels A–B include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each year from 2011–2012 to 2018–2019, respectively. Number of Migrants in Panel A is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. Number of Migrants in Panel B is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the internal baseline specifications from Table 1 columns 3–4. Columns 2-7 retain only specific directions of county pairs. All regressions include precipitation and geographic distances, county-of-origin by decade (Panel A) or year (Panel B) and county-of-destination by decade/year fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.8. Climate Distance and Migration: Dropping Selected Counties

Dep. Variable:	Number of Migrants							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: U.S. Internal Migration, 1850–1940								
Temperature Distance	-0.315***	-0.221***	-0.192***	-0.190***	-0.199***	-0.107***	-0.319***	-0.317***
	(0.019)	(0.011)	(0.010)	(0.010)	(0.012)	(0.011)	(0.019)	(0.019)
Observations	31,092,709	31,011,575	30,857,496	30,858,031	29,956,219	25,859,734	31,079,703	31,075,493
Pseudo R-squared	0.728	0.710	0.691	0.690	0.672	0.668	0.724	0.729
Panel B: U.S. Internal Migration, 2011–2019								
Temperature Distance	-0.237***	-0.074***	-0.033***	-0.043***	-0.042***	-0.041***	-0.247***	-0.253***
	(0.041)	(0.018)	(0.012)	(0.013)	(0.013)	(0.013)	(0.042)	(0.039)
Observations	25,088,941	14,187,151	8,428,499	8,820,539	5,818,035	5,747,484	24,941,618	24,755,769
Pseudo R-squared	0.776	0.795	0.816	0.829	0.763	0.764	0.762	0.774
Sample	Full	Drop Adjacent Counties	Drop Adjacent & Next to Adjacent Counties	Drop Moves < 100km	Between States Only	Drop Adjacent States	Drop California	Drop Florida
Origin×Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The samples in Panels A–B include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each year from 2011–2012 to 2018–2019, respectively. Number of Migrants in Panel A is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. Number of Migrants in Panel B is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the internal baseline specifications from Table 1 columns 3–4. Column 2 drops adjacent county pairs. Column 3 drops adjacent and next-to-adjacent county pairs. Column 4 removes county pairs with a migration distance of less than 100 km. Column 5 retains only county pairs across different states. Column 6 drops county pairs from adjacent states. Column 7 removes county pairs where both origin and destination are in California. Column 8 removes county pairs where both origin and destination are in Florida. All regressions include precipitation and geographic distances, county-of-origin by decade (Panel A) or year (Panel B) and county-of-destination by decade/year fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.9. U.S. Internal Migration: Non-Linear Geographic Distance Controls

Dep. Variable:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Panel A: U.S. Internal Migration, 1850–1940					
Temperature Distance	-0.315***	-0.208***	-0.117***	-0.116***	-0.097***
	(0.019)	(0.015)	(0.011)	(0.010)	(0.009)
Observations	31,092,699	31,092,699	31,092,699	31,092,699	31,092,699
Pseudo R-squared	0.728	0.767	0.794	0.811	0.830
Mean Temp. Dist.	4.923	4.923	4.923	4.923	4.923
SD Temp. Dist.	3.578	3.578	3.578	3.578	3.578
Panel B: U.S. Internal Migration, 2011–2019					
Temperature Distance	-0.237***	-0.045	0.053**	0.018	0.071***
	(0.041)	(0.032)	(0.022)	(0.017)	(0.018)
Observations	25,088,936	25,088,936	25,088,936	25,088,936	25,088,936
Pseudo R-squared	0.776	0.827	0.867	0.885	0.904
Mean Temp. Dist.	4.989	4.989	4.989	4.989	4.989
SD Temp. Dist.	3.632	3.632	3.632	3.632	3.632
Physical distance controls	Linear	Quadratic	Cubic	Quartic	Log
Origin×Decade/Year FE	Yes	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The samples in Panels A–B include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each year from 2011–2012 to 2018–2019, respectively. Number of Migrants in Panel A is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. Number of Migrants in Panel B is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the internal baseline specifications from Table 1 columns 3–4. Columns 2, 3, and 4 control, respectively, for the second, third, and fourth order polynomials of physical distance, always including lower-order terms. Column 5 controls for log geographic distance. All regressions also control for precipitation and geographic distances, county-of-origin by decade (Panel A) or year (Panel B) and county-of-destination by decade/year fixed effects. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C.3 Including Bilateral Controls

In this section, we assess the robustness of the baseline results to the inclusion of bilateral controls. We begin by controlling for additional geographic attributes—such as elevation, ruggedness, and coastal access—that may be correlated with climate distance and may independently shape migration patterns (Albouy et al., 2021). Tables C.10 and C.11 report results for international immigration and U.S. internal migration, respectively. In both tables, Panel A presents estimates for historical migration, while Panel B reports results for modern migration. With the exception of specifications that add ruggedness distance in the modern international context (Table C.10, columns 3 and 5), the point estimates remain in line with the baseline results reported in column 1.

Next, we turn to latitude and longitude. Especially in the context of historical internal migration, latitude was often cited as a key factor shaping predominantly horizontal migration patterns (Steckel, 1983). Tables C.12 and C.13 report results for international and internal migration, respectively. In both tables, Panel A presents estimates for historical migration, while Panel B reports results for modern migration. Given the prominence of these variables in the migration literature, we also report the coefficients on the additional controls. Columns 2 and 3 introduce latitude and longitude distance separately; column 4 includes both measures jointly while excluding geographic distance; column 5 further controls for geographic distance.

Across specifications, the results are broadly robust. The main exception arises for modern international migration (Table C.12, Panel B), where the coefficient on temperature distance becomes statistically insignificant in columns 3–5 once latitude and longitude controls are introduced. For historical internal migration (Table C.13, Panel A), the coefficient on latitude distance is consistently negative and statistically significant, confirming long-standing historical accounts that emphasize the importance of latitudinal sorting in U.S. migration (Steckel, 1983). Notably, this latitudinal pattern persists in the modern internal migration sample as well (Panel B), suggesting a durable role for north–south climatic differences in shaping internal mobility even as other climate-related mechanisms have weakened.

Another feature that may be correlated with climate—and that may matter for migration decisions, particularly among farmers—is soil type. Soil formation is a long-run geological process shaped in part by local biomes and vegetation histories, which themselves reflect historical climatic conditions. Although formal soil science concepts—such as soil pH—were not articulated until the early 20$^{\text{th}}$ century, farmers likely possessed practical knowledge of soil characteristics through experience. Using gridded soil data, we construct county-level averages of four soil attributes: bulk density (column 2), organic matter content (column 3), soil pH (column 4), and water-holding capacity (column 5). Column 6 includes all soil measures jointly.[38] We report results for each of the four contexts—historical international, modern international, historical internal, and modern internal migration—in Tables C.14 through C.17.

Across specifications, the coefficient on temperature distance remains close to its baseline value. By contrast, the coefficients on the various soil-distance measures are unstable in the international immigration setting (Tables C.14 and C.15). For internal migration, however, soil distance is strongly negatively associated with migration in most specifications and is often precisely estimated (Tables C.16 and C.17). This pattern is consistent with historical and anecdotal accounts suggesting that farmer migrants actively sought destinations with soil characteristics similar to those of their places of origin.[39]

In Table C.18, we address the concern that the historical internal migration results may be driven by large changes in transportation costs over the long period we study (1850–1940). In column 2, we interact geographic distance with decade dummies to allow its effect to vary flexibly over time, capturing potential correlations between distance-related frictions and climate distance. Reassuringly, the coefficient on temperature distance remains close to the baseline estimate reported in column 1.

Next, we use data on the U.S. railroad network expansion from Donaldson and Hornbeck (2016) to compute bilateral time-varying transportation costs for each U.S. county-pair. To this end, we augment the railroad network with straight connections within and between county centroids, which are meant to proxy stagecoach connections. Then, we assign a weight to each edge of the resulting network. Transportation costs can be expressed either in hours or in dollars. In the former case, we use as weights 25 miles per hour for a railroad edge and 8 miles per hour for all other edges. In the latter case, we use \$1.5 per mile for a train ride and \$0.03 for all other edges.[40] With this weighted network at hand, we compute the least-cost paths between each county pair, for each decade between 1860 and 1920.[41]

In columns 3 and 4, we augment the baseline specification by controlling for time-varying county-pair travel costs, measured in hours and dollars, respectively. Columns 5 and 6 replicate

[38] Bulk density is the weight of dry soil per unit volume, with higher values indicating compacted soil. Organic matter concentration is the percentage of decomposed plant and animal material in the soil—features that can influence soil fertility and structure. Soil pH is a measure of soil acidity that affects nutrient availability and the types of crops that can be grown. Soil water content refers to the water holding capacity of the soil, which depends on soil texture and structure. Spatial averages of each soil characteristic for each county were computed using the OpenLandMap product available on Google Earth Engine with data from Hengl (2018a), Hengl and Wheeler (2018), Hengl (2018b), and Hengl and Gupta (2019).

[39] Steckel (1983) also conjectured that the east-west migration gradient prevailing in the U.S. during the 19$^{\text{th}}$ century was partly explained by farmers seeking similar soil types. Another anecdotal example is that of the "Cajun Prairie" region of Southwest Louisiana, which was first intensively farmed by Midwestern migrants in the 1880s who were attracted by the similar Mollisol-type prairie soils (see `https://www.loc.gov/item/sn88064676/`).

[40] These numbers were gathered from various sources, including the Poor's Railroad Manuals available at `http://www.pacificng.com/template.php?page=/ref/rrmanuals/index.htm` and the Annual report on the railroads of New York.

[41] Since railroad expansion data are available only through 1920, we assign the 1920 values to 1930 and 1940. The results are unchanged when excluding observations from 1930 and 1940 from the analysis.

columns 3 and 4 but exclude geographic distance. As expected, the coefficient on travel costs is negative and statistically significant. More importantly, the coefficient on temperature distance remains similar in magnitude to that in the baseline specification.

Finally, one may be concerned that origin–destination differences in economic or demographic characteristics correlated with climate distance could confound the relationship between climate and migration. The long-difference analysis in Table 2 mitigates this concern by controlling for origin-by-destination fixed effects. For historical internal migration, we further assess robustness by controlling for a rich set of origin–destination characteristics measured at the beginning of each decade using the full-count U.S. Census of Population. Figure C.1 replicates the baseline specification from Table 1, column 3 (shown by the leftmost dot), and sequentially adds county-pair differences in these characteristics.

The second, third, and fourth dots from the left control for the county-pair difference in labor force participation, manufacturing employment share, and agricultural employment share, respectively.[42] The subsequent dots include the difference in, respectively: the Black, the urban, and the immigrant population share; sex ratios; and population density. Finally, we consider three forces that have been shown to shape population movement and economic activity during our sample period: exposure to the frontier (Bazzi et al., 2020); market access (Donaldson and Hornbeck, 2016); and connection to railroads (Atack and Margo, 2011).[43] The very last dot includes all variables simultaneously. Coefficients on climate distances always remain negative, precisely estimated, and close to those from the baseline specification.

[42] In all cases, the variables are defined for men 15–64 in the baseline decade. Results are unchanged when using different age thresholds, extending the sample to women, or when considering a larger set of economic outcomes.

[43] We measure frontier exposure as the total number of years a county was on the frontier, according to Bazzi et al. (2020). Because this is a time-invariant control, we interact it with decade fixed effects to allow for differential trends over time. Both market access and dummies for being connected to railroads in a given decade are taken from Donaldson and Hornbeck (2016). The measure of market access is constant after 1920, and so we use this value for subsequent decades.

Table C.10. International Immigration: Controlling for Distance in Geographic Features

Dep. Variable:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Panel A: International Immigration, 1880–1920					
Temperature Distance	-0.175***	-0.185***	-0.175***	-0.175***	-0.187***
	(0.021)	(0.020)	(0.021)	(0.021)	(0.021)
Observations	1,939,195	1,914,683	1,914,683	1,939,195	1,914,683
Pseudo R-squared	0.916	0.918	0.916	0.916	0.918
Mean Temp. Dist.	9.622	9.625	9.625	9.622	9.625
SD Temp. Dist.	5.754	5.761	5.761	5.754	5.761
Panel B: International Immigration, 1970–2010					
Temperature Distance	-0.019*	-0.020*	-0.008	-0.019*	-0.009
	(0.011)	(0.011)	(0.009)	(0.011)	(0.009)
Observations	2,475,100	2,445,142	2,445,142	2,475,100	2,445,142
Pseudo R-squared	0.893	0.894	0.895	0.893	0.896
Mean Temp. Dist.	9.621	9.626	9.626	9.621	9.626
SD Temp. Dist.	5.685	5.691	5.691	5.685	5.691
Geographic controls		Elevation	Ruggedness	Coastal	All
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes

Notes: The samples in Panels A–B include U.S. county–country pairs for each decade over the 1880–1920 (except 1890) and 1970–2010 periods, respectively. Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade (except for 1880, when information on year of arrival is not available). Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the international baseline specifications from Table 1 columns 1–2. Column 2 additionally controls for the distance in elevation between o and d. Column 3 additionally controls for the distance in the standard deviation of elevation. Column 4 additionally controls for a dummy equal to one if both origin and destination are coastal (distance to coast equal to zero). Column 5 includes all three geographic controls. All regressions also control for precipitation and geographic distances, country of origin by decade and county of destination by decade fixed effects. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.11. U.S. Internal Migration: Controlling for Distance in Geographic Features

Dep. Variable:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Panel A: U.S. Internal Migration, 1850–1940					
Temperature Distance	-0.315***	-0.302***	-0.315***	-0.315***	-0.300***
	(0.019)	(0.018)	(0.019)	(0.019)	(0.019)
Observations	31,092,699	31,092,699	31,092,699	31,092,699	31,092,699
Pseudo R-squared	0.729	0.730	0.729	0.729	0.730
Mean Temp. Dist.	4.923	4.923	4.923	4.923	4.923
SD Temp. Dist.	3.578	3.578	3.578	3.578	3.578
Panel B: U.S. Internal Migration, 2011–2019					
Temperature Distance	-0.237***	-0.197***	-0.231***	-0.237***	-0.195***
	(0.041)	(0.040)	(0.041)	(0.041)	(0.042)
Observations	25,088,936	25,088,936	25,088,936	25,088,936	25,088,936
Pseudo R-squared	0.776	0.783	0.781	0.776	0.785
Mean Temp. Dist.	4.989	4.989	4.989	4.989	4.989
SD Temp. Dist.	3.632	3.632	3.632	3.632	3.632
Geographic controls		Elevation	Ruggedness	Coastal	All
Origin×Decade/Year FE	Yes	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The samples in Panels A–B include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each year from 2011–2012 to 2018–2019, respectively. Number of Migrants in Panel A is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. Number of Migrants in Panel B is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the internal baseline specifications from Table 1 columns 3–4. Column 2 additionally controls for the distance in elevation between the county of origin and the county of destination. Column 3 additionally controls for the distance in the standard deviation of elevation. Column 4 additionally controls for a dummy equal to one if both origin and destination counties are coastal. Column 5 includes all three geographic controls. All regressions also control for precipitation and geographic distances, county of origin by decade (Panel A) or year (Panel B) and county of destination by decade (Panel A) or year (Panel B) fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.12. International Immigration: Controlling for Latitude and Longitude Distance

Dep. Variable:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Panel A: International Immigration, 1880–1920					
Temperature Distance	-0.175***	-0.166***	-0.174***	-0.168***	-0.173***
	(0.021)	(0.026)	(0.021)	(0.026)	(0.027)
Distance (latitude)		-0.016		-0.013	-0.076***
		(0.025)		(0.024)	(0.028)
Distance (longitude)			0.029	0.029	-0.012
			(0.022)	(0.022)	(0.020)
Observations	1,939,195	1,914,683	1,914,683	1,914,683	1,914,683
Pseudo R-squared	0.916	0.916	0.917	0.917	0.914
Mean Temp. Dist.	9.622	9.625	9.625	9.625	9.625
SD Temp. Dist.	5.754	5.761	5.761	5.761	5.761
Panel B: International Immigration, 1970–2010					
Temperature Distance	-0.019*	-0.049**	-0.011	-0.022	-0.004
	(0.011)	(0.021)	(0.009)	(0.018)	(0.021)
Distance (latitude)		0.036*		0.013	-0.040***
		(0.019)		(0.015)	(0.015)
Distance (longitude)			-0.030***	-0.030***	-0.043***
			(0.004)	(0.004)	(0.005)
Observations	2,475,100	2,445,142	2,445,142	2,445,142	2,445,142
Pseudo R-squared	0.893	0.894	0.898	0.899	0.888
Mean Temp. Dist.	9.621	9.626	9.626	9.626	9.626
SD Temp. Dist.	5.685	5.691	5.691	5.691	5.691
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes

Notes: The samples in Panels A–B include U.S. county-country pairs for each decade over the 1880–1920 (except for 1890) and 1970–2010 periods, respectively. Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade (except for 1880, when information on year of arrival is not available). Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the international baseline specifications from Table 1 columns 1–2. Columns 2–3 add controls for the distance in latitude and longitude, respectively. Column 4 includes both measures simultaneously. Column 5 replicates column 4 without controlling for geographic distance. All regressions also control for precipitation distance, country of origin by decade and county of destination by decade fixed effects. Standard errors, reported in parentheses, are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.13. U.S. Internal Migration: Controlling for Latitude and Longitude Distance

Dep. Variable:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Panel A: U.S. Internal Migration, 1850–1940					
Temperature Distance	-0.315***	-0.156***	-0.188***	-0.155***	-0.158***
	(0.019)	(0.023)	(0.020)	(0.022)	(0.023)
Distance (latitude)		-0.225***		-0.112***	-0.339***
		(0.022)		(0.032)	(0.023)
Distance (longitude)			0.219***	0.140***	-0.131***
			(0.024)	(0.035)	(0.009)
Observations	31,092,699	31,092,699	31,092,699	31,092,699	31,092,699
Pseudo R-squared	0.728	0.733	0.733	0.734	0.730
Mean Temp. Dist.	4.923	4.923	4.923	4.923	4.923
SD Temp. Dist.	3.578	3.578	3.578	3.578	3.578
Panel B: U.S. Internal Migration, 2011–2019					
Temperature Distance	-0.237***	-0.148***	-0.221***	-0.147***	-0.147***
	(0.041)	(0.042)	(0.042)	(0.042)	(0.042)
Distance (latitude)		-0.125**		-0.231***	-0.253***
		(0.049)		(0.052)	(0.050)
Distance (longitude)			0.025	-0.119**	-0.143***
			(0.050)	(0.053)	(0.015)
Observations	25,088,936	25,088,936	25,088,936	25,088,936	25,088,936
Pseudo R-squared	0.776	0.777	0.777	0.778	0.778
Mean Temp. Dist.	4.989	4.989	4.989	4.989	4.989
SD Temp. Dist.	3.632	3.632	3.632	3.632	3.632
Origin×Decade/Year FE	Yes	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The samples in Panels A–B include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each year from 2011–2012 to 2018–2019, respectively. Number of Migrants in Panel A is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. Number of Migrants in Panel B is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. For the modern internal sample (Panel B), latitude/longitude controls are computed using population-weighted coordinates from 2010 and held constant across all years in the sample. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the internal baseline specifications from Table 1 columns 3–4. Columns 2–3 add controls for the distance in latitude and longitude, respectively. Column 4 includes both measures simultaneously. Column 5 replicates column 4 without controlling for geographic distance. All regressions also control for precipitation distance, county of origin by decade (Panel A) or year (Panel B) and county of destination by decade (Panel A) or year (Panel B) fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.14. International Immigration (1880–1920): Controlling for Soil Type Distance

Dep. Variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.175***	-0.176***	-0.176***	-0.176***	-0.176***	-0.176***
	(0.021)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Bulk Density Distance		-0.003**				-0.002
		(0.001)				(0.002)
Organic Matter Distance			-0.015***			-0.013**
			(0.005)			(0.006)
pH Distance				-0.003		-0.003
				(0.003)		(0.003)
Water Content Distance					-0.002	0.002
					(0.005)	(0.005)
Observations	1,939,195	1,780,907	1,780,907	1,780,907	1,780,907	1,780,907
Pseudo R-squared	0.916	0.918	0.918	0.918	0.918	0.918
Mean Temp. Dist.	9.622	9.622	9.622	9.622	9.622	9.622
SD Temp. Dist.	5.754	5.756	5.756	5.756	5.756	5.756
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes U.S. county-country pairs for each decade over the 1880–1920 (except for 1890). Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade (except for 1880, when information on year of arrival is not available). Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the historical international baseline specification from Table 1 column 1. Columns 2–5 add, separately, distance in each soil measure (bulk density, organic matter, pH, and water content). Column 6 includes all four soil distances. All regressions also control for precipitation and geographic distances, country of origin by decade and county of destination by decade fixed effects. Standard errors are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.15. International Immigration (1970–2010): Controlling for Soil Type Distance

Dep. Variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.019*	-0.019*	-0.019*	-0.019*	-0.019	-0.019*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Bulk Density Distance		-0.000				-0.001
		(0.001)				(0.002)
Organic Matter Distance			0.002			0.004
			(0.005)			(0.006)
pH Distance				-0.000		-0.000
				(0.002)		(0.003)
Water Content Distance					-0.001	-0.001
					(0.003)	(0.004)
Observations	2,475,100	2,271,350	2,271,350	2,271,350	2,271,350	2,271,350
Pseudo R-squared	0.893	0.896	0.896	0.896	0.896	0.896
Mean Temp. Dist.	9.621	9.621	9.621	9.621	9.621	9.621
SD Temp. Dist.	5.685	5.686	5.686	5.686	5.686	5.686
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes U.S. county-country pairs for each decade over the 1970–2010 periods. Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the modern international baseline specification from Table 1 column 2. Columns 2–5 add, separately, distance in each soil measure (bulk density, organic matter, pH, and water content). Column 6 includes all four soil distances. All regressions also control for precipitation and geographic distances, country of origin by decade and county of destination by decade fixed effects. Standard errors are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.16. U.S. Internal Migration (1850–1940): Controlling for Soil Type Distance

Dep. Variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.315***	-0.299***	-0.296***	-0.315***	-0.305***	-0.283***
	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)
Bulk Density Distance		-0.057***				-0.041***
		(0.006)				(0.007)
Organic Matter Distance			-0.301***			-0.212***
			(0.032)			(0.033)
pH Distance				-0.022***		-0.011
				(0.007)		(0.007)
Water Content Distance					-0.145***	-0.115***
					(0.013)	(0.013)
Observations	31,092,699	31,092,699	31,092,699	31,092,699	31,092,699	31,092,699
Pseudo R-squared	0.728	0.733	0.732	0.729	0.733	0.738
Mean Temp. Dist.	4.923	4.923	4.923	4.923	4.923	4.923
SD Temp. Dist.	3.578	3.578	3.578	3.578	3.578	3.578
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940. Number of Migrants is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the historical internal baseline specification from Table 1 column 3. Columns 2–5 add, separately, distance in each soil measure (bulk density, organic matter, pH, and water content). Column 6 includes all four soil distances. All regressions also control for precipitation and geographic distances, county of origin by decade and county of destination by decade fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.17. U.S. Internal Migration (2011–2019): Controlling for Soil Type Distance

Dep. Variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.237***	-0.222***	-0.225***	-0.234***	-0.225***	-0.216***
	(0.041)	(0.041)	(0.041)	(0.040)	(0.039)	(0.039)
Bulk Density Distance		-0.064***				-0.056***
		(0.012)				(0.010)
Organic Matter Distance			-0.133**			0.044
			(0.059)			(0.070)
pH Distance				-0.055***		-0.036***
				(0.010)		(0.010)
Water Content Distance					-0.213***	-0.187***
					(0.025)	(0.023)
Observations	25,088,936	25,088,936	25,088,936	25,088,936	25,088,936	25,088,936
Pseudo R-squared	0.776	0.781	0.777	0.778	0.782	0.786
Mean Temp. Dist.	4.989	4.989	4.989	4.989	4.989	4.989
SD Temp. Dist.	3.632	3.632	3.632	3.632	3.632	3.632
Origin×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Year FE	Yes	Yes	Yes	Yes	Yes	Yes

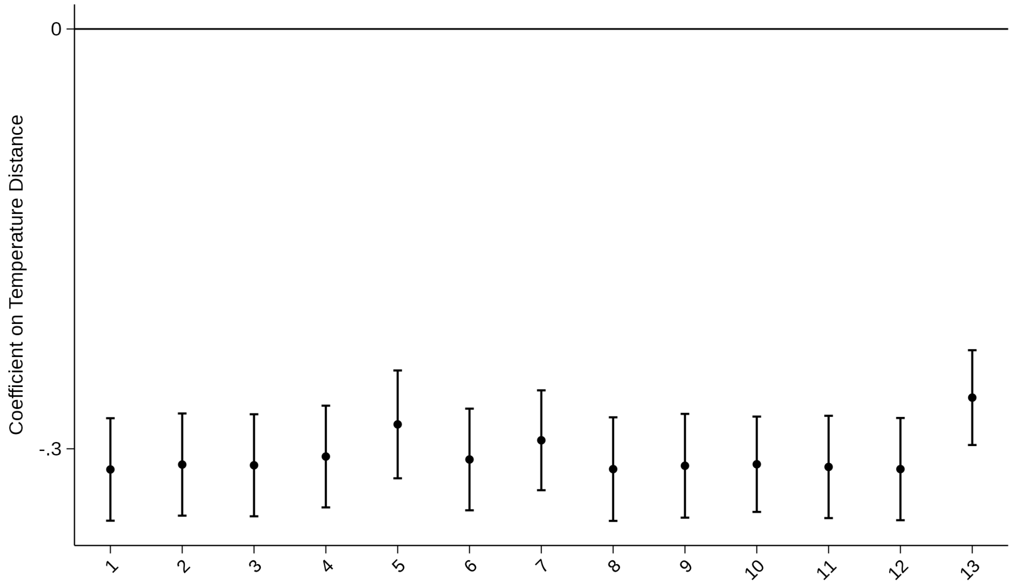
Notes: The sample includes county-pairs in the contiguous U.S. for each year from 2011–2012 to 2018–2019. Number of Migrants is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the modern internal baseline specification from Table 1 column 4. Columns 2–5 add, separately, distance in each soil measure (bulk density, organic matter, pH, and water content). Column 6 includes all four soil distances. All regressions also control for precipitation and geographic distances, county of origin by year and county of destination by year fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.18. U.S. Internal Migration (1850–1940): Controlling for Time-Varying Travel Distance

Dep. Variable:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.315***	-0.312***	-0.296***	-0.321***	-0.299***	-0.357***
	(0.019)	(0.018)	(0.018)	(0.019)	(0.019)	(0.024)
Distance (in 100km)	-0.194***		-0.090***	-0.187***		
	(0.014)		(0.016)	(0.015)		
Transportation Cost			-0.023***	-0.005	-0.037***	-0.082***
			(0.003)	(0.004)	(0.002)	(0.009)
Observations	31,092,699	31,092,699	29,961,412	29,961,412	29,961,412	29,961,412
Pseudo R-squared	0.728	0.729	0.739	0.730	0.733	0.710
Mean Temp. Dist.	4.923	4.923	4.932	4.932	4.932	4.932
SD Temp. Dist.	3.578	3.578	3.583	3.583	3.583	3.583
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940. Number of Migrants is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Column 1 replicates the historical internal baseline specification from Table 1 column 3. Column 2 replicates column 1 by interacting geographic distance with decade dummies. Columns 3–4 add time-varying travel costs (hours or dollars) in addition to physical distance; columns 5–6 replicate columns 3–4 without controlling for geographic distance. All regressions also control for precipitation and geographic distances, county of origin by decade and county of destination by decade fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure C.1. U.S. Internal Migration (1850–1940): Including County-Pair Controls



Notes: The figure plots the coefficient, with corresponding 95% confidence intervals, on the absolute difference in temperature between U.S. origin and destination counties for internal movers. The first dot represents the baseline specification from Table 1, column 3. Dots 2 through 12 represent results obtained by including, one at a time, the absolute value of the difference in the following county-pair variables (measured in the baseline decade): employment share; manufacturing employment share; agriculture employment share; black population share; urban population share; immigrant population share; sex ratios; population density; frontier exposure from Bazzi et al. (2020); market access from Donaldson and Hornbeck (2016); and a dummy equal to one for being connected to railroads. The final dot 13 reports results obtained by including all variables simultaneously. Standard errors are clustered at the U.S. state of origin by U.S. state of destination level.

C.4 Alternative Climate Statistics

In this section, we consider additional measures of climate distance. First, we consider seasonality. Mean annual temperature is the basis for the primary climate distance measure used in all of our analyses. While this measure is one of the simplest and most commonly used climate metrics, it masks a large degree of variation: Washington, DC, and San Francisco, for example, have similar annual mean temperatures, but very different seasonal patterns. To account for this dimension, Tables C.19 and C.20 replicate our preferred specifications for international and internal migration, reporting results for the historical and modern periods in columns 1–3 and 4–6, respectively. We augment the baseline by controlling for distance in climate seasonality using two approaches: the standard deviation of temperature and precipitation over the year (columns 2 and 5), and the difference between average annual maximum and minimum temperature and precipitation (columns 3 and 6).[44]

In all cases, similarity in climate variability increases migration, suggesting that migrants take seasonality into account. However, the coefficient on average temperature remains negative and statistically significant. Moreover, its size is an order of magnitude larger (in absolute value) than that of climate variability.

We also verify that our results are robust to alternative definitions of climate distance. Previous work has linked U.S. population growth patterns to seasonal temperature extremes, often proxied by January and July temperatures (Glaeser and Tobio, 2007). Tables C.21 and C.22 report results for international and internal migration, with historical and modern periods shown in columns 1–4 and 5–8, respectively. The estimated effects remain similar when temperature distance is constructed using the mean of annual maximum temperatures (columns 2 and 6), average summer temperatures from April to September (columns 3 and 7), or average winter temperatures from October to March (columns 4 and 8). These estimates can be directly compared to those based on annual mean temperature in the main paper (Table 1).

[44] Using the above example, San Francisco would have a lower values for its temperature seasonality given its mild climate year round compared to the summer and winter extremes in Washington, DC.

Table C.19. International Immigration: Controlling for Distance in Seasonality

Dep. Variable:	Number of Migrants					
	Historical: 1880–1920			Modern: 1970–2010		
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.175***	-0.192***	-0.190***	-0.019*	-0.010	-0.010
	(0.021)	(0.023)	(0.023)	(0.011)	(0.011)	(0.011)
Temp. SD Distance		0.051			-0.055***	
		(0.037)			(0.016)	
Temp. Range Distance			0.017			-0.019***
			(0.013)			(0.006)
Observations	1,939,195	1,939,195	1,939,195	2,475,100	2,475,100	2,475,100
Pseudo R-squared	0.916	0.917	0.917	0.893	0.894	0.894
Mean Temp. Dist.	9.775	9.775	9.775	9.791	9.791	9.791
SD Temp. Dist.	5.694	5.694	5.694	5.697	5.697	5.697
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The samples in columns 1–3 and 4–6 include U.S. county-country pairs for each decade over the 1880–1920 (except for 1890) and 1970–2010 periods, respectively. Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade (except for 1880, when information on year of arrival is not available). Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Columns 1 and 4 replicate the U.S. international baseline specifications from Table 1 columns 1 and 2, respectively. Columns 2 and 5 add controls for Temperature and Precipitation SD Distance, the absolute difference in the standard deviation of temperature and precipitation over the year. Columns 3 and 6 control for Temperature and Precipitation Range Distance, the absolute difference in the range of temperatures experienced over a year. Each measure of variability is first calculated at the annual level, and then averaged across years. All regressions also control for precipitation and geographic distances, country of origin by decade and county of destination by decade fixed effects. Standard errors are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.20. U.S. Internal Immigration: Controlling for Distance in Seasonality

Dep. Variable:	Number of Migrants					
	Historical: 1850–1940			Modern: 2011–2019		
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.315***	-0.306***	-0.302***	-0.237***	-0.190***	-0.186***
	(0.019)	(0.020)	(0.020)	(0.041)	(0.039)	(0.039)
Temp. SD Distance		-0.058			-0.317***	
		(0.045)			(0.071)	
Temp. Range Distance			-0.030*			-0.117***
			(0.016)			(0.026)
Observations	31,092,709	31,092,709	31,092,709	25,088,941	25,088,941	25,088,941
Pseudo R-squared	0.729	0.729	0.729	0.776	0.779	0.779
Mean Temp. Dist.	5.062	5.062	5.062	5.052	5.052	5.052
SD Temp. Dist.	3.667	3.667	3.667	3.662	3.662	3.662
Origin×Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The samples in columns 1–3 and 4–6 include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each calendar year from 2011–2012 to 2018–2019, respectively. In columns 1–3, Number of Migrants is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. In columns 4–6, Number of Migrants is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Columns 1 and 4 replicate the internal baseline specifications from Table 1 columns 3 and 4, respectively. Columns 2 and 5 add controls for Temperature and Precipitation SD Distance, the absolute difference in the standard deviation of temperature and precipitation over the year. Columns 3 and 6 control for Temperature and Precipitation Range Distance, the absolute difference in the range of temperatures experienced over a year. Each measure of variability is first calculated at the annual level, and then averaged across years. All regressions also control for precipitation and geographic distances, county of origin by decade (columns 1–3) or year (columns 4–6) and county of destination by decade (columns 1–3) or year (columns 4–6) fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.21. International Immigration: Alternative Climate Definitions

Dep. Variable:	Number of Migrants							
	Historical: 1880–1920				Modern: 1970–2010			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature Distance	-0.175***	-0.087***	-0.125***	-0.097***	-0.019*	-0.037***	-0.022**	-0.013**
	(0.021)	(0.022)	(0.023)	(0.015)	(0.011)	(0.011)	(0.011)	(0.006)
Observations	1,939,195	1,939,195	1,939,195	1,939,195	2,475,100	2,475,100	2,475,100	2,475,100
Pseudo R-squared	0.916	0.910	0.911	0.915	0.893	0.893	0.893	0.893
Mean Temp. Dist.	9.784	5.136	6.710	15.16	9.784	5.136	6.710	15.16
SD Temp. Dist.	5.696	4.329	4.839	8.103	5.696	4.329	4.839	8.103
Origin × Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination × Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Temperature Measurement	Yearly Avg.	Yearly Max	Summer	Winter	Yearly Avg.	Yearly Max	Summer	Winter

Notes: The samples in columns 1–4 and 5–8 include U.S. county-country pairs for each decade over the 1880–1920 (except for 1890) and 1970–2010 periods, respectively. Number of Migrants is the number of individuals from origin country o living in destination county d in census year t arrived in the previous decade (except for 1880, when information on year of arrival is not available). Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Columns 1 and 5 replicate the U.S. international baseline specifications from Table 1 columns 1 and 2, respectively. The other columns replicate this specification by calculating temperature distance with alternate definitions of temperature. Columns 2 and 6 use annual maximum, columns 3 and 7 use summer temperature (averaged from April–September), and columns 4 and 8 use winter temperature (averaged from October–March). All regressions also control for precipitation and geographic distances, country of origin by decade and county of destination by decade fixed effects. Standard errors are clustered at the country of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.22. U.S. Internal Migration: Alternative Climate Definitions

Dep. Variable:	Number of Migrants							
	Historical: 1850–1940				Modern: 2011–2019			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature Distance	-0.315***	-0.281***	-0.319***	-0.243***	-0.237***	-0.142***	-0.170***	-0.191***
	(0.019)	(0.023)	(0.022)	(0.016)	(0.041)	(0.040)	(0.043)	(0.033)
Observations	31,092,709	31,092,709	31,092,709	31,092,709	25,088,941	25,088,941	25,088,941	25,088,941
Pseudo R-squared	0.729	0.718	0.724	0.727	0.776	0.771	0.772	0.777
Mean Temp. Dist.	5.062	3.570	4.283	6.337	5.052	3.575	4.288	6.341
SD Temp. Dist.	3.667	2.758	3.178	4.620	3.662	2.762	3.183	4.624
Origin × Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination × Decade/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Temperature Measurement	Yearly Avg.	Yearly Max	Summer	Winter	Yearly Avg.	Yearly Max	Summer	Winter

Notes: The samples in columns 1–4 and 5–8 include county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940 and for each calendar year from 2011–2012 to 2018–2019, respectively. In columns 1–4, Number of Migrants is the number of individuals 15+ in the baseline year who moved from the origin to the destination county in the decade. In columns 5–8, Number of Migrants is the number of people who changed address between origin and destination counties, based on IRS Migration data from individual income tax returns. Temperature (resp., Precipitation) Distance is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination, measured in degrees Celsius (resp., millimeters). Distance is physical distance, expressed in 100 km, between the origin and the destination. Columns 1 and 5 replicate the internal baseline specifications from Table 1 columns 3 and 4, respectively. The other columns replicate this specification by calculating temperature distance with alternate definitions of temperature. Columns 2 and 6 use annual maximum, columns 3 and 7 use summer temperature (averaged from April–September), and columns 4 and 8 use winter temperature (averaged from October–March). All regressions also control for precipitation and geographic distances, county of origin by decade (columns 1–4) or year (columns 5–8) and county of destination by decade (columns 1–4) or year (columns 5–8) fixed effects. Standard errors, reported in parentheses, are clustered at the U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C.5 Additional Robustness Checks (Historical Internal Migration)

In Table C.23, we show that the historical internal migration results are robust to alternative sample restrictions and to using different datasets to measure migration. Column 1 reports the baseline specification, which includes all county pairs in the contiguous United States in each decade. Column 2 restricts the sample to counties already existing in 1860, while columns 3 and 4 further limit the sample to county pairs observed in all decades and to pairs with a strictly positive number of migrants, respectively. Across these exercises, the coefficient on temperature distance remains close to the baseline estimate.

We next address potential concerns related to the linked sample used to construct internal migration flows. In the baseline specification, we include all individuals aged 15 and older. In column 5, we restrict the sample to men aged 15 and above, reflecting the more traditional approach in the linked-census literature (Abramitzky et al., 2021). In column 6, we instead use the linked sample of men aged 15 and above constructed by Abramitzky et al. (2020).[45] Finally, in column 7, we address remaining concerns related to record linkage by exploiting a feature unique to the 1940 Census, which asked respondents about their place of residence five years earlier. We use this information to construct a migration matrix capturing moves between 1935 and 1940.[46] For consistency with the linked-sample analysis, we restrict attention to individuals who were at least 20 years old in 1940 and thus at least 15 years old at their location of origin in 1935. In all cases, the estimated effect of temperature distance remains stable in sign, magnitude, and statistical significance.

[45] The data from Abramitzky et al. (2020) and Price et al. (2021) both draw on the full-count U.S. Census but differ in scope and linking methodology: Price et al. (2021) provide broader coverage using machine-learning approaches, while Abramitzky et al. (2020) construct smaller samples using more conservative linking algorithms.

[46] As with the linked sample, we harmonize destination county boundaries using the Census Place Project for 1940 (Berkes et al., 2023). Because the Census Place Project cannot be applied to counties of origin in 1935, we rely instead on the crosswalk provided by Eckert et al. (2020).

Table C.23. U.S. Internal Migration: Sample Restrictions and Alternative Datasets

Dep. Variable:	Number of Migrants						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Temperature Distance	-0.315***	-0.312***	-0.312***	-0.253***	-0.306***	-0.254***	-0.236***
	(0.019)	(0.023)	(0.023)	(0.025)	(0.018)	(0.014)	(0.022)
Observations	31,092,699	15,660,641	15,573,129	2,387,764	31,086,215	30,992,489	4,825,169
Pseudo R-squared	0.729	0.750	0.750	0.663	0.723	0.726	0.764
Mean Temp. Dist.	4.923	4.590	4.580	2.691	4.923	4.917	5.062
SD Temp. Dist.	3.578	3.321	3.312	2.523	3.578	3.575	3.667
Origin×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes county-pairs in the contiguous U.S. for each decade from 1850–1860 to 1930–1940, unless otherwise stated. Temperature Distance (resp., Precipitation Distance) is the absolute difference in mean annual temperature (resp., precipitation) between the origin and the destination county, measured in degrees Celsius (resp., millimeters). Distance refers to the physical distance between counties (expressed in 100 km). Column 1 reports the baseline specification from Table 1, column 3. Column 2 restricts the sample to counties already existing in 1860. Column 3 and 4 further restricts the sample to county-pairs observed in all decades and with a strictly positive number of migrants, respectively. Columns 5 and 6 use alternative linked-census samples to construct internal migration flows: column 5 restricts the linked sample to men aged 15+ (Abramitzky et al., 2021), while column 6 uses the linked sample of men aged 15+ constructed by Abramitzky et al. (2020). Column 7 uses 1940 Census information on place of residence five years earlier to construct migration flows between 1935 and 1940; destination county boundaries are harmonized using the Census Place Project for 1940 (Berkes et al., 2023), while counties of origin in 1935 are harmonized using the crosswalk provided by Eckert et al. (2020). In column 7, the sample is restricted to individuals aged 20+ in 1940 (and thus 15+ at origin in 1935). All regressions also control for precipitation and geographic distances, county of origin by decade and county of destination by decade fixed effects. Standard errors, reported in parentheses, are clustered at U.S. state of origin by U.S. state of destination level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D Robustness Checks — Effects of Climate Mismatch

D.1 Deriving Predicted Migration from Weather Shocks

To address concerns that aggregate inflows from specific origin countries may be endogenous to U.S. economic conditions, we construct an alternative measure of predicted migration that is driven solely by exogenous weather shocks in origin countries. Our approach follows Sequeira et al. (2020).

We obtain historical temperature data from Luterbacher et al. (2004) and precipitation data from Pauling et al. (2006), measured at a 0.5° spatial resolution and at a seasonal frequency (spring, summer, autumn, winter) for each year between 1850 and 1929. For each origin country o, we compute country-level seasonal averages by aggregating gridcells weighted by the share of land under cultivation, using historical cropland maps from Ramankutty et al. (2002).[47] This yields a country-by-year panel of seasonal temperature and precipitation.

For each country and season, we categorize the distribution of temperature (and precipitation) into six bins: (i) 3+ standard deviations below the historical mean; (ii) 2–3 standard deviations below; (iii) 1–2 standard deviations below; (iv) 1–2 standard deviations above; (v) 2–3 standard deviations above; and (vi) 3+ standard deviations above. The omitted category consists of realizations within one standard deviation of the mean. This yields 24 temperature indicators and 24 precipitation indicators per country and year.

For each origin country o, we estimate the following equation separately:

$$\ln \text{EmigrantFlow}_{o,t+1} = \sum_{s\in\mathcal{S}} \sum_{k\in\mathcal{K}} \beta_{o,s,k}\, \mathbb{1}\{\text{Temp}_{o,s,t} \in k\} + \sum_{s\in\mathcal{S}} \sum_{k\in\mathcal{K}} \gamma_{o,s,k}\, \mathbb{1}\{\text{Precip}_{o,s,t} \in k\} + \varepsilon_{c,t}, \tag{7}$$

where $\ln \text{EmigrantFlow}_{o,t+1}$ is the log number of migrants from origin o entering the United States in year $t+1$, taken from Willcox (1929). The coefficients $\beta_{o,s,k}$ and $\gamma_{o,s,k}$ capture how extreme temperature or precipitation events in season s affect subsequent emigration. Estimating (7) separately for each origin allows the response to weather shocks to vary flexibly across countries.

Using the estimated coefficients, we construct predicted flows for each country and year:

$$\ln \widehat{\text{EmigrantFlow}}_{o,t+1} = \sum_{s,k} \widehat{\beta}_{o,s,k}\, \mathbb{1}\{\text{Temp}_{o,s,t} \in k\} + \sum_{s,k} \widehat{\gamma}_{o,s,k}\, \mathbb{1}\{\text{Precip}_{o,s,t} \in k\}.$$

[47] We use the 1850–1929 time period to calculate the long-run averages, following Sequeira et al. (2020). Results are virtually unchanged when using alternative time windows.

This yields the predicted total inflow of migrants from each country o into the United States in decade t due solely to weather shocks in sending countries.[48]

Finally, we interact $\widehat{\text{ImmigrantFlow}}_{o,t-1}$ with the timing of local railroad access, following our baseline design.

D.2 Tables

[48] Consistent with Sequeira et al. (2020), we find that predicted flows are strongly correlated with actual flows.

Table D.1. The Effects of Climate Mismatch on Agriculture: Robustness

Dep. Variable:	Log Value of Crops	Share Farmland	Log Value per Acre: Crop	Log Value per Acre: Farm	Log Value per Acre: Equipment	Log Number of Farms	Log Avg Farm Size
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Baseline							
Temperature Mismatch	-0.178***	-0.027***	-0.100***	-0.147***	-0.106***	-0.139***	0.024
	(0.029)	(0.007)	(0.022)	(0.022)	(0.019)	(0.018)	(0.017)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.945	0.925	0.948	0.958	0.956	0.924	0.858
KP F-stat	610.1	610.1	610.1	610.1	610.1	610.1	610.1
Panel B: Predicted Migration Flows							
Temperature Mismatch	-0.173***	-0.023***	-0.104***	-0.140***	-0.102***	-0.128***	0.021
	(0.029)	(0.007)	(0.023)	(0.023)	(0.019)	(0.017)	(0.017)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.944	0.924	0.948	0.958	0.956	0.923	0.858
KP F-stat	527.7	527.7	527.7	527.7	527.7	527.7	527.7
Panel C: Long Differences (1920-1880)							
Temperature Mismatch	-0.153***	-0.031***	-0.079***	-0.183***	-0.123***	-0.159***	0.043**
	(0.034)	(0.008)	(0.026)	(0.026)	(0.022)	(0.021)	(0.021)
Observations	2,116	2,116	2,116	2,116	2,116	2,116	2,116
R-squared	0.305	0.427	0.291	0.483	0.272	0.403	0.416
KP F-stat	420.8	420.8	420.8	420.8	420.8	420.8	420.8
Panel D: State×Decade FE							
Temperature Mismatch	-0.238***	-0.052***	-0.093***	-0.072***	-0.048**	-0.186***	0.012
	(0.032)	(0.009)	(0.025)	(0.022)	(0.024)	(0.024)	(0.018)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.954	0.938	0.957	0.969	0.963	0.939	0.876
KP F-stat	286.7	286.7	286.7	286.7	286.7	286.7	286.7
Panel E: Trim Temperature Mismatch							
Temperature Mismatch	-0.169***	-0.024***	-0.100***	-0.126***	-0.097***	-0.126***	0.029*
	(0.028)	(0.007)	(0.021)	(0.021)	(0.017)	(0.017)	(0.017)
Observations	8,283	8,283	8,283	8,283	8,283	8,283	8,283
R-squared	0.946	0.925	0.950	0.959	0.957	0.925	0.861
KP F-stat	701.0	701.0	701.0	701.0	701.0	701.0	701.0
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1880 Dep. Var. Mean	512,108	0.638	2.121	18.59	0.750	1,774	145.9

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 and at least 100 farms in each decade from 1880 to 1920. The year 1890 is excluded because it is not possible to construct the measure of climate mismatch. Panel A reports the 2SLS estimates of Table 3. Panel B replaces the baseline instrument with a version that uses origin-specific migration flows predicted from weather shocks following Sequeira et al. (2020). Panel C reports 1920-1880 long-difference regressions. Panel D replaces census region-by-decade fixed effects with state-by-decade fixed effects. Panel E excludes counties with mismatch values above the 99th percentile or below the 1st percentile. All regressions include county fixed effects, census region by decade fixed effects (except for Panel D), and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.2. The Effects of Climate Mismatch on Agriculture: Robustness/2

Dep. Variable:	Log Value of Crops	Share Farmland	Log Value per Acre: Crop	Log Value per Acre: Farm	Log Value per Acre: Equipment	Log Number of Farms	Log Avg Farm Size
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Baseline							
Temperature Mismatch	-0.178***	-0.027***	-0.100***	-0.147***	-0.106***	-0.139***	0.024
	(0.029)	(0.007)	(0.022)	(0.022)	(0.019)	(0.018)	(0.017)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.945	0.925	0.948	0.958	0.956	0.924	0.858
KP F-stat	610.1	610.1	610.1	610.1	610.1	610.1	610.1
Panel B: 1880 Immigrant Share×Decade FE							
Temperature Mismatch	-0.174***	-0.027***	-0.099***	-0.149***	-0.102***	-0.137***	0.027*
	(0.028)	(0.006)	(0.022)	(0.022)	(0.018)	(0.018)	(0.016)
Observations	8,400	8,400	8,400	8,400	8,400	8,400	8,400
R-squared	0.946	0.927	0.949	0.959	0.956	0.926	0.859
KP F-stat	612.5	612.5	612.5	612.5	612.5	612.5	612.5
Panel C: Crop Suitability×Decade FE							
Temperature Mismatch	-0.143***	-0.021***	-0.088***	-0.141***	-0.097***	-0.117***	0.025
	(0.028)	(0.006)	(0.023)	(0.022)	(0.019)	(0.016)	(0.017)
Observations	8,448	8,448	8,448	8,448	8,448	8,448	8,448
R-squared	0.948	0.927	0.949	0.958	0.957	0.927	0.859
KP F-stat	581.4	581.4	581.4	581.4	581.4	581.4	581.4
Panel D: Frontier Experience×Decade FE							
Temperature Mismatch	-0.174***	-0.026***	-0.098***	-0.138***	-0.104***	-0.134***	0.024
	(0.028)	(0.007)	(0.022)	(0.021)	(0.018)	(0.018)	(0.017)
Observations	8,420	8,420	8,420	8,420	8,420	8,420	8,420
R-squared	0.945	0.925	0.950	0.959	0.957	0.924	0.860
KP F-stat	599.1	599.1	599.1	599.1	599.1	599.1	599.1
Panel E: All Controls							
Temperature Mismatch	-0.139***	-0.019***	-0.088***	-0.137***	-0.095***	-0.112***	0.027
	(0.027)	(0.006)	(0.022)	(0.021)	(0.018)	(0.016)	(0.017)
Observations	8,340	8,340	8,340	8,340	8,340	8,340	8,340
R-squared	0.949	0.929	0.950	0.960	0.958	0.929	0.862
KP F-stat	571.9	571.9	571.9	571.9	571.9	571.9	571.9
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1880 Dep. Var. Mean	512,108	0.638	2.121	18.59	0.750	1,774	145.9

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 and at least 100 farms in each decade from 1880 to 1920. The year 1890 is excluded because it is not possible to construct the measure of climate mismatch. Panel A reports the 2SLS estimates of Table 3. Panels B–D add controls for: (i) 1880 immigrant population share (Panel B), (ii) crop-suitability measure based on climate and soil constraint from Ramankutty et al. (2002) (Panel C), and (iii) total frontier experience from Bazzi et al. (2020) (Panel D). Panel E includes all three controls jointly. All regressions include county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.3. The Effects of Climate Mismatch on Agriculture: Robustness/3

Dep. Variable:	Log Value of Crops	Share Farmland	Log Value per Acre: Crop	Log Value per Acre: Farm	Log Value per Acre: Equipment	Log Number of Farms	Log Avg Farm Size
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Baseline							
Temperature Mismatch	-0.178***	-0.027***	-0.100***	-0.147***	-0.106***	-0.139***	0.024
	(0.029)	(0.007)	(0.022)	(0.022)	(0.019)	(0.018)	(0.017)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.945	0.925	0.948	0.958	0.956	0.924	0.858
KP F-stat	610.1	610.1	610.1	610.1	610.1	610.1	610.1
Panel B: Predicted Precipitation Distance							
Temperature Mismatch	-0.240***	-0.043***	-0.125***	-0.153***	-0.105***	-0.147***	-0.009
	(0.028)	(0.006)	(0.022)	(0.021)	(0.018)	(0.018)	(0.015)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.952	0.936	0.950	0.958	0.956	0.925	0.868
KP F-stat	644.9	644.9	644.9	644.9	644.9	644.9	644.9
Panel C: Predicted Geographic Distance							
Temperature Mismatch	-0.178***	-0.027***	-0.100***	-0.147***	-0.106***	-0.139***	0.024
	(0.029)	(0.007)	(0.022)	(0.022)	(0.019)	(0.018)	(0.017)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.945	0.925	0.948	0.958	0.956	0.924	0.858
KP F-stat	609.9	609.9	609.9	609.9	609.9	609.9	609.9
Panel D: Both Distance Controls							
Temperature Mismatch	-0.240***	-0.043***	-0.125***	-0.153***	-0.105***	-0.147***	-0.009
	(0.028)	(0.006)	(0.022)	(0.021)	(0.018)	(0.018)	(0.015)
Observations	8,464	8,464	8,464	8,464	8,464	8,464	8,464
R-squared	0.952	0.936	0.950	0.958	0.956	0.925	0.868
KP F-stat	644.7	644.7	644.7	644.7	644.7	644.7	644.7
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1880 Dep. Var. Mean	512,108	0.638	2.121	18.59	0.750	1,774	145.9

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 and at least 100 farms in each decade from 1880 to 1920. The year 1890 is excluded because it is not possible to construct the measure of climate mismatch. Panel A reports the 2SLS estimates of Table 3. Panel B replicates Panel A controlling for predicted precipitation distance, constructed analogously to predicted temperature distance (see Section 3.1). Panel C replicates Panel A controlling for predicted geographic distance, constructed analogously to predicted temperature and precipitation distance. Panel D includes both controls simultaneously. All regressions include county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.4. The Effects of Climate Mismatch on Manufacturing and Population: Robustness

		Log Manufacturing				Log
Dep. Variable:	Non-Ag Emp Share	Output per Worker	Avg Wages	Output	# Establish.	Population
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline						
Temperature Mismatch	0.037***	-0.089***	-0.011	-0.141	-0.144*	-0.108***
	(0.007)	(0.029)	(0.015)	(0.096)	(0.083)	(0.028)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.858	0.759	0.906	0.794	0.703	0.894
KP F-stat	590.4	415.6	415.6	415.8	415.6	590.4
Panel B: Predicted Migration Flows						
Temperature Mismatch	0.030***	-0.093***	-0.003	-0.140	-0.116	-0.096***
	(0.007)	(0.028)	(0.015)	(0.089)	(0.078)	(0.026)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.857	0.759	0.906	0.794	0.702	0.894
KP F-stat	502.5	375.6	375.6	375.7	375.6	502.5
Panel C: Long Differences (1920-1880)						
Temperature Mismatch	0.044***	-0.054	0.006	-0.035	-0.078	-0.085***
	(0.008)	(0.034)	(0.017)	(0.109)	(0.096)	(0.031)
Observations	2,303	2,265	2,265	2,266	2,265	2,304
R-squared	0.100	0.034	0.151	0.067	0.097	0.166
KP F-stat	426.0	408.8	408.8	409.1	408.8	426.0
Panel D: State×Decade FE						
Temperature Mismatch	0.018**	-0.101***	-0.032	-0.073	-0.230**	-0.128***
	(0.008)	(0.034)	(0.020)	(0.114)	(0.095)	(0.037)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.873	0.777	0.917	0.855	0.805	0.913
KP F-stat	280.0	244.7	244.7	244.5	244.7	280.0
Panel E: Trim Temperature Mismatch						
Temperature Mismatch	0.030***	-0.089***	-0.022	-0.151*	-0.110	-0.096***
	(0.006)	(0.027)	(0.014)	(0.091)	(0.078)	(0.027)
Observations	9,008	6,706	6,706	6,707	6,706	9,009
R-squared	0.859	0.761	0.908	0.795	0.703	0.896
KP F-stat	665.2	516.6	516.6	516.8	516.6	665.2
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
1880 Dep. Var. Mean	0.389	9,589	265.5	2,253,570	107.2	19,612

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880, for the period 1880–1920. The year 1890 is excluded, because it is not possible to construct the measure of climate mismatch. In columns 2–5, 1910 is also excluded, because data from the Census of Manufactures are missing for that year. Panel A reports the 2SLS estimates of Table 4. Panel B replaces the baseline instrument with a version that uses origin-specific migration flows predicted from weather shocks following Sequeira et al. (2020). Panel C reports 1920–1880 long-difference regressions. Panel D replaces census region-by-decade fixed effects with state-by-decade fixed effects. Panel E excludes counties with mismatch values above the 99th percentile or below the 1st percentile. All regressions include county fixed effects, census region by decade fixed effects (except for Panel D), and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.5. The Effects of Climate Mismatch on Manufacturing and Population: Robustness/2

Dep. Variable:	Non-Ag Emp Share	Log Manufacturing Output per Worker	Log Manufacturing Avg Wages	Log Manufacturing Output	Log Manufacturing # Establish.	Log Population
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline						
Temperature Mismatch	0.037***	-0.089***	-0.011	-0.141	-0.144*	-0.108***
	(0.007)	(0.029)	(0.015)	(0.096)	(0.083)	(0.028)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.858	0.759	0.906	0.794	0.703	0.894
KP F-stat	590.4	415.6	415.6	415.8	415.6	590.4
Panel B: 1880 Immigrant Share×Decade FE						
Temperature Mismatch	0.035***	-0.079***	-0.017	-0.126	-0.137*	-0.095***
	(0.007)	(0.028)	(0.015)	(0.095)	(0.082)	(0.027)
Observations	9,132	6,844	6,844	6,845	6,844	9,133
R-squared	0.862	0.762	0.907	0.797	0.707	0.900
KP F-stat	585.7	418.4	418.4	418.6	418.4	585.8
Panel C: 1880 Non-Ag Share×Decade FE						
Temperature Mismatch	0.035***	-0.089***	-0.010	-0.141	-0.146*	-0.106***
	(0.006)	(0.029)	(0.015)	(0.096)	(0.084)	(0.027)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.877	0.759	0.910	0.794	0.703	0.898
KP F-stat	593.0	417.5	417.5	417.7	417.5	593.1
Panel D: Frontier Experience×Decade FE						
Temperature Mismatch	0.036***	-0.090***	-0.007	-0.154	-0.153*	-0.108***
	(0.007)	(0.028)	(0.014)	(0.095)	(0.081)	(0.028)
Observations	9,116	6,838	6,838	6,839	6,838	9,117
R-squared	0.858	0.763	0.907	0.792	0.699	0.893
KP F-stat	591.0	429.2	429.2	429.5	429.2	591.1
Panel E: All Controls						
Temperature Mismatch	0.036***	-0.080***	-0.010	-0.132	-0.135*	-0.096***
	(0.007)	(0.028)	(0.014)	(0.094)	(0.080)	(0.027)
Observations	9,052	6,787	6,787	6,788	6,787	9,053
R-squared	0.879	0.766	0.911	0.795	0.704	0.901
KP F-stat	577.4	424.3	424.3	424.5	424.3	577.5
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
1880 Dep. Var. Mean	0.389	9,589	265.5	2,253,570	107.2	19,612

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880, for the period 1880–1920. The year 1890 is excluded, because it is not possible to construct the measure of climate mismatch. In columns 2–5, 1910 is also excluded, because data from the Census of Manufactures are missing for that year. Panel A reports the 2SLS estimates of Table 4. Panels B–D add controls for: (i) the 1880 immigrant population share (Panel B), (ii)the 1880 non-agricultural employment share (Panel C), and (iii) total frontier experience from Bazzi et al. (2020) (Panel D). Panel E includes all three controls jointly. All regressions include county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.6. The Effects of Climate Mismatch on Manufacturing and Population: Robustness/3

Dep. Variable:	Non-Ag Emp Share	Log Manufacturing: Output per Worker	Log Manufacturing: Avg Wages	Log Manufacturing: Output	Log Manufacturing: # Establish.	Log Population
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline						
Temperature Mismatch	0.037***	-0.089***	-0.011	-0.141	-0.144*	-0.108***
	(0.007)	(0.029)	(0.015)	(0.096)	(0.083)	(0.028)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.858	0.759	0.906	0.794	0.703	0.894
KP F-stat	590.4	415.6	415.6	415.8	415.6	590.4
Panel B: Predicted Precipitation Distance						
Temperature Mismatch	0.040***	-0.104***	-0.010	-0.217**	-0.226***	-0.136***
	(0.006)	(0.027)	(0.015)	(0.093)	(0.080)	(0.028)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.860	0.761	0.906	0.798	0.710	0.898
KP F-stat	623.5	452.6	452.6	452.8	452.6	623.5
Panel C: Predicted Geographic Distance						
Temperature Mismatch	0.036***	-0.089***	-0.011	-0.140	-0.144*	-0.108***
	(0.007)	(0.029)	(0.015)	(0.095)	(0.083)	(0.028)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.858	0.759	0.906	0.794	0.703	0.894
KP F-stat	590.8	416.1	416.1	416.4	416.1	590.9
Panel D: Both Distance Controls						
Temperature Mismatch	0.040***	-0.104***	-0.010	-0.217**	-0.226***	-0.137***
	(0.006)	(0.027)	(0.015)	(0.093)	(0.079)	(0.028)
Observations	9,196	6,895	6,895	6,896	6,895	9,197
R-squared	0.860	0.761	0.906	0.798	0.710	0.898
KP F-stat	623.9	453.3	453.3	453.5	453.3	623.9
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Region×Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
1880 Dep. Var. Mean	0.389	9,589	265.5	2,253,570	107.2	19,612

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880, for the period 1880–1920. The year 1890 is excluded, because it is not possible to construct the measure of climate mismatch. In columns 2–5, 1910 is also excluded, because data from the Census of Manufactures are missing for that year. Panel A reports the 2SLS estimates of Table 4. Panel B replicates Panel A controlling for predicted precipitation distance, constructed analogously to predicted temperature distance (see Section 3.1) and entered directly as a control (i.e., we do not instrument actual precipitation distance). Panel C replicates Panel A controlling for predicted geographic distance, constructed analogously to predicted temperature and precipitation distance. Panel D includes both controls simultaneously. All regressions include county fixed effects, census region by decade fixed effects, and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.7. Climate Mismatch and 1940 Income: Robustness

Dep. Variable:	All men		U.S.-born men		Immigrant men	
	Log Average Income	Log Hourly Earnings	Log Average Income	Log Hourly Earnings	Log Average Income	Log Hourly Earnings
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline						
Temperature Mismatch	-0.027***	-0.053***	-0.024***	-0.050***	-0.036***	-0.061***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Observations	2,890	2,867	2,890	2,867	2,752	2,707
R-squared	0.317	0.378	0.325	0.381	0.182	0.160
KP F-stat	644.98	646.36	644.98	646.36	569.77	560.23
Panel B: Predicted Migration Flows						
Temperature Mismatch	-0.028***	-0.054***	-0.024***	-0.051***	-0.036***	-0.061***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Observations	2,890	2,867	2,890	2,867	2,752	2,707
R-squared	0.317	0.379	0.325	0.382	0.182	0.160
KP F-stat	665.98	667.61	665.98	667.61	588.58	578.87
Panel C: State FE						
Temperature Mismatch	-0.035***	-0.041***	-0.026**	-0.033***	-0.101***	-0.093***
	(0.012)	(0.011)	(0.011)	(0.011)	(0.025)	(0.024)
Observations	2,889	2,866	2,889	2,866	2,751	2,706
R-squared	0.437	0.500	0.451	0.508	0.228	0.200
KP F-stat	137.63	141.34	137.63	141.34	123.00	124.24
Panel D: Trim Temperature Mismatch						
Temperature Mismatch	-0.033***	-0.060***	-0.029***	-0.056***	-0.038***	-0.064***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.011)	(0.010)
Observations	2,829	2,806	2,829	2,806	2,691	2,646
R-squared	0.318	0.376	0.326	0.379	0.184	0.159
KP F-stat	534.06	534.51	534.06	534.51	460.15	450.51
Panel E: Excluding Agricultural Workers						
Temperature Mismatch	-0.016***	-0.044***	-0.014***	-0.042***	-0.027***	-0.060***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)
Observations	2,890	2,867	2,890	2,867	2,727	2,672
R-squared	0.316	0.427	0.325	0.429	0.141	0.136
KP F-stat	644.98	646.36	644.98	646.36	556.05	543.56

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 for 1940. Panel A reports the 2SLS estimates of Table 5. Panel B replaces the baseline instrument with a version that uses origin-specific migration flows predicted from weather shocks following Sequeira et al. (2020). Panel C replaces census-region fixed effects with state fixed effects (with all other aspects of the specification unchanged). Panel D trims counties with mismatch values above the 99th percentile or below the 1st percentile. Panel E excludes farmers. All regressions include census region fixed effects (except for Panel C) and an indicator equal to one if the county was newly connected to the railroad network in the previous decade. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.8. Climate Mismatch and 1940 Income: Robustness/2

	All men		U.S.-born men		Immigrant men	
Dep. Variable:	Log Average Income	Log Hourly Earnings	Log Average Income	Log Hourly Earnings	Log Average Income	Log Hourly Earnings
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline						
Temperature Mismatch	-0.027***	-0.053***	-0.024***	-0.050***	-0.036***	-0.061***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Observations	2,890	2,867	2,890	2,867	2,752	2,707
R-squared	0.317	0.378	0.325	0.381	0.182	0.160
KP F-stat	644.98	646.36	644.98	646.36	569.77	560.23
Panel B: 1880 Immigrant Share						
Temperature Mismatch	-0.029***	-0.056***	-0.028***	-0.054***	-0.014	-0.051***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Observations	2,288	2,270	2,288	2,270	2,176	2,140
R-squared	0.399	0.434	0.413	0.440	0.158	0.135
KP F-stat	787.27	778.74	787.27	778.74	680.14	660.32
Panel C: 1880 Non-Ag Share						
Temperature Mismatch	-0.038***	-0.065***	-0.037***	-0.063***	-0.020**	-0.057***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.009)	(0.009)
Observations	2,299	2,281	2,299	2,281	2,184	2,148
R-squared	0.472	0.512	0.483	0.520	0.171	0.153
KP F-stat	700.24	691.59	700.24	691.59	597.50	577.79
Panel D: Frontier Experience						
Temperature Mismatch	-0.019***	-0.048***	-0.015***	-0.044***	-0.036***	-0.060***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Observations	2,845	2,822	2,845	2,822	2,710	2,666
R-squared	0.347	0.396	0.356	0.400	0.175	0.156
KP F-stat	591.50	590.85	591.50	590.85	534.69	525.25
Panel E: All Controls						
Temperature Mismatch	-0.034***	-0.063***	-0.032***	-0.060***	-0.020**	-0.057***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Observations	2,271	2,253	2,271	2,253	2,161	2,125
R-squared	0.484	0.513	0.497	0.520	0.173	0.158
KP F-stat	706.78	696.77	706.78	696.77	624.97	606.73

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 for 1940. Panel A reports the 2SLS estimates of Table 5. Panels B–D add controls for: (i) the 1880 immigrant population share (Panel B), (ii) the 1880 non-agricultural employment share (Panel C), and (iii) total frontier experience from Bazzi et al. (2020) (Panel D). Panel E includes all three controls jointly. All regressions include census region fixed effects and the year in which a county was first connected to the railroad. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.9. Climate Mismatch and 1940 Income: Robustness/3

	All men		U.S.-born men		Immigrant men	
Dep. Variable:	Log Average Income	Log Hourly Earnings	Log Average Income	Log Hourly Earnings	Log Average Income	Log Hourly Earnings
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Baseline						
Temperature Mismatch	-0.027***	-0.053***	-0.024***	-0.050***	-0.036***	-0.061***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)
Observations	2,890	2,867	2,890	2,867	2,752	2,707
R-squared	0.317	0.378	0.325	0.381	0.182	0.160
KP F-stat	644.98	646.36	644.98	646.36	569.77	560.23
Panel B: Predicted Precipitation Distance						
Temperature Mismatch	-0.021***	-0.050***	-0.016***	-0.046***	-0.056***	-0.075***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.011)	(0.010)
Observations	2,890	2,867	2,890	2,867	2,752	2,707
R-squared	0.324	0.382	0.334	0.386	0.195	0.163
KP F-stat	479.86	481.01	479.86	481.01	427.07	422.70
Panel C: Predicted Geographic Distance						
Temperature Mismatch	-0.033***	-0.055***	-0.031***	-0.053***	-0.020*	-0.045***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.012)	(0.012)
Observations	2,890	2,867	2,890	2,867	2,752	2,707
R-squared	0.317	0.379	0.325	0.381	0.184	0.162
KP F-stat	435.50	436.78	435.50	436.78	395.34	390.41
Panel D: Both Distance Controls						
Temperature Mismatch	-0.026***	-0.052***	-0.023***	-0.049***	-0.042***	-0.060***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.013)	(0.013)
Observations	2,890	2,867	2,890	2,867	2,752	2,707
R-squared	0.324	0.383	0.334	0.386	0.197	0.166
KP F-stat	315.69	317.43	315.69	317.43	284.52	283.42

Notes: The sample includes counties in the contiguous U.S. with positive population in 1880 for 1940. Panel A reports the 2SLS estimates of Table 5. Panel B replicates Panel A controlling for predicted precipitation distance, constructed analogously to predicted temperature distance (see Section 3.1). Panel C replicates Panel A controlling for predicted geographic distance, constructed analogously to predicted temperature and precipitation distance. Panel D includes both controls simultaneously. All regressions include census region fixed effects and the year in which a county was first connected to the railroad. The KP F-statistic refers to the Kleibergen–Paap test for weak instruments. Standard errors are clustered at the county level and further adjusted using the baseline block bootstrap procedure. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

E A Simple Two-Sector Model

This appendix presents a highly stylized model that formalizes the mechanisms discussed in Section 3.

Environment. A county is endowed with a fixed labor force $\bar{L}$, allocated between agriculture (A) and manufacturing (M):

$$L_A + L_M = \bar{L}.$$

Let δ denote the average *climate mismatch* (i.e., the absolute difference in temperature) between local conditions and the climates of resident workers. Mismatch weakly reduces effective labor input in sector $s \in \{A, M\}$ through $m_s(\delta)$, where $m_s(\delta) \leq 0$.

Technology. Agriculture uses land T and manufacturing uses capital K. Output in each sector is Cobb–Douglas:

$$Y_A = A_A \, [m_A(\delta) L_A]^{\alpha} T^{1-\alpha}, \tag{8}$$

$$Y_M = A_M \, g(Y_A) \, [m_M(\delta) L_M]^{\beta} K^{1-\beta}. \tag{9}$$

The term $g(Y_A)$ captures input–output linkages: higher agricultural output boosts manufacturing productivity such that $g'(Y_A) > 0$.

Sectoral wages equal marginal products:

$$w_A = \alpha A_A m_A(\delta)^{\alpha} L_A^{\alpha-1} T^{1-\alpha}, \tag{10}$$

$$w_M = \beta A_M g(Y_A) m_M(\delta)^{\beta} L_M^{\beta-1} K^{1-\beta}. \tag{11}$$

Labor is perfectly mobile across sectors, so wages equal marginal products in equilibrium: $w_A = w_M = w$.

E.1 Labor Reallocation

To understand how climate mismatch affects the allocation of labor across sectors, note that equilibrium requires equality of wages in agriculture and manufacturing. Let

$$F(L_A, \delta) \equiv w_A(L_A, \delta) - w_M(L_A, \delta)$$

denote the wage gap. The equilibrium labor allocation solves $F(L_A, \delta) = 0$. Differentiating this condition with respect to d gives

$$\frac{dL_A}{d\delta} = -\frac{F_\delta}{F_{L_A}}.$$

Since $F_{L_A} < 0$ (marginal products decline in own labor), the sign of $\frac{dL_A}{d\delta}$ depends entirely on F_δ—that is, on how mismatch shifts the relative marginal products of labor in the two sectors.

Differentiating F with respect to δ gives

$$\begin{aligned} F_\delta = \alpha A_A \left(\frac{m_A(\delta)L_A}{T}\right)^{\alpha-1} m'_A(\delta) - \beta A_M g(Y_A) \left(\frac{m_M(\delta)L_M}{K}\right)^{\beta-1} m'_M(\delta) \\ - A_M m_M(\delta)^\beta \left(\frac{L_M}{K}\right)^{\beta-1} g'(Y_A)\frac{\partial Y_A}{\partial \delta} \end{aligned} \tag{12}$$

Three forces shape the sign of F_δ. First, mismatch directly reduces agricultural labor productivity. Because agriculture relies heavily on climate-specific human capital, even modest mismatch can lower the marginal product of labor in this sector. This tends to lower the agricultural wage relative to the manufacturing wage, pushing labor out of agriculture.

Second, mismatch may reduce productivity in manufacturing, though the degree of sensitivity may differ across sectors. If manufacturing is also climate-sensitive, mismatch lowers its marginal product of labor as well, raising the relative attractiveness of agriculture. This pushes labor in the opposite direction.

Third, input–output linkages propagate agricultural shocks into manufacturing. Because local manufacturing productivity depends on agricultural output through $g(Y_A)$, any decline in Y_A induced by mismatch (see Section E.2) reduces manufacturing productivity. This indirect channel again raises agricultural wages relative to manufacturing wages, counteracting the initial outflow of labor from agriculture.

Taken together, these forces imply that the effect of mismatch on labor allocation is theoretically ambiguous. In the empirically relevant case we emphasize in the main text, agriculture is sufficiently more climate-sensitive than manufacturing and input–output linkages are not strong enough to overturn this:

$$|m'_A(\delta)| \text{ large relative to } |m'_M(\delta)| \quad \text{and} \quad g'(Y_A) \text{ moderate.}$$

Under this parametric restriction,

$$F_\delta < 0 \quad \Rightarrow \quad \frac{dL_A}{d\delta} < 0, \qquad \frac{dL_M}{d\delta} > 0,$$

so climate mismatch reallocates labor from agriculture to manufacturing.

E.2 Output

Agriculture. Agricultural output depends on effective labor input $m_A(\delta)L_A$. Differentiating $\ln Y_A$ yields

$$\frac{d\ln Y_A}{d\delta} = \alpha\frac{m'_A(\delta)}{m_A(\delta)} + \alpha\frac{d\ln L_A}{d\delta}.$$

The first term reflects the direct effect of mismatch on agricultural productivity, which is strongly negative given the high climate sensitivity of the sector. The second term captures labor reallocation. Under our maintained assumption that the direct productivity loss dominates labor-adjustment forces, agricultural output declines with mismatch. Thus,

$$\frac{dY_A}{d\delta} < 0.$$

Manufacturing. In addition to effective labor, manufacturing output depends on agricultural output through the linkage $g(Y_A)$. Differentiating $\ln Y_M$ gives

$$\frac{d\ln Y_M}{d\delta} = \beta\frac{m'_M(\delta)}{m_M(\delta)} + \beta\frac{d\ln L_M}{d\delta} + \frac{g'(Y_A)}{g(Y_A)}\frac{dY_A}{d\delta}.$$

The first term reflects direct climate sensitivity in manufacturing, which is small but weakly negative. The second term captures the inflow of labor into manufacturing, which pushes output upward. The third term captures the spillover from agriculture and is negative because $g'(Y_A) > 0$ and we have established $\frac{dY_A}{d\delta} < 0$. Hence, the net effect of mismatch on manufacturing output is ambiguous.

E.3 Wages

Mismatch affects wages through its impact on effective labor productivity and through the reallocation of workers across sectors. Given the comparative static results above—namely, that mismatch reduces agricultural productivity, draws labor out of agriculture, and induces a corresponding inflow into manufacturing—we can characterize the wage responses in each sector.

Agriculture. Agricultural wages depend on both the effective productivity of agricultural workers and the amount of labor employed in the sector. Differentiating the expression for w_A

yields

$$\frac{\partial \ln w_A}{\partial \delta} = \alpha \frac{m_A'(\delta)}{m_A(\delta)} + (\alpha - 1) \frac{d \ln L_A}{d\delta}.$$

The first term is strictly negative because agricultural productivity declines with mismatch. The second term is positive under our maintained assumption that mismatch reduces agricultural employment (i.e., $dL_A/d\delta < 0$), which raises marginal products in a concave production technology. These two forces work in opposite directions; as a result, the net effect on agricultural wages is ambiguous.

Manufacturing. Manufacturing wages respond to mismatch through three channels:

$$\frac{d \ln w_M}{d\delta} = \beta \frac{m_M'(\delta)}{m_M(\delta)} + \frac{g'(Y_A)}{g(Y_A)} \frac{dY_A}{d\delta} + (\beta - 1) \frac{d \ln L_M}{d\delta}. \tag{13}$$

The first term reflects the direct climate sensitivity of manufacturing and is weakly negative. The second term captures input–output spillovers: because $g'(Y_A) > 0$ and we have established $\frac{dY_A}{d\delta} < 0$, agricultural decline reduces manufacturing productivity, making this term negative as well. The final term is also negative: mismatch induces an inflow of labor into manufacturing ($dL_M/d\delta > 0$), and with $\beta < 1$, this lowers marginal products. Under the parametric restrictions used throughout this appendix, all three forces move manufacturing wages in the same direction. Hence,

$$\frac{dw_M}{d\delta} < 0.$$

E.4 Additional predictions

Farmland. Let p_A denote the price of agricultural output and let r_T be the rental rate of land. The first-order condition for land in the agricultural sector implies

$$r_T = (1 - \alpha)\, p_A A_A (m_A(\delta) L_A)^{\alpha} T^{-\alpha}.$$

Solving for land demand yields

$$T(\delta) = \left(\frac{(1 - \alpha) p_A A_A}{r_T} \right)^{1/\alpha} m_A(\delta)\, L_A(\delta). \tag{14}$$

The effect of mismatch on farmland use therefore depends on how mismatch alters both effective labor productivity and the amount of labor employed in agriculture. Under the assumption that mismatch reduces agricultural productivity and induces an outflow of labor from agriculture—both $m_A(\delta)$ and $L_A(\delta)$ decline with mismatch, farmland demand falls as well:

$$\frac{dT}{d\delta} < 0.$$