Climate Matching, European Overseas Settlements, and Long-Run Development*

Alessandra Nicolo', Marco Tabellini, and Charles A. Taylor†

March 2026

Abstract

We test Alfred Crosby's *Ecological Imperialism* hypothesis that ecological compatibility shaped European overseas settlement, and examine its consequences for long-run economic development. We combine genealogical microdata on more than 300,000 Europeans born between 1500 and 1850 who died outside Europe with high-resolution climate data to measure climatic similarity between origins and destinations. We first show that climate matching strongly predicts migration: increasing temperature distance reduces the number of settlers. We then document that climate mismatch affected settlers' demographic viability: migrants to less similar environments lived shorter lives and had fewer children, with particularly large penalties in hotter destinations. Turning to long-run outcomes, we find that locations settled by more climate-matched Europeans exhibit larger shares of European ancestry today. These differences extend to economic development: locations where European settlers were better matched climatically exhibit higher levels of nighttime luminosity, even within subnational regions.

Keywords: Climate matching, European overseas settlement, long-run economic growth.
JEL Codes: N0, O1, O40, O43.

*Michael Cleary, Gunnhildur Hallgrimsdottir, Jacki Huang, and Gabrielle Levy provided outstanding research assistance. All remaining errors are ours.

†Nicolo': Bocconi University (alessandra.nicolo@studbocconi.it); Tabellini: Harvard Business School, NBER, CEPR, RF Berlin, and LISER (mtabellini@hbs.edu); Taylor: Harvard Kennedy School, Harvard University, and NBER (ctaylor@hks.harvard.edu).

1 Introduction

In influential work, Alfred Crosby argued that permanent European settlement depended on ecological compatibility between Europe and its overseas territories (Crosby, 1986). European plants, livestock, pathogens, and people—the so-called "portmanteau biota"—thrived in temperate regions that resembled Europe, allowing settlers to reproduce familiar production systems and sustain demographic growth. In environments that differed markedly from European conditions—whether hotter or colder—European crops often failed, livestock struggled to adapt, and settlers faced higher mortality from unfamiliar diseases and environmental stressors. Demographic viability—the ability of Europeans to survive, reproduce, and sustain settlement over generations—thus constituted a central constraint on colonial expansion. Where ecological conditions permitted demographic self-replication, durable settler societies emerged; where they did not, European presence remained sparse and extractive.

In this paper, we bring Crosby's hypothesis to the data and extend it to examine the implications of climatic similarity between European settlers' origins and destinations for long-run economic development. We measure European overseas settlement using population-scale genealogical records from Kaplanis et al. (2018), which report individuals' birth and death locations and dates and allow us to trace long-distance moves of Europeans born between 1500 and 1850 who died outside Europe. Throughout the analysis, we define origins as European birthplaces and destinations as non-European locations identified as colonies in Acemoglu et al. (2001), in which at least one European death is recorded. To the best of our knowledge, this is the first paper to provide systematic, individual-level evidence on Crosby's hypothesis and to trace its implications from migration patterns through demographic outcomes to long-run economic development.

Because the data from Kaplanis et al. (2018) are constructed from user-submitted family trees, they do not capture the full universe of European settlers. We therefore assess external validity in two ways. First, aggregating our data to the country-of-destination level, we show that settlement patterns closely track independent historical estimates of European population shares between 1500 and 1900 (Easterly and Levine, 2016). Second, we compare the migration patterns of Spanish migrants in our data to newly compiled archival records on Spanish transatlantic migration for 1492–1830 (Arroyo Abad and Espín-Sánchez, 2025), and document similar timing and regional origins.

We then link these genealogical records to high-resolution climate data on temperature, precipitation, and other characteristics. To harmonize migration and ecological data, we aggregate origins and destinations to $0.5° \times 0.5°$ grid cells. We study the relationship be-

tween climate at origin and destination by estimating a gravity model of origin–destination migration flows. Following the gravity framework widely used in trade and, more recently, in migration research (Anderson and Van Wincoop, 2003; Obolensky et al., 2024), we relate the number of individuals born in a European origin cell in a given 50-year period and dying in an overseas destination cell to geographic and ecological distance between the two locations. We estimate the model using Poisson pseudo–maximum likelihood (PPML), which accommodates the large number of zero flows and the potential heteroskedasticity in the migration matrix. Building on Obolensky et al. (2024), our baseline measure of ecological distance is the absolute difference in temperature and precipitation between origin and destination cells. We control for geographic distance and its square to capture transportation costs and other spatial frictions that may vary nonlinearly with distance, such as the fixed costs of long oceanic crossings and inland travel.

The preferred specification includes country-of-origin-by-period and destination-cell-by-period fixed effects. The former absorb time-varying push factors at the origin, such as population growth, wars, and economic crises; the latter absorb all time-varying pull factors at the destination, including changes in labor demand, colonial policies, and local mortality environments. Identification therefore comes from variation in ecological similarity across origin-destination pairs after accounting for shocks at both ends of the migration corridor. This stringent structure also mitigates concerns about selection in the genealogical records, since bias would require climate-correlated reporting differences within origin-destination-period groups—a form of selection that seems unlikely in practice.

We find that permanent European settlements were more likely to arise in destinations that are climatically more similar to the origin. According to our estimates, a 5°C increase in temperature distance—comparable to moving from London to Buenos Aires or from Madrid to Johannesburg—reduces migration flows by roughly 56%. This relationship likely reflects the combined influence of complementary mechanisms: climate distance may have discouraged migration, reduced the likelihood that settlers remained in poorly matched environments, and lowered survival in such locations. Decomposing temperature distance by direction reveals a marked asymmetry: the negative effect is substantially larger and more precisely estimated when destinations are hotter than the origin than when they are colder. Consistent with evidence on climate matching in migration in the post-industrial US context (Obolensky et al., 2024), the effects for precipitation distance are instead weak and unstable.

These results are robust to alternative measures of geographic distance and travel costs, sample restrictions, and the exclusion of canonical settler societies such as the Neo-Europes (Easterly and Levine, 2016). They are also robust to including origin-country by destination-

country fixed effects. These fixed effects absorb all time-invariant factors specific to each country pair—such as historical ties, migration networks, language, and institutional linkages. Because migration flows are measured at the origin–destination cell level, identification in this specification comes from cross-sectional variation within country pairs across origin and destination locations. In practice, this amounts to comparing flows from different regions within the same origin country (e.g., Madrid versus Seville) to different locations within the same destination country (e.g., Buenos Aires versus other cells in Argentina), holding constant all factors common to the Spain–Argentina pair. The coefficient on temperature distance remains negative and statistically significant.

We also examine whether the relationship between climate similarity and migration extends beyond temperature to other dimensions of ecological similarity. Crosby (1986)'s hypothesis extends beyond temperature, which we treat as a summary measure of climatic similarity. We therefore incorporate a broader set of ecological controls, including a soil similarity index, absolute latitude, coastal access, elevation, terrain ruggedness, and additional temperature and precipitation moments such as minimum, maximum, and seasonal range. The coefficient on temperature distance remains large and precisely estimated across all specifications. At the same time, other ecological dimensions, including latitude and elevation, are independently predictive of migration flows.

We then leverage information on age at death and, for a subset of records, the number of children born overseas. We estimate individual-level OLS regressions relating settlers' age at death and completed fertility to the temperature distance between origin and destination. As in the migration analysis, we control for a quadratic in geographic distance, as well as origin-by-period and destination-by-period fixed effects. We find that settlers who moved to climates farther from their origin lived shorter lives and had fewer children. According to our estimates, a 5°C increase in temperature distance is associated with approximately 310 fewer days of life (1.26% relative to the sample mean) and about 0.16 fewer children (4.56% relative to the sample mean). The effect of precipitation distance is again weak or unstable. Decomposing temperature distance by direction, demographic penalties are largest for moves to hotter destinations but also present at extreme cold, producing an inverted U-shaped pattern. This contrasts with the migration results, where the deterrent effect is more pronounced for moves to hotter environments, suggesting that climate mismatch affected both selection into destinations and outcomes conditional on arrival.

These estimates should be interpreted in light of several considerations. First, for mortality, the effect of climate mismatch becomes small and statistically indistinguishable from zero once destination-cell fixed effects are included, indicating sensitivity to the level of spatial

aggregation. Our preferred specification for this analysis thus absorbs broad local shocks at the subnational level, and results remain robust when controlling for destination characteristics associated with settler mortality and long-run development, such as malaria ecology, absolute latitude, and pre-colonial population density. At the same time, it does not condition on the most granular destination-by-period effects. Second, the demographic results may partly reflect selection into climate matching—for example, if migrants who settled in climatically similar destinations were systematically healthier or wealthier. Finally, fertility is observed only for a subset of individuals in the genealogical records. For these cases, we cannot distinguish between genuinely childless settlers and instances of non-reporting. If reporting quality were correlated with climate distance, fertility estimates could be biased. Taken together, however, the patterns are consistent with Crosby's central insight. They suggest that ecological compatibility influenced both the geography of settlement and the demographic viability of European populations overseas.

Having established that climate matching shaped the geography of European settlement during the colonial period, we turn to its long-run consequences. If ecological compatibility facilitated durable settler societies, then territories settled by Europeans who were better matched climatically should exhibit greater persistence of European populations. To test this implication, we use country-level estimates of the European population share in former colonies as of 2000 from Putterman and Weil (2010). We find that countries where European settlers were more closely matched to local climates have larger European population shares today. A 5°C reduction in temperature mismatch among settlers between 1500 and 1850 is associated with an increase of 9 percentage points (about 50% of the mean) in the European population share in 2000.

Although this cross-sectional evidence is necessarily suggestive, the relationship is robust to controlling for country-level characteristics commonly emphasized in the literature, including malaria ecology, absolute latitude, average temperature, precipitation, and other standard geographic controls, and to excluding the Neo-Europes or other potentially distinctive destinations such as African countries. One possible interpretation is that lower climate distance led to larger initial European settlements, which then persisted over time. However, the results remain robust when controlling directly for the number of European settlers between 1500 and 1850, suggesting that climate matching influenced not only the initial scale of settlement but also its long-run persistence.

We then examine whether these patterns extend to economic development. Following the literature (Michalopoulos and Papaioannou, 2013, 2014; Alesina et al., 2016; Henderson et al., 2018; Chiovelli et al., 2026), we proxy for long-run economic performance using satellite-

based nighttime luminosity. This approach allows us to measure economic activity at a fine spatial scale within countries and to exploit variation across locations while holding constant country-level factors. We estimate OLS regressions relating present-day luminosity in each $0.5° \times 0.5°$ destination cell to the average climate distance between European settlers' origins and that location. The specification includes subnational region fixed effects, which absorb broad institutional and colonial differences within countries, as well as a rich set of controls associated with long-run development and colonial expansion (e.g., absolute latitude, malaria ecology, elevation, terrain ruggedness, soil suitability, distance to coast, and population density in 1500).

Identification therefore comes from comparing locations within the same subnational region (equivalent to a U.S. state) that differ in the extent to which European settlers were climatically matched to local conditions. Because the climate distance measure is constructed using migrant shares, this variation reflects differences in the composition of settlers rather than the overall scale of European settlement. Two considerations are important for interpretation. First, the estimates may partly reflect selection into climate matching: if more skilled or wealthier migrants were better able to settle in climatically similar environments, the results may capture differences in the composition of settlers rather than the causal effect of climate compatibility per se. Second, despite the inclusion of detailed controls and subnational fixed effects, omitted local characteristics may still bias the estimates if locations with greater climate mismatch also differ systematically along other dimensions that affect long-run development.

We find that a 5°C increase in temperature distance among settlers is associated with a 25% reduction in present-day luminosity. One interpretation is that ecological compatibility influenced the types of institutions that emerged, consistent with the institutional view of colonial development (Acemoglu et al., 2001). Locations where European settlers could more easily adapt may have supported more inclusive institutional arrangements, whereas areas with greater mismatch may have favored more extractive systems. Consistent with this interpretation, the coefficient declines by about 40% when we include country fixed effects, suggesting that part of the relationship operates through cross-country institutional differences.

To examine the institutional channel more directly, we estimate cross-country regressions relating temperature mismatch to contemporary measures of institutional quality. Countries with greater climate mismatch exhibit lower democracy scores and weaker constraints on the executive. While these patterns are only suggestive, they are consistent with the institutional mechanism. At the same time, the relationship between climate distance and development

remains economically large and statistically significant even when we include subnational fixed effects, which absorb national institutions. This indicates that the results cannot be explained solely by cross-country institutional variation, and are also consistent with differences in local institutional environments (Dell, 2010; Michalopoulos and Papaioannou, 2013, 2014).

A second channel is that climatic similarity may have facilitated the transfer and adoption of European technologies and production systems. Relatedly, climatic similarity may have encouraged trade and investment flows by lowering the costs of operating in similar environments. Third, by enabling larger and more persistent settler populations, climate matching may have generated scale effects and agglomeration economies. While we do not disentangle these channels, the evidence is consistent with all three mechanisms.

Together, our findings highlight the importance of ecological compatibility in shaping both European settlement patterns and their long-run consequences. They complement existing work on the determinants of European settlement that emphasizes violence (Lindqvist, 1992; Stannard, 1992; Elkins, 2022), military and technological advantages (Thomas, 1993), capitalist expansion (Wallerstein, 1974; Pomeranz, 2000; Elliott, 2006), and deep geographic and biogeographic determinants of European dominance (Diamond, 1997). They also relate to the large literature on the long-run consequences of colonial expansion and the persistence of economic development.[1] This literature has linked settler mortality to patterns of settlement and institutional development (Acemoglu et al., 2001), documented the emergence of extractive institutions and the "reversal of fortune" (Acemoglu et al., 2002), the persistence of local economic activity and locational advantages (Maloney and Valencia Caicedo, 2016), and the long-run effects of specific institutional arrangements and colonial interventions (Banerjee and Iyer, 2005; Nunn and Wantchekon, 2011; Michalopoulos and Papaioannou, 2016; Dell and Olken, 2020). We contribute to this literature by bringing Crosby (1986)'s hypothesis to the data, identifying ecological compatibility as a deeper geographic determinant of both settlement patterns and long-run outcomes. While Acemoglu et al. (2001) emphasize settler mortality as a key driver of cross-country institutional divergence, our results show that climatic similarity shaped the broader demographic viability of European settlement and that this ecological channel has persistent effects on development even within subnational regions—a pattern that cannot be explained by national institutions alone.

Our paper also relates to Feyrer and Sacerdote (2009) and Easterly and Levine (2016), who show that the length of exposure to colonialism and the European share of the population

[1] See Fenske et al. (2025) for a recent review of this literature.

during colonization, respectively, are positively associated with economic development today. We complement these works by showing that it is not only the extent of European settlement or the length of colonial rule that matters, but also the degree of climatic compatibility between settlers and their destinations. Climatic compatibility may have influenced the types of institutions and activities established by colonizers, thereby shaping long-run development (Engerman and Sokoloff, 2000; Acemoglu et al., 2001; Bruhn and Gallego, 2012).

2 Historical Background

2.1 European Overseas Expansion and Settlement

From the late 15th to the late 20th century, colonialism transformed global demographics, economies, and ecologies, leaving legacies that continue to shape the contemporary world. Osterhammel (1997) defines colonialism as "a relationship between an indigenous (or forcibly imported) majority and a minority of foreign invaders. The fundamental decisions affecting the lives of the colonized people are made and implemented by the colonial rulers in pursuit of interests that are often defined in a distant metropolis." As noted by Manjapra (2020), this asymmetrical relationship of domination was not unique to the modern era. Earlier empires—such as the Roman, Mongol, and Ottoman—also ruled over conquered territories far from their political centers. What distinguished the modern era, however, was the combination of maritime technology and oceanic expansion. Beginning in the 15th century, European powers established colonies separated from their metropoles by vast oceans, creating for the first time a truly global system that integrated the Americas, Africa, Asia, and Oceania.

European overseas expansion unfolded gradually from the late 15th century onward. Iberian powers first established Atlantic outposts and, by the early 16th century, conquered large territories in the Americas (Dunn and Kelley, 1989; Subrahmanyam, 1997). From the 17th century, northern European powers—the Dutch, English, and French— entered the imperial arena, creating trading posts, plantation colonies, and settler outposts across the Atlantic, Indian Ocean, and Asia (Osterhammel, 1997). In Africa, large-scale territorial control expanded later, particularly during the late nineteenth-century "Scramble for Africa," when European powers partitioned most of the continent into formal colonies (Michalopoulos and Papaioannou, 2014, 2016, 2020). By the 18th and 19th centuries, European empires spanned the globe, though their effective control and demographic presence remained highly uneven across regions.

Forms of colonial rule varied widely, reflecting both metropolitan strategy and local ecological and demographic conditions (Osterhammel, 1997). In some territories—such as North America, Australia, New Zealand, and temperate regions of South America, particularly Argentina and Uruguay—Europeans migrated in large numbers and established enduring settler societies (Acemoglu et al., 2001). In others, particularly in the Caribbean and Brazil, European elites developed plantation economies reliant on enslaved African labor (Engerman and Sokoloff, 2000, 2002). Still elsewhere, colonial activity centered on mineral extraction, as in Potosí, or on strategic ports and trading hubs such as Goa, Batavia in Java, and Singapore, where European presence was often commercial and administrative rather than demographic (Dell, 2010; Dell and Olken, 2020). Similar patterns characterized much of tropical Africa and South Asia, where European control often extended over vast territories without large-scale settler migration (Easterly and Levine, 2016).

Institutional arrangements mirrored these differences. In "extractive states," colonial rulers imposed few checks on executive power and directed resources to the metropole, often through coerced or slave labor. By contrast, where Europeans settled in large numbers, they established institutions that protected property rights and constrained executive authority—institutions that were inclusive primarily for European settlers. Across both extractive and settler settings, however, colonial rule entailed profound disruption for indigenous populations through coercion, displacement, extermination, and political exclusion. Acemoglu et al. (2001, 2002) emphasize settler mortality and pre-colonial population density as key determinants of this divergence and its long-run consequences. A complementary perspective stresses factor endowments and geographic conditions (Diamond, 1997; Engerman and Sokoloff, 1997, 2002): where climate and soils favored plantation crops or mineral extraction, production relied on coerced labor and entrenched inequality; where smallholder farming was viable, more egalitarian institutions emerged.

2.2 Crosby's Hypothesis

While the perspectives described earlier emphasize disease environments, factor endowments, or institutions, they do not directly examine how ecological compatibility between Europe and overseas territories shaped the scale and sustainability of European settlement. In *Ecological Imperialism*, Alfred W. Crosby advanced precisely this argument. He contended that the success of European colonization depended on whether local environments could sustain the transfer and reproduction of European organisms (Crosby, 1986). As Crosby put it, "the prospective settlement had to be placed where the land and climate were similar

to those in some part of Europe," and sufficiently removed from the Old World to limit exposure to predators and pathogens adapted to European species (Crosby, 1986, p.102).

Central to Crosby's framework is the notion of "Neo-Europes:" geographically distant territories united by temperate climates and, ultimately, by majority-European populations. These lands shared ecological conditions that allowed European settlers to replicate familiar crops and livestock. Even within these broader temperate zones, settlement clustered where local conditions most closely resembled those of Europe—the eastern seaboard of North America, southeastern Australia, much of New Zealand, and the Río de la Plata grasslands of South America.

At its core, Crosby's argument implies that climatic similarity between origin and destination should predict the intensity of European migration and settlement. The mechanism operated through what he termed the *portmanteau biota*: the bundle of crops, livestock, weeds, and pathogens that Europeans carried abroad. In temperate settings, European grains and animals flourished and often displaced native species. Mortality shocks among native populations, combined with the ecological dominance of European species, tilted the balance decisively in favor of settlers.[2] In tropical environments, by contrast, European crops frequently failed, livestock succumbed to local diseases, and settlers faced high mortality from malaria, yellow fever, and other pathogens.

Crosby's argument resonates with contemporary records stressing the relevance of environmental fit. The *Relaciones Geográficas* (1577–1586), a systematic survey commissioned by the Spanish crown, asked local officials to document climate, soils, flora and fauna, and the healthfulness of their jurisdictions. Climate and disease environments were central to assessments of colonial viability: questions inquired whether provinces were "cold or hot, dry or damp," whether towns were "healthful or unhealthful," and how well European animals bred in local conditions. Similar concerns appear in English sources. In 1622, Sir Ferdinando Gorges described New England as "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" (Kupperman, 1984). A few decades later, Thomas Newe observed from Carolina that settlers typically endured a "seasoning" but that mortality was limited (Newe, 1911). Such testimonies reveal a shared awareness across empires that climate compatibility shaped the sustainability of colonial

[2] A large epidemiological-historical literature argues that Euroasian populations evolved and (maintained) crowd diseases because of higher population densities sustained urbanization, and close contact with domesticated animals, yielding pathogens capable of large outbreaks (smallpox, measles, influenza) to which many indigenous populations outside the Old World had limited prior exposure (McNeill, 1976; Crosby, 1986; Diamond, 1997). This asymmetry in exposure and immunity helps explain why pathogen transmission after contact produced catastrophic mortality among Native populations, reinforcing the "one-sided" ecological advantage emphasized by Crosby.

projects, anticipating Crosby's later synthesis.

While Crosby (1986) focused primarily on the ecological and demographic dynamics of European expansion, his framework carries implications that extend beyond settlement patterns. If ecological compatibility determined where Europeans could survive and reproduce at scale, it likely shaped not only the geography of migration but also the subsequent institutional and economic trajectories of those territories. Regions where climate matching facilitated large and persistent European populations may have developed different political and economic structures than regions where European presence remained thin or ecologically constrained. In the empirical analysis that follows, we test both dimensions of this argument. First, we evaluate Crosby's central insight that climatic similarity between European origins and overseas destinations predicts the scale and demographic viability of settlement. Second, we examine whether ecological matching is associated with differences in contemporary economic development, consistent with the idea that settlement feasibility influenced long-run trajectories.

3 Data and Measurement

This section describes the main datasets used in the analysis. Section 3.1 presents the genealogical data, Section 3.2 the climate data, and Section 3.3 nighttime luminosity. Additional data sources are described in Section 3.4.

3.1 Genealogy Data

Measuring European overseas settlement. To measure historical overseas migration of European settlers, we rely on the FamiLinx database. FamiLinx is an online genealogical resource that records individual-level demographic information from user-submitted family trees hosted on major genealogy platforms, harmonizing details on birth and death dates and locations, sex, and kinship ties over several centuries (Colasurdo and Omenti, 2024). The dataset was first assembled by Kaplanis et al. (2018) and has since been extended and applied in subsequent work, including Blanc and Wacziarg (2025). It links together population-scale family trees covering approximately 13 million individuals, compiled from 86 million user-contributed profiles on Geni.com.

The sample is restricted to 328,067 individuals born in Europe between 1500–1850 who died in non-European territories classified as colonies in Acemoglu et al. (2001). This restriction

isolates cases of long-distance, permanent migration during the early modern colonial period and excludes temporary sojourns or return movements.[3] Figure 1, Panel A, displays the distribution of European settlers in our sample across destinations. Settlement is highly concentrated in North America—particularly the United States—as well as in temperate regions of South America, Australia, and New Zealand. By contrast, European presence is more limited and spatially dispersed across Africa and South and Southeast Asia, consistent with lower levels of permanent settlement in these regions. Table A.1 reports the distribution of individuals by European country of origin. The bulk of observations come from a small number of sending countries, most notably Great Britain, France, Germany, Ireland, and the Netherlands, which together account for the vast majority of cases. Scandinavian countries also contribute a non-trivial number of migrants. While Scandinavian overseas expansion during the early modern period was limited, these flows reflect both Denmark–Norway's colonial possessions in the Caribbean and, more importantly, the large-scale 19th-century emigration from Sweden and Norway to North America and other settler destinations (Abramitzky and Boustan, 2017). Southern European countries such as Spain and Portugal are also represented, though in smaller numbers, while origins from the rest of Europe are relatively rare.

Figure A.1 displays settlement patterns for four major colonial European origins: Spain (Panel A), Portugal (Panel B), Great Britain (Panel C), and the Netherlands (Panel D). Migration flows are substantially larger for Great Britain and the Netherlands, as indicated by the thicker and more numerous connections, while flows from Spain and Portugal are smaller in our data. There is also marked variation across empires in the geography of settlement. Spanish and Portuguese migrants are concentrated in Latin America, reflecting early colonial expansion in the Americas, whereas British and Dutch settlers are more widely dispersed, with flows extending to North America, Africa, Asia, and Oceania. Notably, even within each empire there is considerable heterogeneity in destination choices: migrants from the same origin country move to a wide range of locations across continents. This within-origin variation is central to our empirical strategy, as it allows us to compare settlement patterns across destinations that differ in ecological characteristics while holding origin-specific factors constant.

Data validation and potential limitations. Because FamiLinx does not capture the full universe of European settlers, we assess how closely the observed settlement patterns align with independent historical evidence on European presence overseas. First, we aggregate

[3] Europe is defined as the intersection of a bounding box (−10° to 40° longitude, 30° to 71° latitude) with the Natural Earth "Europe" polygons (WGS84). Individuals are classified as European-born if their origin lies within this area. Country boundaries are harmonized to contemporary Natural Earth definitions.

our sample to the country-of-destination level and correlate the log number of settlers in our data with the log European population during the colonial period as estimated by Easterly and Levine (2016). Figure A.2 plots the resulting relationship and reveals a strong positive correlation between the two independent sources.[4] The relationship remains robust when excluding the four countries with especially large European populations—United States, Canada, New Zealand, and Australia (dashed line).

Second, we focus on migrants born in Spain and compare our counts to the comprehensive archival estimates in Arroyo Abad and Espín-Sánchez (2025).[5] Figure A.3 presents this comparison. Panel A displays the number of Spanish-born individuals in our data by year of birth, while Panel B maps their regional origins within Spain. Panels C and D report the corresponding figures from Arroyo Abad and Espín-Sánchez (2025). The temporal dynamics and regional composition align closely across datasets. The peak in migration appears slightly later in the archival data (Panel C) than in our series (Panel A), as expected: their data record year of departure, whereas ours record year of birth.

One remaining caveat to the Familinx data is that missingness is selective. Individuals with reported demographic information are more likely to have other variables recorded, and completeness tends to cluster within families. Compared to official registers, FamiLinx population display systematically higher survival, with underrepresentation of women and individuals who died young, leading to biased age-sex distributions (Colasurdo and Omenti, 2024). Mortality patterns derived from FamiLinx, for example, are lower than those in official registers, particularly for earlier cohorts. While such selection affects representativeness and thus external validity, it would threaten identification only if missingness were systematically correlated with climatic distance between European origins and overseas destinations.

[4] Specifically, for each country we compute the average European population across the benchmark years reported in Easterly and Levine (2016) for the period 1500–1900. There are no former colonies for which Easterly and Levine (2016) report a strictly positive European population between 1500 and 1900 but for which we observe no genealogical migrants. Conversely, we identify 50 former colonies for which we observe at least one migrant in our data but for which Easterly and Levine (2016) either lack data or record a zero European population over the same period. Three of these countries—Indonesia (566 migrants), India (528 migrants), and Uruguay (393 migrants)—account for the largest shares. Easterly and Levine (2016) explicitly mention India, noting that their measure of European population excludes individuals who served temporary terms in the colonial administration, as it counts only those identified as permanent private settlers. The remaining 47 countries exhibit considerably smaller migrant counts, with a maximum of 49 and a mean of 9.4. Collectively, these 50 countries account for only 0.59% of our total genealogical sample.

[5] Arroyo Abad and Espín-Sánchez (2025) compile a comprehensive dataset of Spanish migration to Spanish America for the period 1492–1830, drawing on licenses, ship manifests, notarial records, and other archival sources from the *Archivo General de Indias* and related repositories. Their dataset records over 250,000 migrants and documents geographic origins, destinations, gender, and occupational characteristics.

3.2 Climate Data

We measure climate using gridded data from TerraClimate (Abatzoglou et al., 2018), accessed and processed in Google Earth Engine. TerraClimate provides monthly mean temperature and total precipitation at approximately 4 km spatial resolution globally from 1958 onward. To obtain measures of local climate, we average monthly values over the period 1960–2000, and aggregate them to a 0.5° grid. Mean annual temperature and precipitation are defined, respectively, as the average of monthly maximum and minimum temperatures (expressed in degrees Celsius) and the sum of monthly precipitation totals (expressed in millimeters per year). These long-run means are held fixed across cohorts and are used to characterize climatic conditions at both European origins and overseas destinations.

Figure 1, Panel B, displays the spatial distribution of mean annual temperature across European origins and overseas regions in our data. Temperate zones—such as North America, southern South America, Australia, and New Zealand—exhibit moderate temperatures similar to those in Europe, while tropical regions in Africa, South Asia, and parts of Southeast Asia are characterized by substantially higher temperatures. Contrasting Panel B with Panel A reveals that European settlement is disproportionately concentrated in regions with climates closer to those of Europe, particularly in temperate zones, and more limited in hotter tropical environments. This visual evidence is consistent with Crosby (1986)'s hypothesis, and suggests that climatic similarity shaped the geography of European overseas settlement—something we examine more systematically in Section 4.

For each origin–destination pair, we define temperature and precipitation distances as the absolute differences between the corresponding origin and destination grid cells. Figures 2 and 3 illustrate these measures using two examples: Argentina and Australia. In each figure, Panels A–B plot temperature distance from London to each grid cell alongside the number of settlers born in Great Britain, while Panels C–D present the analogous comparison for Seville and Spanish-born settlers.

Three patterns emerge. First, as expected, temperature distance is lower between London and Australia, and between Seville and Argentina; conversely, the pairs London–Argentina and Seville–Australia exhibit larger temperature differences. Second, there is substantial within-country variation: some areas of Australia are much closer in temperature to London than others, and a similar pattern holds for Argentina relative to Seville. Third, migration flows track these patterns. British settlements are, on average, higher in Australia than in Argentina, while the opposite holds for Spanish migrants. Moreover, within each destination country, regions that are closer in temperature to the origin attract more settlers, whereas

climatically more distant areas receive fewer migrants. These patterns provide suggestive evidence of climate matching, which we quantify in the regression analysis in Section 4.

3.3 Long-Run Demographic and Economic Outcomes

To evaluate the long-run implications of European overseas settlement and climate matching, we rely on two main datasets. First, we use data from Putterman and Weil (2010), who construct a migration matrix tracing the share of each country's year-2000 population descended from populations living in different source countries in 1500. Using this matrix, we compute both the share of each country's population descended from specific European source countries and the aggregate share descended from any European country, as in Easterly and Levine (2016). These measures enable us to test whether countries that attracted settlers who were more closely matched climatically exhibit greater long-run European demographic persistence.

Second, following the existing literature (Michalopoulos and Papaioannou, 2013; Henderson et al., 2018), we proxy for long-run economic performance using satellite-based nighttime luminosity from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band. Nighttime lights provide a spatially disaggregated measure of economic activity that allows us to exploit within-country variation while holding national institutions and policies constant. We use annual composite data for 2015, which provides largely cloud-free global coverage at approximately 500-meter resolution. To harmonize luminosity with the genealogical and climate data, we aggregate pixel-level values to $0.5° \times 0.5°$ grid cells by computing mean luminosity within each cell.

Figure 4 displays the spatial distribution of nighttime luminosity across our sample, revealing substantial variation in economic activity both across and within countries.[6]

3.4 Additional Data Sources

Soil characteristics. We measure soil conditions using data from SoilGrids, a global gridded soil database that provides predictions of soil properties at fine spatial resolution based on machine learning models trained on soil profile observations and environmental covariates. We extract four soil characteristics at each $0.5°$ grid cell: clay content, soil organic matter, pH, and water-holding capacity. Soil distance between origin and destination is computed as

[6] The map is shown at a coarser $2.5° \times 2.5°$ resolution for visualization purposes; all analysis is conducted at the $0.5° \times 0.5°$ grid-cell level.

the absolute difference in each soil characteristic, as well as a composite Euclidean distance across all four dimensions. We also construct a soil quality index using principal component analysis (PCA) across the four soil variables and compute the distance in the first principal component between origin and destination.

Travel costs. As a robustness check to straight-line geographic distance, we derive a stylized proxy for historical travel costs. This measure decomposes movement into three segments: (i) the inland distance from each origin grid-cell centroid to the principal seaport of the origin country, (ii) the oceanic distance between the principal origin and destination seaports, and (iii) the inland distance from the destination seaport to the destination grid-cell centroid. Principal seaports for each country are selected based on their documented role as long-distance gateways during the period and are held fixed across periods. This proxy is intentionally coarse—it does not attempt to reproduce specific sailing routes, prevailing winds, or riverine access—but it captures two margins that straight-line distance omits: port accessibility at origin and the tendency to settle nearer to ports at destination. Empirically, results using this measure closely track the baseline estimates.

Institutions. We measure institutional quality using four standard indicators widely employed in the literature.[7] First, we use the Polity2 score from the Polity IV project, which ranges from -10 (full autocracy) to 10 (full democracy) and captures the overall openness and competitiveness of political regimes. Second, we use the liberal democracy index from the Varieties of Democracy (V-Dem) dataset, a continuous measure that aggregates information on electoral processes, civil liberties, and constraints on executive authority. Third, we include the binary democracy indicator from Acemoglu et al. (2019). Finally, we use the Polity IV measure of constraints on the executive, which captures the extent of institutionalized limits on executive power.

4 Climate Matching and the Geography of European Settlement

This section tests the idea that European overseas settlements were more successful in areas whose ecological conditions resembled those of the European metropole. Section 4.1 outlines the empirical strategy. Section 4.2 presents the main findings, showing that settlers systematically selected destinations with climates closer to those of their origins. Section 4.3 demonstrates that the results are robust to alternative specifications and sample restric-

[7] All institutional variables are measured in 2010 and merged with our country-level dataset.

tions. Section 4.4 explores heterogeneity across birth cohorts, countries of origin, continents of destination, and the direction of temperature distance (hot versus cold).

4.1 Empirical Strategy

We aggregate the data described in Section 3.1 to the origin–destination–cohort level. Origins correspond to $0.5° \times 0.5°$ grid cells in Europe, and destinations to $0.5° \times 0.5°$ grid cells in non-European territories classified as colonies in Acemoglu et al. (2001). Cohorts are defined as 50-year birth intervals (e.g., 1500–1549, 1550–1599) spanning 1500 to 1850. Following the trade literature (Anderson and Van Wincoop, 2003), we estimate gravity models that relate migration flows to both geographic and climatic distances. While the standard gravity framework explains flows primarily as a function of geographic distance, we build on recent work by Obolensky et al. (2024) by incorporating ecological similarity, adding temperature and precipitation distance alongside geography. Because the migration matrix is sparse, we estimate the model using Poisson pseudo-maximum likelihood (PPML), following Silva and Tenreyro (2006). The estimating equation is:

$$M_{odt} = \exp[\alpha_{c(o)t} + \delta_{dt} + \beta \text{Dist}_{od}^{\text{Climate}} + \theta \text{Dist}_{od}^{\text{Physical}}]\epsilon_{odt} \tag{1}$$

where M_{odt} denotes the count of individuals born in European origin cell o in period t who died in destination cell d, and $Dist_{od}^{Climate}$ is the absolute difference in temperature and precipitation between origin o and destination d. We also control for geographic distance and its square ($\text{Dist}_{od}^{\text{Physical}}$), to capture migration costs and other unobserved correlates of spatial separation. The preferred specification also includes country of origin by period fixed effects ($\alpha_{c(o)t}$) and $0.5° \times 0.5°$ destination cell by period fixed effects (δ_{dt}). The former absorb time-varying push factors at the origin (e.g., demographic shocks, political instability, conflicts), while the latter account for time-varying pull factors at the destination (e.g., ecological conditions, colonial policies, institutional shifts). We cluster standard errors at the country-of-origin $\times$ $2.5°$ destination-cell level—a spatial unit large enough to capture serial correlation in unobserved local conditions (e.g., regional shocks or measurement noise in genealogical records).

The coefficient of interest is β, which captures the relationship between climate distance and overseas permanent settlements. Crosby (1986)'s hypothesis implies a negative coefficient on climate distance: the greater the climatic divergence between a given European origin and an overseas location, the less likely Europeans from that origin were to establish permanent settlements in that destination.

4.2 Main Results

We report estimates from equation (1) in Table 1. Column 1 presents a parsimonious specification including temperature distance, precipitation distance, and geographic distance. Both climate variables enter with negative and statistically significant coefficients, while geographic distance is negative but imprecisely estimated. Column 2 adds the squared term of geographic distance to allow for non-linear transport costs; the coefficients on temperature and precipitation distance remain stable, while geographic distance becomes large and statistically significant.

Columns 3 through 8 introduce progressively more demanding fixed effects. Column 3 includes 50-year period fixed effects to absorb global shocks common across all locations, such as technological change or shifts in trade conditions. Column 4 adds country-of-origin and country-of-destination fixed effects, controlling for time-invariant characteristics such as baseline population size, institutional capacity, or average mortality environments. Column 5 replaces destination-country fixed effects with administrative-level-1 fixed effects, capturing finer spatial heterogeneity within countries.[8] Column 6 interacts period dummies with origin-country and administrative-level-1 fixed effects, allowing for time-varying shocks specific to origin countries and destination subnational units. Column 7 further includes $0.5° \times 0.5°$ destination-cell fixed effects, and Column 8 replaces the period-by-administrative-level-1 interactions with period-by-$0.5° \times 0.5°$ destination-cell fixed effects.

Column 8 is our preferred specification. The inclusion of period-by-destination-cell fixed effects absorbs all time-varying local pull factors at the most granular spatial level—such as changes in settlement attractiveness, mortality environments, or resource discoveries. Thus, identification comes from within-period variation in ecological similarity across origin–destination pairs that share the same destination cell. Across all specifications, the coefficient on temperature distance remains negative, statistically significant, and remarkably stable in magnitude. The estimated effect implies that a 5°C increase in temperature distance—comparable to the average annual temperature difference between Madrid and Johannesburg, or between London and Buenos Aires—reduces the number of settlers by approximately 56%. By contrast, the coefficient on precipitation distance becomes smaller and less precisely estimated as additional controls and fixed effects are introduced. This pattern is consistent with Obolensky et al. (2024), who find that temperature plays a more robust role than precipitation in shaping migration flows in the U.S. context.

[8] Administrative-level-1 (from Natural Earth) refers to the first-level subnational administrative divisions, corresponding to states, provinces, regions, or equivalent jurisdictions depending on the country (e.g., states in the United States and Brazil, provinces in Argentina and Canada).

Overall, the results provide support for the climate-matching hypothesis: temperature similarity between origin and destination determined European settlement patterns overseas between 1500 and 1850. This relationship likely reflects several complementary mechanisms. First, climate distance may have deterred migration ex ante. Second, it may have increased return or onward migration from poorly matched destinations. Third, it may have reduced survival, limiting the likelihood that settlers left observable genealogical traces. While we cannot disentangle these channels, the estimated coefficient captures their combined effect, providing a summary measure of the overall influence of climate distance on settlement patterns.

4.3 Robustness Checks

In this section, we assess the robustness of the baseline specification (column 8 of Table 1) to alternative samples and functional forms.

Geographic distance and travel costs. First, we address concerns that the results may be sensitive to the functional form used to measure distance or to spatial correlation in climate. Results are reported in Table A.2. Column 1 reproduces the baseline specification, which includes geographic distance and its square. Column 2 replaces these terms with the logarithm of geographic distance. Column 3 additionally log-transforms temperature and precipitation distance, consistent with the standard gravity specification (Anderson and Van Wincoop, 2003). Column 4 replaces geographic distance with the travel-cost measure described in Section 3.4, which captures inland access, oceanic crossing, and inland travel at destination. Column 5 takes the logarithm of this travel-cost measure. Across all specifications, the coefficient on temperature distance remains negative, statistically significant, and quantitatively similar to the baseline.

Alternative samples. Next, we assess robustness to alternative samples. Figure A.4 reports these results, with the baseline specification shown as the leftmost point for comparison. We first restrict the sample to migrants born in the six major colonial empires—Spain, Portugal, Great Britain, France, Germany, and the Netherlands. We then exclude origin countries with fewer than 500 and 1,000 settlers, respectively. Next, we exclude the four "Neo-European" countries commonly referenced in the literature—Australia, Canada, New Zealand, and the United States (Easterly and Levine, 2016). Finally, we exclude African destinations, given their distinct colonial histories and climatic environment. Across all subsamples, the temperature-distance estimate remains negative and close to the baseline magnitude.

Country-of-origin by country-of-destination fixed effects. A remaining concern is that, despite the rich set of fixed effects in the baseline specification, unobserved factors at the country-pair level may still bias the estimates. For example, historical colonial ties, language proximity, or other bilateral relationships between countries may correlate with both migration flows and climatic similarity. To address this issue, column 2 of Table A.3 augments the preferred specification by including country-of-origin × country-of-destination fixed effects. This specification absorbs all time-invariant factors specific to each country pair. Because migration flows are measured at the origin–destination cell level, identification comes from variation within these country pairs across locations.[9] The results remain negative, statistically significant, and quantitatively similar to the baseline (reported in column 1 for comparison), indicating that the climate-matching effect is not driven by omitted country-pair characteristics.

Additional bilateral differences. Our analysis thus far has focused on mean annual temperature as a measure of climatic similarity. We now examine whether the results are robust to additional bilateral characteristics. This exercise serves two purposes. First, these factors may be correlated with temperature distance and independently influence migration patterns. Second, Crosby (1986)'s hypothesis emphasizes a broader notion of ecological compatibility beyond temperature alone. We report results from column 3 onward in Table A.3, where we sequentially add these variables. Column 3 includes absolute latitude distance. Column 4 adds elevation distance, column 5 incorporates terrain ruggedness, and column 6 introduces a dummy indicating whether both origin and destination have coastal access.[10]

Across all specifications, the coefficient on temperature distance remains stable in magnitude and statistically significant. In addition, several of these variables are independently predictive of migration flows. In particular, latitude distance enters with a negative and statistically significant coefficient, indicating that migrants were more likely to settle in regions with similar latitudinal positions. This likely reflects broader climatic and ecological similarities—such as seasonality, day length, and agricultural cycles—that are not fully captured by temperature alone. Elevation distance and coastal access also exhibit significant effects, consistent with the role of geographic accessibility and environmental conditions in shaping settlement patterns.

Soil characteristics. Another potentially important dimension of ecological compatibility is soil quality and composition. As emphasized by Crosby (1986), agro-climatic condi-

[9] In practice, this means comparing flows from different regions within the same origin country (e.g., Madrid versus Seville) to different locations within the same destination country (e.g., Buenos Aires versus other regions in Argentina), while holding constant all factors common to the Spain–Argentina pair.

[10] We define a cell as having coastal access if its centroid lies within 100 kilometers of the coastline.

tions—such as soil type, nutrient content, and drainage—play a central role in determining the viability of European crops and agricultural practices. If soil similarity is correlated with temperature distance, part of the estimated temperature effect may in fact reflect underlying differences in soil characteristics. To address this possibility, we construct several measures of bilateral soil distance and verify that our main result is robust to controlling for these measures (see also Section 3.4 for more details). We compute soil distance across four dimensions: clay content, soil organic matter, pH level, and water-holding capacity.

Column 2 of Table A.4 includes a raw Euclidean distance measure, standardizing each soil characteristic before taking differences between origin and destination. Since soil characteristics may themselves be determined by climate, column 3 uses soil distance computed from residuals after regressing each soil variable on origin and destination climate levels. Column 4 adopts an even more conservative approach, constructing soil distance from residuals after regressing soil characteristic distances on climate distances, thereby isolating variation in soil that is orthogonal to climate distance. Finally, column 5 includes all four individual residualized soil distances separately (clay, organic matter, pH, and water content). Across all specifications, the coefficient on temperature distance remains negative, statistically significant, and of similar magnitude to the baseline estimate in column 1. This indicates that temperature distance exerts an independent effect on migration flows beyond any correlation with soil characteristics.

Alternative climate statistics. Finally, Table A.5 assesses robustness to alternative climate measures. Column 2 augments the specification with temperature and precipitation ranges, with results remaining unchanged. Columns 3 and 4 replace mean temperature and precipitation distances with their minimum and maximum values, respectively, while column 5 includes both simultaneously. Across these specifications, the coefficients on temperature distance remain negative and statistically significant, indicating that both colder and warmer seasonal differences—captured by minimum and maximum temperatures—are relevant for settlement patterns.

4.4 Heterogeneity

We next examine how the relationship between climate distance and migration varies across contexts.

Hot vs cold destinations. We begin by decomposing temperature distance into its signed components. This decomposition can be informative because the constraints faced by settlers may differ across these two margins: moving to colder environments may limit the viability

of European crops and production systems, while moving to hotter environments may expose settlers to unfamiliar disease environments and physiological stress. As a result, the overall effect of temperature distance may mask fundamentally asymmetric responses to "cold" versus "hot" moves.

We replace absolute temperature distance with the signed temperature difference, defined at the 0.5° grid-cell level as the difference between destination and origin temperatures, so that moves to colder destinations take negative values. We then re-estimate the baseline specification using this temperature-difference measure. Figure 5 reports the coefficients by temperature-difference bins, using $(-1, 1]$ as the reference category. The estimates reveal a pronounced asymmetry. For the coldest bin, the coefficient is positive and statistically significant, indicating that migrants were relatively more likely to move to destinations that were substantially colder than their origin regions. For moderate cold differences, the estimates are close to zero and imprecisely estimated, suggesting limited sensitivity along this margin. The pattern for hotter destinations is markedly different. As temperature differences increase, coefficients become progressively more negative and statistically significant, indicating that migrants were increasingly less likely to move to substantially hotter environments. The decline is monotonic and steep in the highest temperature bins, pointing to a strong deterrent effect of large positive temperature differences.

These results suggest that the aggregate negative effect of temperature distance is primarily driven by aversion to hotter destinations rather than colder ones. While moves to colder climates do not appear to systematically deter migration, large increases in temperature relative to the origin substantially reduce migration flows. This pattern may reflect both a lower willingness of settlers to move to hotter environments and higher mortality in such locations, which would reduce the likelihood that migrants leave observable genealogical traces. Consistent with this interpretation, hotter environments likely imposed stronger constraints on European settlers—through disease exposure, agricultural incompatibility, or physiological limits—than colder environments, at least within the range of destinations observed in the data.

Main European origins. We next examine heterogeneity across major European origins, focusing on the six principal empires: Spain, Portugal, France, Great Britain, Germany, and the Netherlands. We present results in Figure A.5, standardizing temperature distance within each subsample to have mean zero and standard deviation one and reporting the full-sample estimate as the first dot on the left. Across empires, most point estimates are negative, though their magnitude and precision vary. The coefficients for Germany, Great Britain, and Portugal are negative and statistically significant, while Spain is also negative

but smaller in magnitude. For the Netherlands, the estimate is negative but very imprecisely estimated. France stands out as the main exception: the coefficient is close to zero and statistically insignificant.

Main regions of destination. We next examine heterogeneity across major regions of destination. We focus on five broad areas: North America, Central and Latin America, Africa, Asia, and Oceania. Figure A.6 reports the results.[11] All point estimates are negative and statistically significant, though their magnitude varies across regions. The coefficients for Africa, Asia, and especially Oceania are larger in absolute value, indicating a stronger role for climate matching in these destinations. By contrast, estimates for North America and Central and Latin America are smaller in magnitude.

Time periods. Finally, Figure A.7 reports estimates separately by 50-year birth cohorts.[12] The negative relationship between temperature distance and migration is present in most periods, but its magnitude varies over time. The coefficient is largest in absolute value for cohorts born in 1550–1599, declines sharply for cohorts born in the seventeenth century, is close to zero for 1650–1699, and then becomes negative again for cohorts born after 1700. The pattern is therefore not monotonic, but it does suggest that climate distance shaped migration throughout much of the colonial period, with especially strong effects in the late sixteenth century and again in the eighteenth and early nineteenth centuries.

5 Mortality and Fertility Consequences of Climate Matching

Having established that climate distance is an important determinant of European overseas settlement, we now examine the consequences of climate matching for settlers themselves. We focus on two demographic indicators of settlement success: mortality and fertility. Section 5.1 presents the empirical specification, and Section 5.2 reports the results.

5.1 Empirical Strategy

We consider the same sample of the analysis in Section 4, namely European settlers born between 1500 and 1850 who died outside Europe in territories classified as colonies in Ace-

[11] As before, we standardize temperature distance within each subsample to have mean zero and standard deviation one, and report the full-sample estimate as the first dot on the left.

[12] As in previous figures, we standardize temperature distance within each subsample to have mean zero and standard deviation one, and report the full-sample estimate as the first dot on the left.

moglu et al. (2001). We estimate individual-level regressions using OLS. The model takes the following form:

$$Y_{iodt} = \alpha_{c(o)t} + \delta_{dt} + \beta \, \text{Dist}_{od}^{\text{Climate}} + \theta \, \text{Dist}_{od}^{\text{Physical}} + \gamma \, Male_i + e_{iodt} \tag{2}$$

where i indexes individuals, o denotes place of birth in Europe, d denotes place of death, and t indicates the birth cohort. All variables on the right-hand side follow the notation introduced in equation (1), with the addition of an indicator equal to one if the individual is male ($Male_i$). Standard errors are clustered at the country-of-origin $\times$ 2.5° destination-cell level, as in the migration analysis.

For mortality, the dependent variable is age at death; for fertility, it is the number of children. Both outcomes can be observed in the genealogical records. Age at death is available for approximately 94% of settlers in our sample, while completed fertility is observed for only 33.2%. Two important caveats apply. First, the estimates may reflect selection into climate matching. For example, settlers with better health or greater resources may have been more able to move to climates similar to those of their origin regions. If this is the case, the estimated relationship between climate distance and mortality may partly capture differences in migrant characteristics rather than purely ecological constraints. Second, for fertility we cannot distinguish between genuine childlessness and non-reporting of children in family trees, and such reporting differences may be selective. Although we include a rich set of controls described below, these considerations suggest interpreting the demographic estimates as suggestive.

5.2 Results

Age at death. Table 2 reports the results from equation (2) for the effects of climate mismatch on mortality. The structure of the table mirrors that of Table 1. Column 1 presents a parsimonious specification including temperature distance, precipitation distance, and geographic distance. The coefficient on temperature distance is negative and statistically significant. By contrast, and consistent with the patterns observed in the migration analysis, the coefficient on precipitation distance is small and imprecisely estimated. The coefficient on temperature distance implies that increasing temperature distance by 5°C reduces age at death by 1.3 years (or, 1.93% relative to the sample mean). Results are similar when including geographic distance squared (column 2), 50-year period fixed effects (column 3), country-of-origin and country-of-destination fixed effects (column 4), replacing country-of-destination fixed effects with administrative-level-1 fixed effects (column 5), and interacting

period dummies with both country-of-origin and administrative-level-1 destination fixed effects (column 6). According to the coefficient in column 6, a 5°C increase in temperature distance reduces age at death by about 310 days (or, 1.26% relative to the sample mean).

The results are not robust to replacing country-of-destination fixed effects with $0.5° \times 0.5°$ destination-cell fixed effects (column 7), or to interacting these fixed effects with period dummies (column 8). In both cases, the coefficient on temperature distance changes sign and becomes statistically insignificant, likely reflecting limited within-cell variation across origins. This suggests that the estimates should be interpreted with caution. To partly address concerns about destination-specific omitted variables, Table A.6 replicates the specification in Table 2, column 6, while progressively adding destination-level controls associated with mortality environments. Column 2 adds mean annual temperature and precipitation. Column 3 includes slope, elevation, and terrain ruggedness. Column 4 adds malaria ecology, and column 5 further includes log population in 1500. Across all specifications, the coefficient on temperature distance remains negative, statistically significant, and similar in magnitude to that reported in column 1.

These findings suggest that the baseline relationship is not driven by omitted environmental characteristics of the destination. However, an important caveat remains: the estimates may reflect selection into climate matching. In particular, migrants who were healthier, wealthier, or otherwise better equipped to undertake long-distance migration may have been more likely to settle in destinations whose climates resembled those of their origin regions. If such positive selection is present, part of the estimated effect of climate distance on mortality may capture differences in the characteristics of settlers rather than purely the causal effect of ecological mismatch.

Completed fertility. We next examine the relationship between completed fertility and climate distance. Because the genealogical records only report positive numbers of children, observations with zero reported children are excluded from the analysis. Table 3 reports the results, following the same structure as the mortality specifications (Table 2). As before, the coefficient on temperature distance is negative and statistically significant: settlers who moved to destinations with climates farther from those of their origins had fewer children. This relationship remains robust when progressively adding fixed effects, up to country-of-origin-by-period and administrative-level-1-destination-by-period fixed effects (columns 2–6). A 5°C increase in temperature distance reduces the number of children by 0.16 (or, 4.56% relative to the sample mean).

The results are also robust to alternative definitions of destination fixed effects. Replacing

administrative-level-1 fixed effects with $0.5° \times 0.5°$ destination-cell fixed effects (column 7), or interacting them with period dummies (column 8), yields similar or larger coefficients, which remain negative and statistically significant. These findings should nevertheless be interpreted with caution. First, as discussed in the mortality analysis, the estimates may reflect selection into climate matching: migrants who were healthier or wealthier may have been more able to settle in destinations with climates similar to those of their origin regions. Second, the fertility measure may be subject to selective reporting. Completed fertility is observed only when it is positive, and we cannot distinguish between genuine childlessness and missing reports of children in genealogical records. If such reporting is non-random, the estimates may be biased.

Heterogeneity by signed temperature difference. Finally, we examine whether the demographic effects of climate mismatch differ depending on whether settlers moved to hotter or colder environments. This analysis parallels the heterogeneity exercise conducted earlier for migration flows (Figure 5). Because we lack sufficient variation to estimate the most demanding specification (column 8 of Table 2), we use the more parsimonious specification from column 6, which includes country-of-origin-by-period and administrative-level-1-destination-by-period fixed effects.[13] As in Figure 5, we divide temperature differences into bins and estimate separate coefficients for each bin.

Figure 6 reports the results for age at death (Panel A) and completed fertility (Panel B). The estimates display an inverse U-shaped pattern. On the cold side (negative temperature differences), age at death is lower in the coldest bins, with the strongest effects at large negative differences, while estimates for moderate cold differences are close to zero and imprecisely estimated. Fertility shows a similar but weaker pattern, with negative but generally imprecise coefficients. On the hot side (positive temperature differences), the decline is more pronounced and systematic. Both age at death and completed fertility fall as destinations become hotter relative to the origin, with a monotonic and steep gradient for age at death and a similarly negative—though smaller—pattern for fertility.

These results complement the migration patterns documented in Figure 5. While migrants were substantially less likely to move to hotter destinations, conditional on moving, hotter environments are also associated with worse demographic outcomes. In contrast, colder destinations do not appear to strongly deter migration, but extreme cold is nonetheless associated with lower longevity and, to a lesser extent, reduced fertility. Taken together,

[13] In principle, for fertility we could estimate the more demanding specification in column 8. However, for consistency with the mortality analysis, we use the specification in column 6 instead. The results are very similar when employing the more saturated specification for fertility.

these findings suggest that climate mismatch affected both selection into destinations and outcomes conditional on arrival. The demographic penalties are largest for moves to hotter environments, but are also present—though less precisely estimated—for extreme cold.

Although results in this section should be interpreted cautiously given the limitations discussed above, they are consistent with Crosby (1986)'s hypothesis that ecological compatibility shaped the demographic viability of European settlement. If Section 4 documented the role of climate matching in determining where Europeans chose to settle, the analysis in this section highlights an additional channel: climate compatibility also affected settlers' survival and fertility.

6 The Persistence of European Ancestry

A central implication of Crosby (1986)'s hypothesis is not only that ecological compatibility shaped where Europeans settled and how they fared, but also that it influenced the long-run demographic composition of these regions. If climatic similarity facilitated survival and fertility—as documented in Sections 4 and 5—then areas that were more ecologically compatible with European settlers should have been more likely to sustain large and persistent European populations over time. In this sense, climate matching should translate into long-run "population success," measured by the share of the population today that descends from European migrants.

To examine this implication, we rely on country-level data on the share of the population in 2000 that is of European ancestry from Putterman and Weil (2010) as described in Section 3.3. We estimate cross-country regressions using the 74 non-European countries in our sample. The key explanatory variable is the migrant-weighted average temperature distance between European origins and each destination, computed using settlers born between 1500 and 1850 who died in that country. This exercise is necessarily suggestive and should be interpreted with caution. First, the analysis is cross-sectional and based on 74 observations. Second, climate mismatch is not randomly assigned: migrants to more dissimilar environments may have been negatively selected, which could bias OLS estimates downward. Third, countries with higher average mismatch may differ systematically along other dimensions—such as disease environments, pre-colonial development, or geography—that also influence long-run population composition. While we control for a rich set of observables, omitted variable bias remains a concern.

With these caveats in mind, Table 4 reports the results. Column 1 includes temperature

distance alongside precipitation and geographic distance. The coefficient on temperature distance is negative and statistically significant. The estimate implies that a 5°C increase in climate mismatch is associated with a 15.5 percentage point lower European population share in 2000—a large effect relative to the sample mean of 19%. Columns 2 through 5 show that this relationship is robust to the inclusion of progressively richer controls. In column 2, we add destination climate characteristics (mean temperature, precipitation, and absolute latitude). Column 3 further controls for geographic features such as elevation, terrain ruggedness, slope, and soil characteristics. Column 4 introduces proxies for malaria ecology and population density in 1500. Column 5 adds legal origin and continent fixed effects. The coefficient on temperature distance declines in magnitude—from -0.041 to -0.018—but remains statistically significant. The estimate in column 5 implies that a 5°C increase in mismatch reduces the European population share by approximately 9 percentage points, or about 50% of the sample mean.

The remainder of the table assesses the robustness of this finding. Column 6 excludes the four "Neo-Europes" (Australia, Canada, New Zealand, and the United States). These destinations received exceptionally large and sustained European settler flows relative to most former colonies, and their inclusion could mechanically drive the relationship between climate matching and long-run population shares. The coefficient remains similar in magnitude, though statistically significant only at the 10% level. Column 7 excludes African destinations, to address concerns that the distinctive dynamics of Africa's colonial history—large-scale territorial control concentrated in the late-nineteenth-century Scramble for Africa and limited permanent settler migration—may unduly influence the results. Reassuringly, the coefficient is negative, statistically significant, and larger in magnitude.

Finally, we address the concern that results may capture the persistent effect of the scale of European settlement (Easterly and Levine, 2016), rather than the role of climate compatibility per se. Our measure of climate distance is constructed as a migrant-weighted average across European origins, so that it varies with the composition of settlers—specifically, how climatically similar their origins were to the destination—rather than with the total number of migrants. To the extent that more numerous settlers simply happened to come from more climatically similar origins, however, scale effects could still confound the estimates. In column 8, we therefore augment the preferred specification with the log number of European settlers recorded in each destination country.[14] The coefficient on temperature distance remains negative and in line with that in column 5, though now statistically significant only

[14] This is a potentially "bad control" (Angrist and Pischke, 2009): migration flows are themselves shaped by temperature distance, as shown in Section 4, so conditioning on the number of settlers may absorb part of the very effect we seek to identify. The estimates in column 8 should thus be interpreted with caution.

at the 10% level. As expected, the coefficient on log migrants is positive and statistically significant, consistent with the findings of Easterly and Levine (2016).

Taken together, these results provide suggestive evidence that climate matching shaped not only the geography of European settlement and settlers' demographic outcomes, but also the long-run persistence of European populations. Countries that attracted settlers facing lower climatic mismatch during the colonial period tend to exhibit higher shares of European ancestry today. While the cross-sectional nature of the analysis precludes strong causal interpretation, the findings are consistent with the notion that ecological compatibility was an important determinant of the demographic success of European expansion.

7 Climate Matching and Long-Run Economic Development

The preceding sections have shown that climatic similarity shaped both the geography of European settlement and the long-run persistence of European populations. We now examine whether these patterns extend to economic development. If climate matching facilitated durable settler societies—and if European settlement influenced local institutional, technological, and demographic trajectories—then areas settled by better-matched Europeans may have higher levels of economic activity today. Following the literature (Michalopoulos and Papaioannou, 2013, 2014; Henderson et al., 2018), we use satellite-based nighttime luminosity as a proxy for subnational economic performance. Luminosity data provide a spatially disaggregated measure of economic activity that can be matched to our gridcell-level data on European settlement and allow us to exploit variation within countries and subnational regions.

Section 7.1 outlines the empirical strategy. Section 7.2 presents the results, showing that luminosity is lower in areas where European settlers experienced greater climatic mismatch. Section 7.3 demonstrates that these findings are robust to alternative samples and specifications.

7.1 Empirical Strategy

We restrict the analysis to $0.5° \times 0.5°$ destination cells that received at least one European settler born between 1500 and 1850. We estimate cross-sectional OLS regressions relating present-day luminosity to the average climate distance between European settlers' origins

and each destination:

$$\log(L_d) = \alpha + \beta \overline{\text{Dist}}_d^{\text{Temp}} + \gamma \overline{\text{Dist}}_d^{\text{Precip}} + \theta \overline{\text{Dist}}_d^{\text{Geo}} + \mathbf{X}_d'\delta + \mu_{r(d)} + \varepsilon_d \tag{3}$$

The dependent variable is the log of nighttime luminosity from VIIRS 2015 data.[15] The main explanatory variable of interest is the settler-weighted average temperature distance at destination d, $\overline{\text{Dist}}_d^{\text{Temp}}$, computed as:

$$\overline{\text{Dist}}_d^{\text{Temp}} = \sum_o \omega_{od} \times \text{TempDist}_{od}. \tag{4}$$

where TempDist_{od} is the absolute temperature difference between each settler's European origin o and destination d, and ω_{od} denotes the share of European settlers in destination d who were born in origin o between 1500 and 1850. This measure differs from the bilateral temperature distance used in the gravity analysis in two respects. First, it is a destination-level summary that averages over all European origins sending settlers to d, weighting by migrant shares. Second, it reflects both the climatic similarity of the destination to European conditions and the composition of the settler population: destinations that attracted settlers from more climatically similar origins will exhibit lower average temperature distance.

Equation (3) also controls for average precipitation distance and geographic distance, constructed in the same way as temperature distance. In our preferred specification, we further include a vector of destination-level controls, $\mathbf{X}_d$: mean temperature and precipitation, absolute latitude, slope, log elevation, roughness, malaria ecology, and log population density in 1500.[16] Finally, we control for fixed effects at either the country or administrative-level-1 level, $\mu_{r(d)}$, which absorb all common institutional, historical, and geographic factors shared within these regions, such as colonial regimes, national policies, and broad differences in state capacity. Standard errors are clustered at the $2.5° \times 2.5°$ grid cell level.

Because both the dependent variable and the key regressor vary at the destination level, equation (3) does not exploit bilateral variation as in Sections 4 and 5. Instead, identification comes from comparing locations within the same subnational region (equivalent to a U.S. state) that received settlers with different average levels of climate mismatch, conditional on observed characteristics and fixed effects. Since the temperature distance measure is constructed using migrant shares, this variation reflects differences in the climatic composition of settlers rather than the overall scale of European settlement. Two considerations are important for interpretation. First, the estimates may partly capture selection into cli-

[15] As in Michalopoulos and Papaioannou (2013), we add 0.01 to nighttime luminosity.
[16] Following Michalopoulos and Papaioannou (2013), we add 0.01 to 1500 population to retain observations with zero values.

mate matching—for example, if more skilled or wealthier migrants were better able to settle in climatically similar environments. Second, despite the inclusion of detailed controls and subnational fixed effects, omitted local characteristics may still bias the estimates if locations with greater mismatch differ systematically along other dimensions that affect long-run development.

7.2 Main Results

Table 5 reports the results. Column 1 presents a parsimonious specification including temperature distance, precipitation distance, geographic distance, and a limited set of climate controls—mean temperature, mean precipitation, and absolute latitude. The coefficient on temperature distance is negative and statistically significant, indicating that locations where European settlers were more mismatched in temperature exhibit lower luminosity today. Columns 2 and 3 progressively add geographic controls (slope, elevation, and terrain ruggedness) and proxies for the disease environment and pre-colonial population density. The coefficient remains stable and economically meaningful, implying that a 5°C increase in temperature distance among settlers is associated with a 38.7% reduction in nighttime luminosity.

Column 4 introduces country fixed effects. The coefficient declines by about 40%—from -0.098 to -0.061—but remains statistically significant at the 1% level. Column 5 replaces country fixed effects with administrative-level-1 fixed effects. The coefficient on temperature distance remains negative and statistically significant, indicating that even within subnational regions, locations that received settlers from more climatically similar origins exhibit higher luminosity today. According to the specification in column 5, a 5°C increase in temperature distance among settlers is associated with a 24.8% reduction in present-day luminosity.

One interpretation of these findings is that ecological compatibility influenced the types of institutions that emerged, consistent with the institutional view of colonial development (Acemoglu et al., 2001). Locations where European settlers could more easily adapt may have supported more inclusive institutional arrangements, whereas areas with greater mismatch may have favored more extractive systems. Consistent with this interpretation, the coefficient declines when we include country fixed effects, suggesting that part of the relationship operates through cross-country institutional differences.

To examine this channel more directly, Table 6 relates temperature mismatch to a range of contemporary measures of institutional quality (see Section 3.4), using the same country-level

specification as in Table 4. Given the cross-sectional nature of the analysis, these results should be interpreted as suggestive. In column 1, the dependent variable is the Polity2 score. The coefficient is negative and statistically significant at the 10% level, indicating that countries where European settlers were more climatically mismatched exhibit lower levels of democracy today. A 5°C increase in mismatch is associated with a reduction of 3.34 points (80% of the sample mean). Results are similar when using the V-Dem liberal democracy index in column 2. The coefficients remain negative but become imprecise in column 3, where we use the binary democracy indicator from Acemoglu et al. (2019), suggesting that the relationship operates more along the intensive margin of institutional quality than at the extensive margin between democracy and autocracy. Finally, temperature mismatch is negatively associated with constraints on the executive in column 4, although statistically significant only at the 10% level. According to our estimates, a 5°C increase in mismatch is associated with a reduction of -1.09 in executive constraints (22% of the sample mean).

At the same time, national institutions are unlikely to fully account for the relationship between climate mismatch and long-run development. Indeed, the estimates in Table 5 remain economically large and statistically significant even when we include subnational fixed effects, which absorb cross-country institutional differences and other national-level factors. This pattern suggests that the observed relationship is not driven solely by variation in national institutions. Instead, it is also consistent with differences in local institutional environments—such as subnational variation in governance, property rights, and state capacity—that may have been shaped by the degree of ecological compatibility faced by early European settlers (Dell, 2010; Michalopoulos and Papaioannou, 2013, 2014).

Other mechanisms may also contribute to the observed relationship. Climatic similarity may have facilitated the transfer and adoption of European technologies and production systems, or encouraged trade and investment by lowering the costs of operating in similar environments. In addition, by enabling larger and more persistent settler populations, climate matching may have generated scale effects and agglomeration economies. While we do not attempt to disentangle these channels, the evidence is consistent with a combination of these mechanisms.

7.3 Robustness Checks

We now assess the robustness of the results reported in Section 7.2. Table A.7 presents the results, reproducing the baseline specification (Table 5, column 5) in column 1 to facilitate comparisons. We begin by addressing a key concern: locations with low temperature distance

from European settlers may simply be those with intermediate or temperate climates. While in the gravity framework (Section 4) we can absorb such destination-specific characteristics through fixed effects, this is not feasible here. Although we control for mean temperature, absolute latitude, and subnational fixed effects, these controls may not fully capture non-linearities in climate. To address this concern, column 2 replaces mean temperature with temperature bins, allowing for a more flexible specification. The coefficient on temperature distance declines somewhat in absolute value, but remains statistically significant and close to the baseline estimate.

We next address the concern that our results may primarily reflect scale effects—that is, that locations with lower climate mismatch simply attracted more European settlers and, as a result, exhibit higher development today. To assess this possibility, columns 3 and 4 introduce controls that absorb variation in the scale of settlement, albeit at the cost of potentially over-controlling. In column 3, we control for the log number of European migrants. This is a demanding specification, since migration flows are themselves a function of temperature distance (as shown in Section 4), and our mismatch measure is constructed using migrant shares. Nevertheless, the coefficient on temperature distance remains negative and statistically significant, though smaller in magnitude (-0.037). In column 4, we replace 1500 population density with its contemporary counterpart. This is also a stringent test, as modern population may itself be an outcome of climate distance. Despite this, the coefficient remains negative, though further attenuated and statistically significant at the 10% level.[17]

The remainder of the table examines robustness to alternative samples. Column 5 excludes the four Neo-European countries—Australia, Canada, New Zealand, and the United States. The coefficient becomes somewhat smaller and is statistically significant at the 10% level, but remains broadly in line with the baseline estimates. Column 6 excludes African destinations, to ensure that the results are not driven by the distinct historical dynamics of the Scramble for Africa (Michalopoulos and Papaioannou, 2016). The coefficient remains negative and statistically significant.

In column 7, we restrict the sample to destinations that received at least 50 European settlers in our data. In this case, the coefficient increases substantially in absolute value. One interpretation is that this reflects the attenuation of measurement error or noise in locations with very small numbers of settlers, which may bias baseline estimates toward zero. Another possibility is that climate mismatch has stronger effects precisely in locations

[17] These specifications should be interpreted with caution, as both controls may be endogenous and potentially constitute "bad controls" in the sense of Angrist and Pischke (2009). Even so, the persistence of the results under these demanding specifications provides additional reassurance that our findings are not driven solely by scale effects.

where settlement was more substantial, consistent with complementarity between climate compatibility and the scale of European presence. While this pattern may also partly reflect scale effects, the fact that the coefficient increases when focusing on more intensively settled areas suggests that our baseline estimates are, if anything, conservative.

8 Conclusions

The scale, geography, and durability of European overseas expansion varied widely across regions. In some areas, Europeans settled in large numbers and established enduring societies; in others, their presence remained limited, transient, or extractive. A large literature has sought to explain these differences, emphasizing factors such as disease environments and settler mortality (Acemoglu et al., 2001), factor endowments and inequality (Engerman and Sokoloff, 2000), military and technological advantages (Thomas, 1993), and broader geographic and biogeographic conditions (Diamond, 1997). Complementing these perspectives, Alfred Crosby argued that European expansion depended on the extent to which overseas environments could sustain European populations, crops, and livestock (Crosby, 1986).

In this paper, we bring Crosby (1986)'s hypothesis to the data by combining genealogical records on more than 300,000 Europeans born between 1500 and 1850 who died overseas with high-resolution climate data. Estimating gravity regressions building on the trade and migration literature (Anderson and Van Wincoop, 2003; Obolensky et al., 2024), we show that temperature similarity strongly predicts European overseas settlements. We further document that climate mismatch had meaningful consequences for settlers themselves. Europeans who moved to more climatically dissimilar environments experienced lower demographic success, living shorter lives and having fewer children.

We then show that these historical patterns have persistent consequences. Locations where European settlers were more closely matched to local climates exhibit larger shares of European ancestry today, indicating that ecological compatibility shaped the long-run demographic success of European populations overseas. These differences extend to economic development: areas settled by more climate-matched Europeans display significantly higher levels of economic development, proxied by nighttime luminosity, even within subnational regions. One interpretation of these findings is that ecological compatibility influenced the types of institutions that emerged, consistent with the institutional view of colonial development (Acemoglu et al., 2001). In line with this view, we find that countries with greater climate mismatch exhibit lower institutional quality today. However, because the relationship between economic development and climate mismatch remains strong even after controlling

for country fixed effects, this channel cannot fully account for the patterns we document.

Our analysis leaves several important questions open for future research. First, the long-run results are based on cross-sectional variation and do not rely on an instrument for climate matching. As a result, they may partly reflect selection—for example, if more skilled or better-resourced migrants were more likely to settle in climatically similar environments. Future work could exploit natural experiments or plausibly exogenous variation in settlement patterns to strengthen causal interpretation. Second, while our results are consistent with multiple mechanisms—including institutional development, technology transfer, and scale effects—we do not disentangle their relative importance. Advancing our understanding of these channels, and how they interact with ecological conditions, remains an important avenue for future research.

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.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*.

——— (2002): "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*.

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2019): "Democracy does cause growth," *Journal of political economy*, 127, 47–100.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2016): "Ethnic inequality," *Journal of Political Economy*, 124, 428–488.

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

Angrist, J. D. and J.-S. Pischke (2009): *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Arroyo Abad, L. and J.-A. Espín-Sánchez (2025): "The First Numbers of the Americas," CEPR Discussion Paper DP20591, CEPR.

Banerjee, A. and L. Iyer (2005): "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *American Economic Review*.

Blanc, G. and R. Wacziarg (2025): "Malthusian Migrations," *National Bureau of Economic Research*.

Bruhn, M. and F. A. Gallego (2012): "Good, bad, and ugly colonial activities: do they matter for economic development?" *Review of economics and statistics*, 94, 433–461.

Chiovelli, G., S. Michalopoulos, E. Papaioannou, and T. Regan (2026): "Illuminating the global south," *The Economic Journal*, ueaf134.

Colasurdo, A. and R. Omenti (2024): "Using online genealogical data for demographic research: An empirical examination of the FamiLinx database," *Demographic Research*.

Crosby, A. W. (1986): *Ecological Imperialism: The Biological Expansion of Europe, 900–1900*, Cambridge University Press.

Dell, M. (2010): "The persistent effects of Peru's mining Mita," *Econometrica*.

Dell, M. and B. A. Olken (2020): "The development effects of the extractive colonial economy: The dutch cultivation system in java," *The Review of Economic Studies*, 87, 164–203.

Diamond, J. (1997): *Guns, Germs, and Steel: The Fates of Human Societies*, W. W. Norton & Company.

Dunn, O. and J. E. Kelley (1989): *The diario of Christopher Columbus's first voyage to America, 1492–1493*, University of Oklahoma Press.

Easterly, W. and R. Levine (2016): "The European origins of economic development," *Journal of Economic Growth*, 21, 225–257.

Elkins, C. (2022): *Legacy of Violence: A History of the British Empire*, Alfred A. Knopf.

Elliott, J. H. (2006): *Empires of the Atlantic World: Britain and Spain in America 1492–1830*, Yale University Press.

Engerman, S. L. and K. L. Sokoloff (1997): "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States," in *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800–1914*, ed. by S. Haber, Stanford, CA: Stanford University Press, 260–304.

——— (2000): "Institutions, Factor Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives*.

——— (2002): "Factor endowments, inequality, and paths of development among new world economics," .

Fenske, J., B. Gupta, and A. Mukhopadhyay (2025): "Colonial persistence," .

Feyrer, J. and B. Sacerdote (2009): "Colonialism and modern income: islands as natural experiments," *The Review of Economics and Statistics*, 91, 245–262.

Henderson, J. V., T. Squires, A. Storeygard, and D. Weil (2018): "The global distribution of economic activity: nature, history, and the role of trade," *The Quarterly Journal of Economics*, 133, 357–406.

Kaplanis, J., A. Gordon, T. Shor, O. Weissbrod, D. Geiger, M. Wahl, M. Gershovits, B. Markus, M. Sheikh, M. Gymrek, G. Bhatia, D. G. MacArthur, A. L. Price, and Y. Erlich (2018): "Quantitative analysis of population-scale family trees with millions of relatives," *Science*, 360, 171–175.

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. D. Sachs (2004): "A global index representing the stability of malaria transmission," *The American Journal of Tropical Medicine and Hygiene*, 70, 486–498.

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

Lindqvist, S. (1992): *Exterminate All the Brutes*, The New Press.

Maloney, W. F. and F. Valencia Caicedo (2016): "The persistence of (subnational) fortune," *The Economic Journal*, 126, 2363–2401.

Manjapra, K. (2020): *Colonialism in Global Perspective*, Cambridge University Press.

McNeill, W. H. (1976): *Plagues and Peoples*, Garden City, NY: Anchor Press/Doubleday.

Michalopoulos, S. and E. Papaioannou (2013): "Pre-colonial ethnic institutions and contemporary African development," *Econometrica*, 81, 113–152.

——— (2014): "National institutions and subnational development in Africa," *The Quarterly journal of economics*, 129, 151–213.

——— (2016): "The long-run effects of the scramble for Africa," *American Economic Review*, 106, 1802–1848.

——— (2020): "Historical legacies and African development," *Journal of Economic Literature*, 58, 53–128.

Newe, T. (1911): "Letter from Carolina, 1682," in *Narratives of Early Carolina, 1650–1708*, ed. by A. S. Salley, Charles Scribner's Sons, 131–139.

Nunn, N. and L. Wantchekon (2011): "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*.

Obolensky, M., M. Tabellini, and C. Taylor (2024): "Migration, Climate Similarity, and the Consequences of Climate Mismatch," Tech. rep., National Bureau of Economic Research.

Osterhammel, J. (1997): *Colonialism: A Theoretical Overview*, Markus Wiener Publishers.

Pomeranz, K. (2000): *The Great Divergence: China, Europe, and the Making of the Modern World Economy*, Princeton University Press.

Putterman, L. and D. N. Weil (2010): "Post-1500 population flows and the long-run determinants of economic growth and inequality," *The Quarterly journal of economics*, 125, 1627–1682.

Silva, J. S. and S. Tenreyro (2006): "The log of gravity," *The Review of Economics and statistics*, 641–658.

Stannard, D. E. (1992): *American Holocaust: The Conquest of the New World*, Oxford University Press.

Subrahmanyam, S. (1997): *The Career and Legend of Vasco da Gama*, Cambridge University Press.

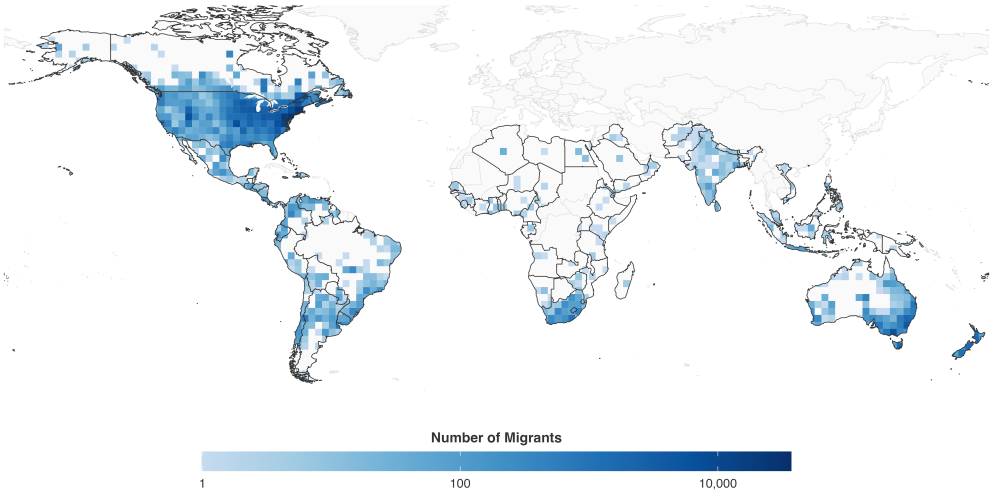
Thomas, H. (1993): *Conquest: Montezuma, Cortés, and the Fall of Old Mexico*, Simon & Schuster.

Wallerstein, I. (1974): *The Modern World-System, Vol. I: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*, Academic Press.

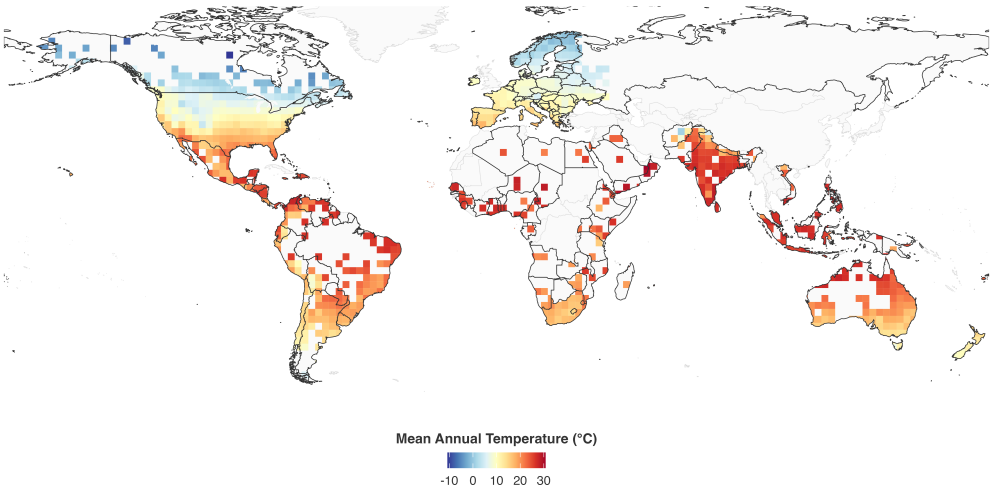
Figures and Tables

Figure 1. Number of Migrants and Average Annual Temperature

(A) Number of Migrants



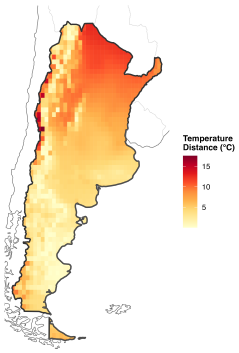
(B) Average Annual Temperature



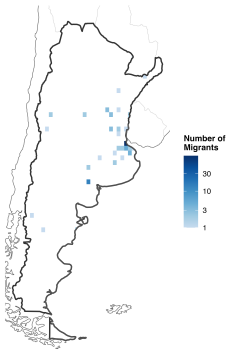
Notes: Panel A reports the number of European migrants (1500–1850) by destination grid cell using FamiLinx data. Panel B reports mean annual temperature (°C) from TerraClimate (1960–2000 averages). Both panels are aggregated to 2.5° × 2.5° cells for visualization.

Figure 2. Temperature Distance and Migrant Settlement Patterns: Argentina

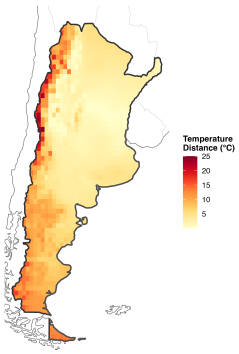
(A) Temperature Distance from London



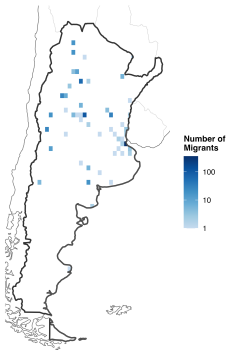
(B) Number of Settlers Born in Great Britain



(C) Temperature Distance from Seville



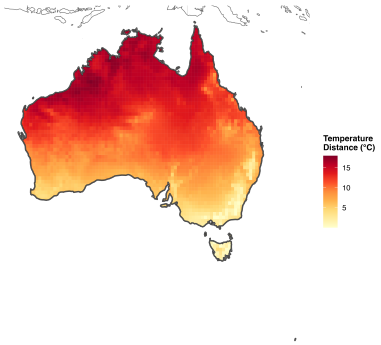
(D) Number of Settlers Born in Spain



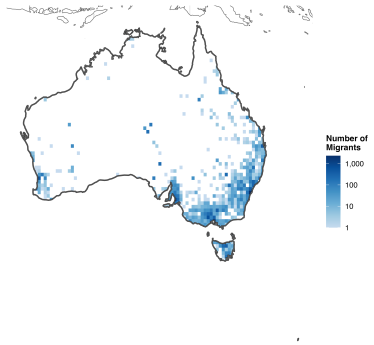
Notes: Panels A and B display, respectively, temperature distance from London to each 0.5° × 0.5° grid cell and the number of settlers born in Great Britain and dying in each destination cell. Panels C and D present the analogous maps for Seville and Spanish-born settlers. Mean annual temperature at the origin is 10.9 °C for London and 18.3 °C for Seville. Temperature distance is defined as the absolute difference in mean annual temperature (in °C) between origin and destination grid cells. The total number of migrants from each origin represented in the figure is 252 for Great Britain and 2,050 for Spain.

Figure 3. Temperature Distance and Migrant Settlement Patterns: Australia

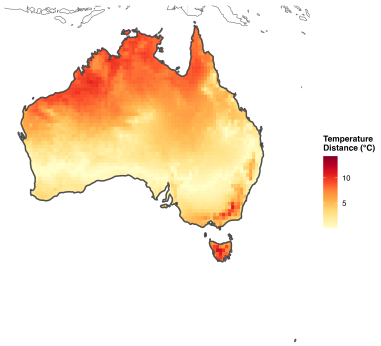
(A) Temperature Distance from London



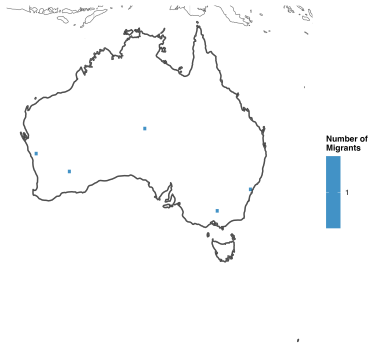
(B) Number of Settlers Born in Great Britain



(C) Temperature Distance from Seville

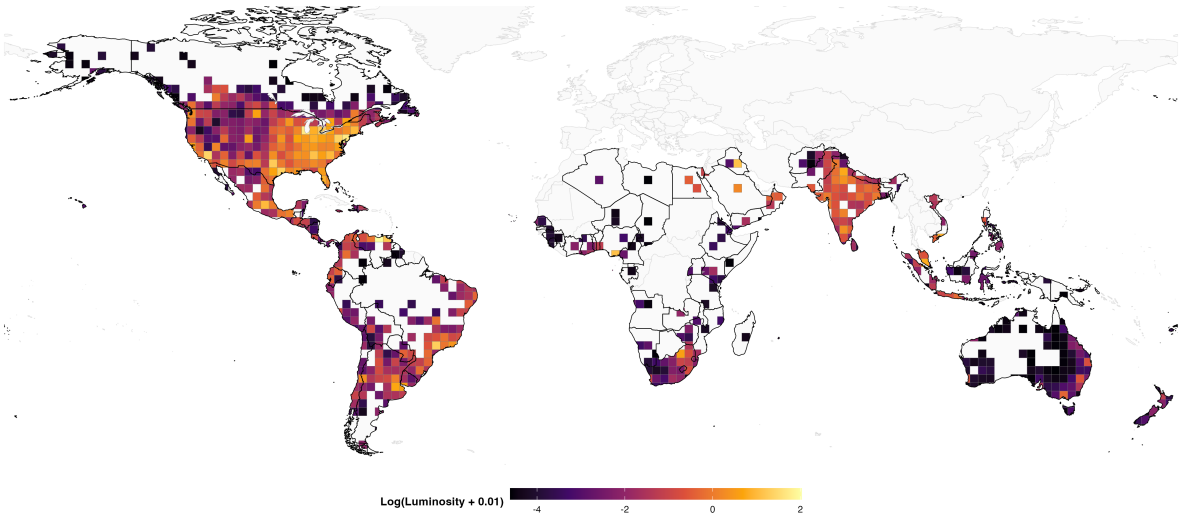


(D) Number of Settlers Born in Spain



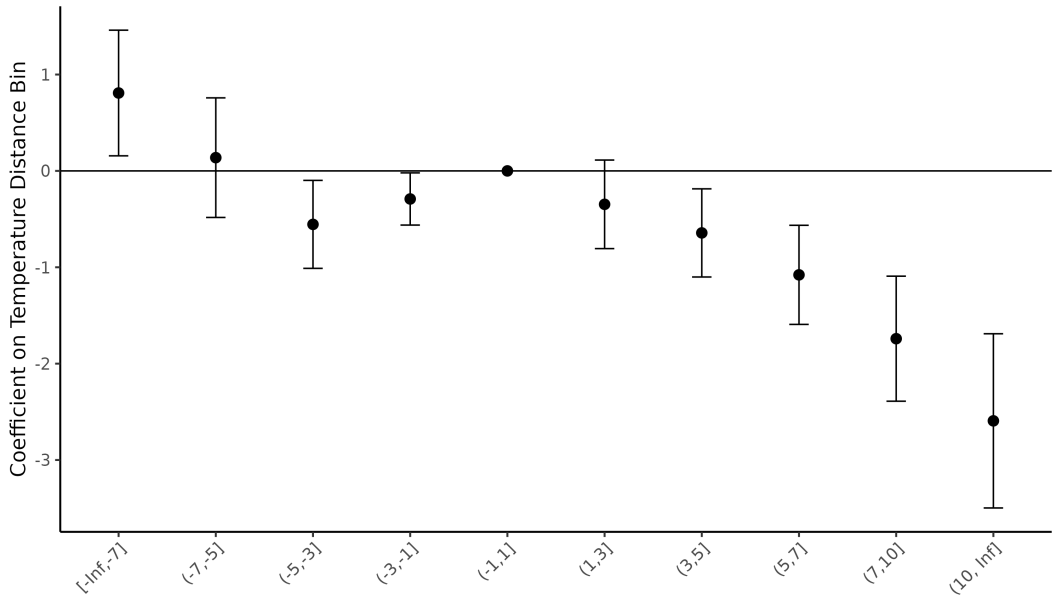
Notes: Panels A and B display, respectively, temperature distance from London to each 0.5° × 0.5° grid cell and the number of settlers born in Great Britain and dying in each destination cell. Panels C and D present the analogous maps for Seville and Spanish-born settlers. Mean annual temperature at the origin is 10.9 °C for London and 18.3 °C for Seville. Temperature distance is defined as the absolute difference in mean annual temperature (in °C) between origin and destination grid cells. The total number of migrants from each origin represented in the figure is 21,514 for Great Britain and 5 for Spain.

Figure 4. Nighttime Luminosity



Notes: This map displays mean nighttime luminosity from VIIRS 2015 data, aggregated to 2.5° × 2.5° grid cells by computing mean luminosity within each cell". The sample includes grid cells in former colonies where at least one European settler born between 1500 and 1850 is recorded to have died in the FamiLinx genealogical data. The color scale represents log(Luminosity + 0.01).

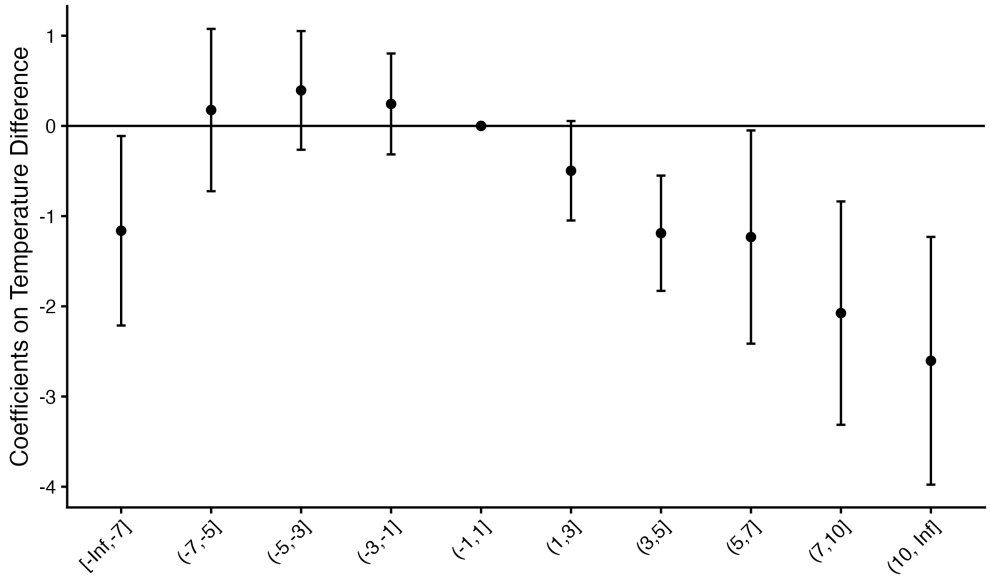
Figure 5. Non-linear Effects of Temperature Distance on Migration Flows



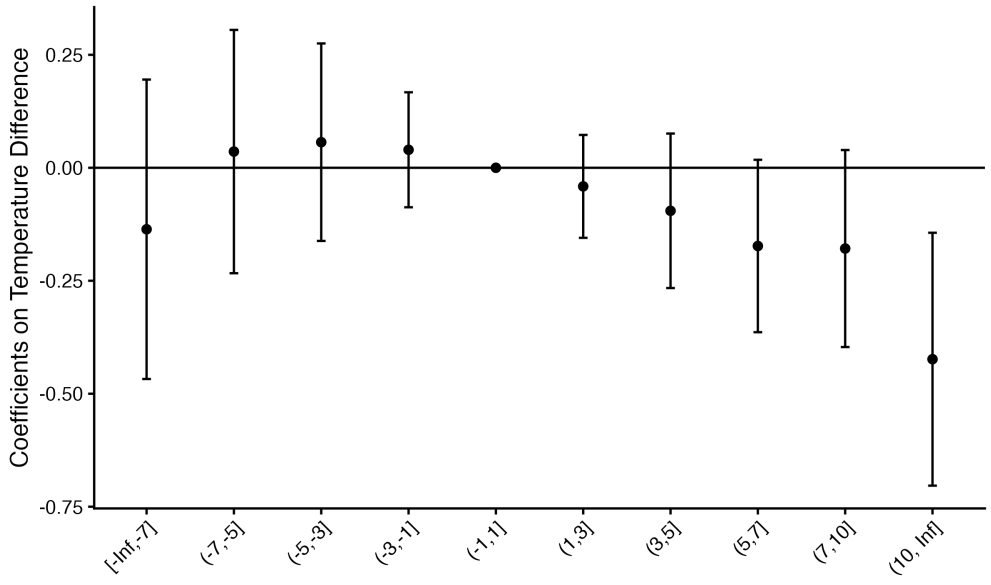
Notes: The figure plots estimated coefficients on temperature-difference bins from a Poisson regression. The dependent variable is the number of migrants in the FamiLinx database born in Europe and dying overseas, aggregated to the origin–destination–period cell at 0.5° × 0.5° resolution for 50-year periods between 1500 and 1850. Temperature difference is defined as the signed difference in mean annual temperature between destination and origin grid cells (destination minus origin, in °C). The omitted category is the bin containing zero (i.e., destinations with similar temperatures to the origin). The specification includes precipitation distance, geographic distance and its square, country-of-origin-by-period fixed effects, and destination-cell-by-period fixed effects. Confidence intervals at the 95% level are reported. Standard errors are clustered at the country-of-origin × 2.5° × 2.5° destination-cell level.

Figure 6. Temperature Distance and Demographic Outcomes: Binned Analysis

(A) Age at Death



(B) Completed Fertility



Notes: The figure plots estimated coefficients from regressions of age at death (Panel A) and completed fertility (Panel B) on temperature-difference bins. Temperature difference is defined as the signed difference in mean annual temperature between destination and origin grid cells (destination minus origin, in °C). The omitted category is the bin containing zero (i.e., destinations with similar temperatures to the origin). All specifications include precipitation distance, geographic distance and its square, country-of-origin-by-period fixed effects, and administrative-level-1-by-period fixed effects. Panel A additionally controls for an indicator for male individuals. Confidence intervals at the 95% level are reported. Standard errors are clustered at the country-of-origin × 2.5° × 2.5° destination-cell level.

Table 1. Climate Distance and Migration Flows

Dep. var.:	Number of Migrants							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature Distance	-0.193***	-0.169***	-0.162***	-0.139***	-0.094***	-0.100***	-0.168***	-0.166***
	(0.036)	(0.039)	(0.036)	(0.035)	(0.032)	(0.029)	(0.040)	(0.041)
Precipitation Distance	-0.006***	-0.007***	-0.006**	-0.004	-0.003	-0.004	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Geographic Distance	-0.091	-1.144***	-1.014***	-0.943***	0.301	0.345	1.141	1.145
	(0.058)	(0.158)	(0.130)	(0.205)	(0.495)	(0.578)	(0.757)	(0.755)
Observations	13,793,028	13,793,028	13,793,028	13,793,028	13,793,028	13,786,220	13,786,220	13,786,220
R^2	0.000	0.001	0.003	0.007	0.024	0.038	0.132	0.133
Mean Temp. Distance	7.12	7.12	7.12	7.12	7.12	7.12	7.12	7.12
Mean Prec. Distance	38.87	38.87	38.87	38.87	38.87	38.87	38.87	38.87
Mean Geographic Distance	8.70	8.70	8.70	8.70	8.70	8.70	8.70	8.70
Geog. Dist ^{2}		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE			Yes	Yes	Yes			
Country of origin FE				Yes	Yes			
Country of destination FE				Yes				
Administrative level 1 FE					Yes			
Country of origin × period FE						Yes	Yes	Yes
Administrative level 1 × period FE						Yes	Yes	
Destination cell FE (0.5° × 0.5°)							Yes	
Destination cell (0.5° × 0.5°) × period FE								Yes

Notes: The sample includes all origin–destination pairs formed by European grid cells (origins) and overseas grid cells (destinations) at a 0.5° × 0.5° resolution, aggregated into 50-year periods between 1500 and 1850. The dependent variable, Number of Migrants, is the count of individuals in the FamiLinx database born in Europe and dying overseas, collapsed to the origin–destination–period cell. Temperature Distance is the absolute difference in mean annual temperature between origin and destination grid cells. Precipitation Distance is defined analogously. Geographic Distance is the great-circle distance between the centroids of origin and destination grid cells, expressed in 10^3 km. Column (1) reports a parsimonious specification. Column (2) adds the square of geographic distance. Column (3) includes period fixed effects. Column (4) adds country-of-origin and country-of-destination fixed effects. Column (5) replaces destination-country fixed effects with administrative-level-1 fixed effects. Column (6) interacts period fixed effects with country-of-origin and administrative-level-1 fixed effects. Column (7) adds destination-cell fixed effects. Column (8) replaces administrative-level-1-by-period fixed effects with destination-cell-by-period fixed effects. Standard errors, reported in parentheses, are clustered at the country-of-origin × 2.5° × 2.5° destination-cell level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2. Climate Distance and Mortality of Settlers

Dep. var.:	Age at Death							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature Distance	-0.260***	-0.350***	-0.306***	-0.355***	-0.185***	-0.170***	0.073	0.077
	(0.074)	(0.071)	(0.054)	(0.067)	(0.052)	(0.050)	(0.051)	(0.052)
Precipitation Distance	-0.035***	-0.033***	-0.016***	-0.008**	-0.002	-0.001	-0.001	0.001
	(0.006)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Geographic Distance	0.303***	1.971***	-0.561*	1.460***	-1.990***	-1.292*	-2.545***	-2.150**
	(0.062)	(0.396)	(0.299)	(0.480)	(0.736)	(0.775)	(0.882)	(0.862)
Observations	312,758	312,758	312,758	312,749	312,633	312,326	311,745	310,408
R^2	0.007	0.009	0.036	0.044	0.056	0.076	0.088	0.131
Mean Temp. Distance	3.337	3.337	3.337	3.337	3.337	3.337	3.337	3.337
Mean Prec. Distance	29.23	29.23	29.23	29.23	29.23	29.23	29.23	29.23
Mean Geographic Distance	7.453	7.453	7.453	7.453	7.453	7.453	7.453	7.453
Mean Age at Death	67.41	67.41	67.41	67.41	67.41	67.41	67.41	67.41
Geog. Dist ^{2}		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE			Yes	Yes	Yes			
Country of origin FE				Yes	Yes			
Country of destination FE				Yes				
Administrative level 1 FE					Yes			
Country of origin × period FE						Yes	Yes	Yes
Administrative level 1 × period FE						Yes	Yes	
Destination cell FE (0.5° × 0.5°)							Yes	
Destination cell (0.5° × 0.5°) × period FE								Yes

Notes: The dependent variable, Age at Death, is the age at death of individuals in the FamiLinx database born in Europe and dying overseas. Temperature Distance is the absolute difference in mean annual temperature between origin and destination grid cells. Precipitation Distance is defined analogously. Geographic Distance is the great-circle distance between the centroids of origin and destination grid cells, expressed in 10^3 km. All specifications include a male indicator. Standard errors, reported in parentheses, are clustered at the country-of-origin × 2.5° × 2.5° destination-cell level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3. Climate Distance and Fertility of Settlers

Dep. var.:	Number of Children							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature Distance	-0.025	-0.043**	-0.034***	-0.050***	-0.036***	-0.032***	-0.037***	-0.042***
	(0.018)	(0.018)	(0.009)	(0.011)	(0.010)	(0.010)	(0.013)	(0.014)
Precipitation Distance	-0.007***	-0.007***	-0.000	0.001	-0.001	-0.000	-0.001	-0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Geographic Distance	0.208***	0.634***	-0.120*	0.110	-0.094	0.018	0.128	0.217
	(0.016)	(0.146)	(0.068)	(0.150)	(0.205)	(0.212)	(0.233)	(0.257)
Observations	108,573	108,573	108,573	108,562	108,488	108,224	107,786	106,638
R^2	0.028	0.031	0.091	0.096	0.103	0.110	0.136	0.159
Mean Temp. Distance	3.48	3.48	3.48	3.48	3.48	3.48	3.48	3.48
Mean Prec. Distance	29.04	29.04	29.04	29.04	29.04	29.04	29.04	29.04
Mean Geographic Distance	7.13	7.13	7.13	7.13	7.13	7.13	7.13	7.13
Mean Number of Children	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51
Geog. Dist ^{2}		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE			Yes	Yes	Yes			
Country of origin FE				Yes	Yes			
Country of destination FE				Yes				
Administrative level 1 FE					Yes			
Country of origin × period FE						Yes	Yes	Yes
Administrative level 1 × period FE						Yes	Yes	
Destination cell FE (0.5° × 0.5°)							Yes	
Destination cell (0.5° × 0.5°) × period FE								Yes

Notes: The dependent variable, Number of Children, is the number of children of individuals in the FamiLinx database born in Europe and dying overseas, observed only when strictly positive. Temperature Distance is the absolute difference in mean annual temperature between origin and destination grid cells. Precipitation Distance is defined analogously. Geographic Distance is the great-circle distance between the centroids of origin and destination grid cells, expressed in 10^3 km. Standard errors, reported in parentheses, are clustered at the country of origin × 2.5° × 2.5° destination-cell level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4. European Share and Climate Distance: Country-Level Analysis

Dep. var.:	European Share (P&W 2000)							
						Excluding Neo-Europes	Excluding Africa	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Temperature Distance	-0.031***	-0.038***	-0.033***	-0.041***	-0.018**	-0.017*	-0.039**	-0.015*
	(0.006)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.017)	(0.008)
Precipitation Distance	-0.000	-0.001	-0.000	-0.000	-0.000	0.001	-0.003	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Geographic Distance	0.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mean Temp. Dest.		0.017	0.013	0.022*	-0.001	0.006	0.015	0.001
		(0.013)	(0.012)	(0.011)	(0.011)	(0.012)	(0.021)	(0.009)
Mean Precip. Dest.		0.001	-0.001	-0.000	-0.000	-0.000	0.003**	-0.000
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Absolute Latitude		0.655**	0.713**	0.890***	0.474	0.495*	0.738	0.365
		(0.281)	(0.315)	(0.308)	(0.295)	(0.278)	(0.545)	(0.274)
Log Elevation			-0.027	-0.033	-0.020	-0.009	-0.004	-0.022
			(0.027)	(0.025)	(0.021)	(0.021)	(0.042)	(0.021)
Roughness			0.000	0.000*	0.000	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Slope			-0.110	-0.166*	-0.109	-0.073	-0.243	-0.084
			(0.085)	(0.092)	(0.088)	(0.095)	(0.153)	(0.089)
Soil PCA			-0.066**	-0.061**	-0.036*	-0.025	0.002	-0.026
			(0.029)	(0.030)	(0.021)	(0.022)	(0.028)	(0.019)
Malaria Index				0.001	0.002	0.002	0.004	0.001
				(0.003)	(0.003)	(0.003)	(0.028)	(0.003)
Pop. Density 1500				-0.001***	-0.001**	-0.001*	-0.001**	-0.001*
				(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Dummy British Legal Origin					-0.065	-0.068*	-0.088	-0.073*
					(0.039)	(0.037)	(0.080)	(0.038)
Asia					-0.300***	-0.315***	-0.179**	-0.263***
					(0.067)	(0.069)	(0.078)	(0.067)
Oceania					-0.018	-0.340**	0.089	0.009
					(0.187)	(0.145)	(0.269)	(0.168)
Africa					-0.301***	-0.287***		-0.236***
					(0.064)	(0.071)		(0.067)
Log Migrants								0.026**
								(0.010)
Observations	74	74	74	74	74	70	43	74
R^2	0.468	0.514	0.586	0.631	0.783	0.748	0.802	0.805
Mean Temp. Distance	12.85	12.85	12.85	12.85	12.85	12.85	12.85	12.85
Mean Precip. Distance	68.56	68.56	68.56	68.56	68.56	68.56	68.56	68.56
Mean European Share	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19

Notes: The sample includes former colonies that received European migrants between 1500 and 1850. The dependent variable is the European population share in 2000 from Putterman and Weil (2010). Temperature Distance is the migrant-weighted absolute difference in mean annual temperature between each destination and European origins, where weights correspond to migrant shares from each origin. Precipitation Distance and Geographic Distance (in 10^3 km) are constructed analogously. Column (1) includes only climate distances. Column (2) adds destination climate controls (mean temperature, precipitation, and absolute latitude). Column (3) adds geographic controls (log elevation, terrain ruggedness, slope, and soil PCA). Column (4) further includes controls for the disease environment and pre-colonial development (malaria ecology index from Kiszewski et al. (2004), and population density in 1500). Column (5) adds legal origin and continent fixed effects, with the Americas as the omitted category. Column (6) excludes Neo-European countries (United States, Canada, Australia, and New Zealand). Column (7) excludes African countries. Column (8) adds the log number of migrants.
Standard errors, reported in parentheses, are robust (HC1). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5. Climate Distance and Long-Run Development

Dep. var.:	Log(Luminosity + 0.01)				
	(1)	(2)	(3)	(4)	(5)
Temperature Distance	-0.116***	-0.112***	-0.098***	-0.061***	-0.057***
	(0.013)	(0.013)	(0.014)	(0.015)	(0.015)
Precipitation Distance	-0.013***	-0.013***	-0.011***	-0.012***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Geographic Distance	-0.229***	-0.228***	-0.221***	-0.077	0.078
	(0.011)	(0.011)	(0.013)	(0.050)	(0.056)
Mean Temp. Dest.	0.167***	0.172***	0.146***	0.139***	0.133***
	(0.014)	(0.015)	(0.016)	(0.017)	(0.024)
Mean Precip. Dest.	0.010***	0.011***	0.009***	0.008***	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Slope		-0.214***	-0.162**	-0.131**	0.012
		(0.068)	(0.068)	(0.065)	(0.077)
Log elevation		0.046	0.019	0.042	0.032
		(0.032)	(0.032)	(0.031)	(0.034)
Roughness		0.001*	-0.000	-0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Malaria Index			-0.129***	-0.081***	-0.016
			(0.013)	(0.025)	(0.055)
Log Pop. 1500 Dest.			0.151***	0.235***	0.411***
			(0.024)	(0.032)	(0.045)
Absolute Latitude	0.042***	0.046***	0.042***	0.033***	0.024
	(0.007)	(0.009)	(0.008)	(0.012)	(0.025)
Observations	4,781	4,780	4,656	4,640	4,404
R^2	0.269	0.272	0.305	0.361	0.488
Mean Temp. Distance	6.36	6.36	6.36	6.36	6.36
Mean Prec. Distance	38.79	38.79	38.79	38.79	38.79
Country of Destination FE				Yes	
Administrative level 1 FE					Yes

Notes: The sample includes destination grid cells at a 0.5° × 0.5° resolution that received European migrants between 1500 and 1850. The dependent variable is log(Luminosity + 0.01), where luminosity is nighttime light intensity from VIIRS 2015 data. Temperature Distance is the migrant-weighted absolute difference in mean annual temperature between each destination and European origins, where weights correspond to migrant shares from each origin. Precipitation Distance and Geographic Distance are constructed analogously. Geographic controls include slope, log(elevation + 0.01), and terrain ruggedness. Climate controls include mean temperature and precipitation at destination. Disease and pre-colonial controls include the malaria ecology index Kiszewski et al. (2004) and log(population density in 1500 + 0.01). Standard errors, reported in parentheses, are clustered at the 2.5° × 2.5° grid-cell level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6. Climate Distance and Democracy

	Polity2 (Polity IV)	V-Dem Polyarchy	Democracy (ANRR 2019)	Exec. Constraints (Polity IV)
	(1)	(2)	(3)	(4)
Temperature Distance	-0.668*	-0.022**	-0.025	-0.218*
	(0.355)	(0.010)	(0.030)	(0.117)
Precipitation Distance	0.014	0.001	0.001	0.005
	(0.026)	(0.001)	(0.002)	(0.008)
Geographic Distance (1000 km)	0.001	0.000*	0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Mean Temp. Dest.	-0.016	-0.010	0.015	0.082
	(0.429)	(0.012)	(0.028)	(0.138)
Mean Precip. Dest.	0.008	-0.001	-0.000	0.001
	(0.025)	(0.001)	(0.002)	(0.008)
Absolute Latitude	-2.703	-0.189	0.735	2.189
	(10.573)	(0.308)	(0.837)	(3.489)
Log Elevation	-1.333*	-0.067**	-0.049	-0.390*
	(0.746)	(0.029)	(0.061)	(0.226)
Roughness	-0.010	-0.001**	-0.001	-0.003
	(0.009)	(0.000)	(0.001)	(0.002)
Slope	4.188	0.245**	0.326	1.765*
	(3.126)	(0.098)	(0.292)	(0.917)
Soil PCA	0.884	-0.003	-0.018	0.223
	(0.771)	(0.023)	(0.070)	(0.234)
Malaria Index	0.272**	0.012**	0.019*	0.067*
	(0.124)	(0.005)	(0.011)	(0.039)
Pop. Density 1500	0.006	0.000	0.002	0.001
	(0.020)	(0.001)	(0.002)	(0.007)
Dummy British Legal Origin	2.131	0.068	0.185	0.878*
	(1.560)	(0.054)	(0.121)	(0.482)
Asia	-4.089*	-0.188***	-0.442**	-1.378**
	(2.191)	(0.069)	(0.192)	(0.616)
Oceania	-9.968**	-0.391*	-0.747*	-3.635**
	(4.272)	(0.196)	(0.434)	(1.530)
Africa	-3.839*	-0.177***	-0.279*	-1.010
	(2.208)	(0.066)	(0.159)	(0.662)
Observations	71	73	74	69
R^2	0.480	0.592	0.366	0.493
Mean Temp. Distance	12.85	12.85	12.85	12.85
Mean Precip. Distance	68.56	68.56	68.56	68.56
Mean Dep. Var.	4.17	0.51	0.66	5.00

Notes: The sample includes former colonies that received European migrants between 1500 and 1850. Each column reports the full-controls specification (Column 5 from Table 4) for a different democracy outcome measured in 2010. Polity2 and executive constraints are from Polity IV; the polyarchy index is from V-Dem; the democracy indicator is from Acemoglu et al. (2019). Temperature Distance is the migrant-weighted average absolute temperature difference between destination and European origins. Standard errors, reported in parentheses, are robust (HC1). Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix

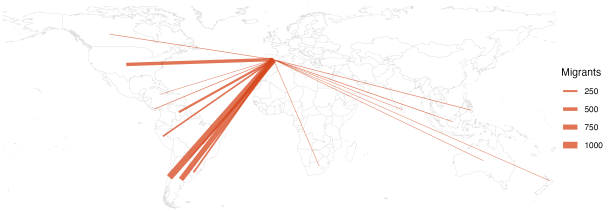
Table of Contents

A Additional Table and Figures **2**

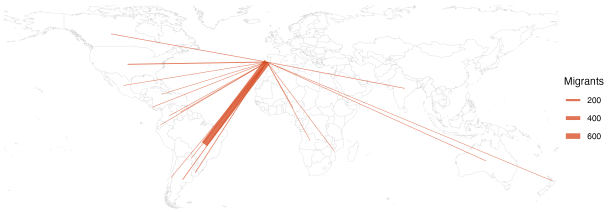
A Additional Table and Figures

Figure A.1. Migration Flows from European Origins to Colonial Destinations

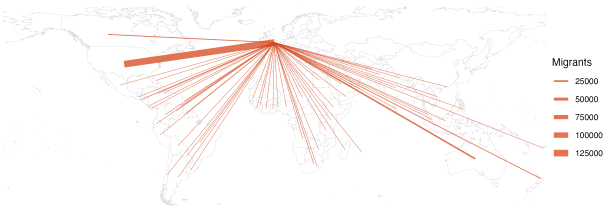
(a) Spain



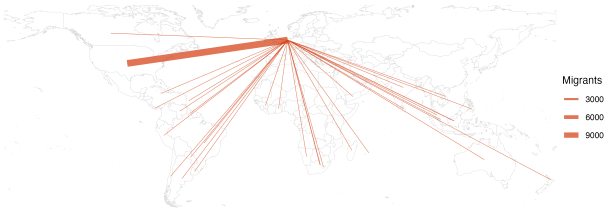
(b) Portugal



(c) Great Britain

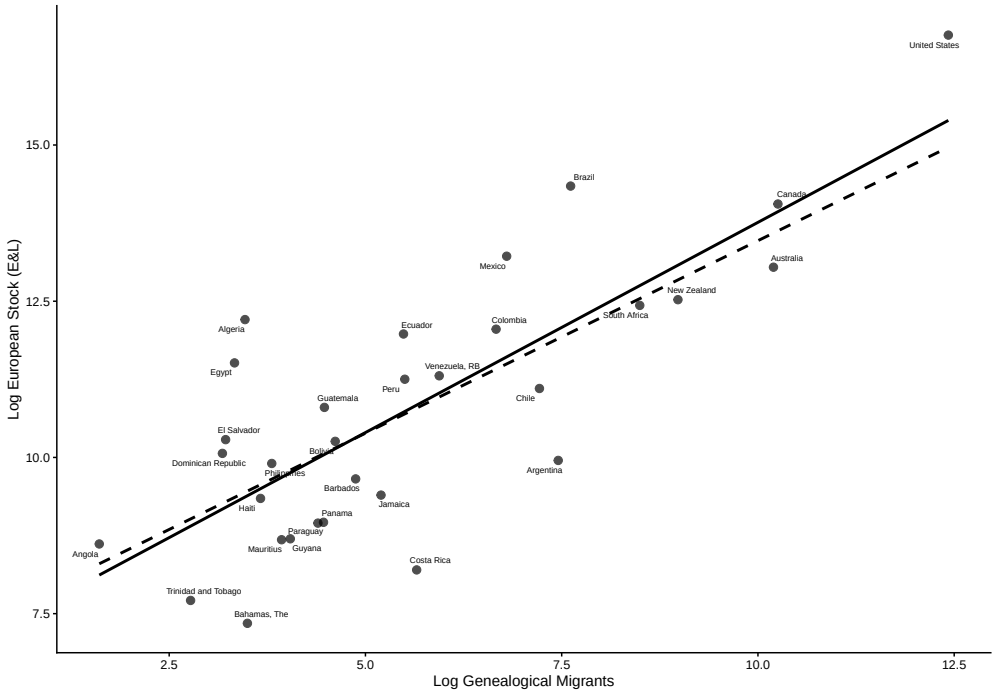


(d) The Netherlands



Notes: The figure displays the geographic distribution of settlers across destination countries by origin country. Lines connect origin capitals (Madr Lisbon, London, and Amsterdam) to destination countries. Line thickness is proportional to the number of migrants in the genealogical sample. Par (a)–(d) correspond to Spain, Portugal, Great Britain, and the Netherlands.

Figure A.2. Validation: Genealogical Migrants vs. Historical European Population

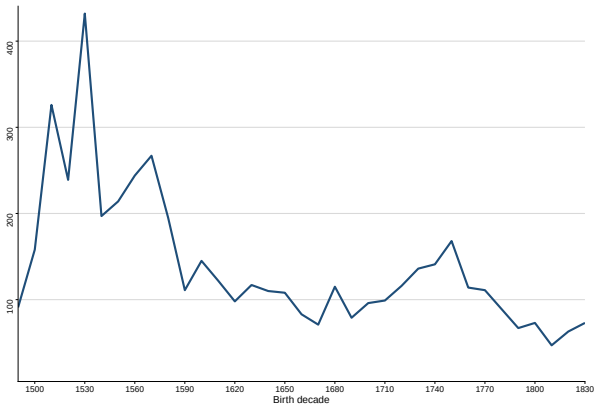


Notes: The figure plots the log number of European settlers in our genealogical data (1500–1850) against the log of the average European populat reported by Easterly and Levine (2016) over the period 1500–1900. Each point represents a former colony. The solid line shows the linear fit the full sample, while the dashed line excludes the Neo-Europes—namely, the United States, Canada, Australia, and New Zealand. The estima regression coefficient (robust standard errors) is 0.673 (0.089) for the full sample and 0.616 (0.164) for the sample excluding Neo-Europes.

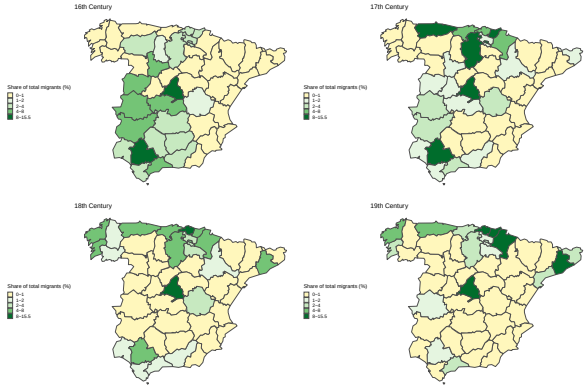
Figure A.3. Comparison of Spanish Migration Patterns: Genealogical Data vs. Historical Records

Genealogical Data

(A) Temporal Distribution



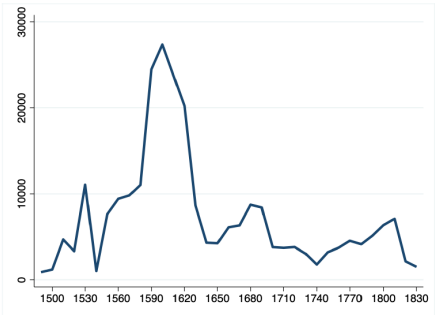
(B) Regional Origins



Historical Records (Arroyo Abad and Espín-Sánchez, 2025)

(C) Temporal Distribution

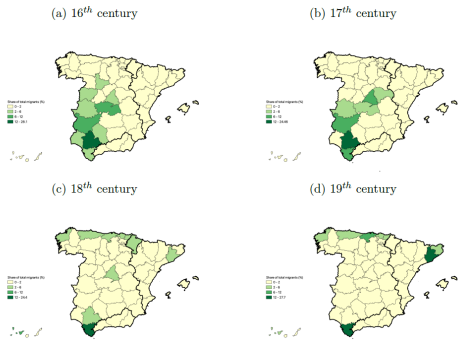
Figure 3: Total migrant flows per decade, 1490-1830



Notes: see Section 3 for details.

(D) Regional Origins

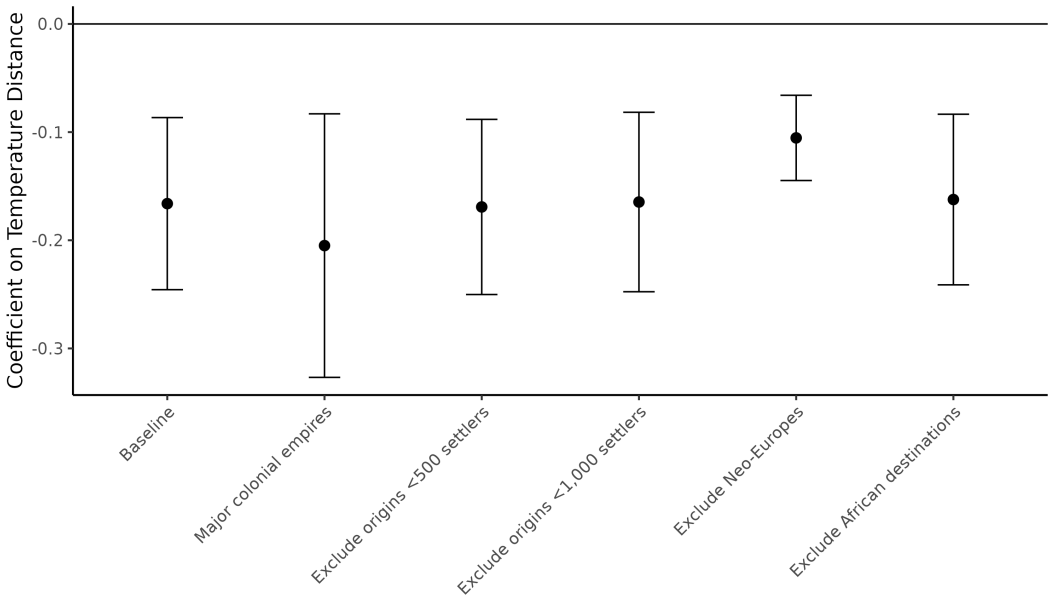
Figure 5: Migration share by province.



Notes: 16^{th} century includes 1492-1599, the 19^{th} century includes 1800-1830. Share of total migrant flows—with known origin—for the period.

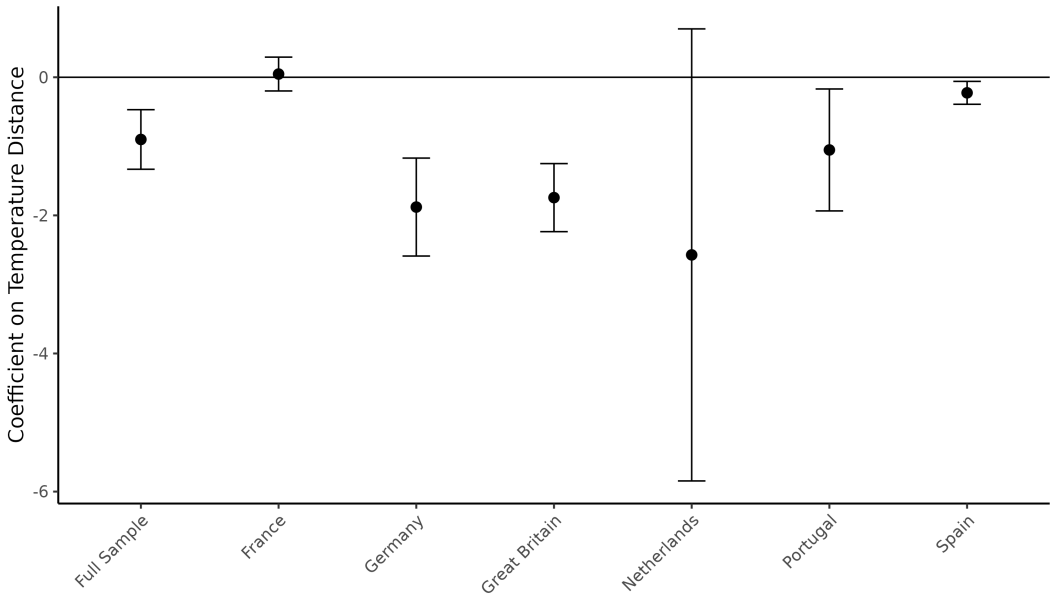
Notes: This figure compares Spanish migration patterns from our genealogical data (Panels A–B) with historical records from Arroyo Abad a Espín-Sánchez (2025) (Panels C–D). Panel A displays the number of Spanish-born individuals by decade of birth, while Panel B maps their regio origins within Spain. Panels C and D report the corresponding patterns using archival data from Arroyo Abad and Espín-Sánchez (2025).

Figure A.4. Robustness: Climate Matching Across Alternative Samples



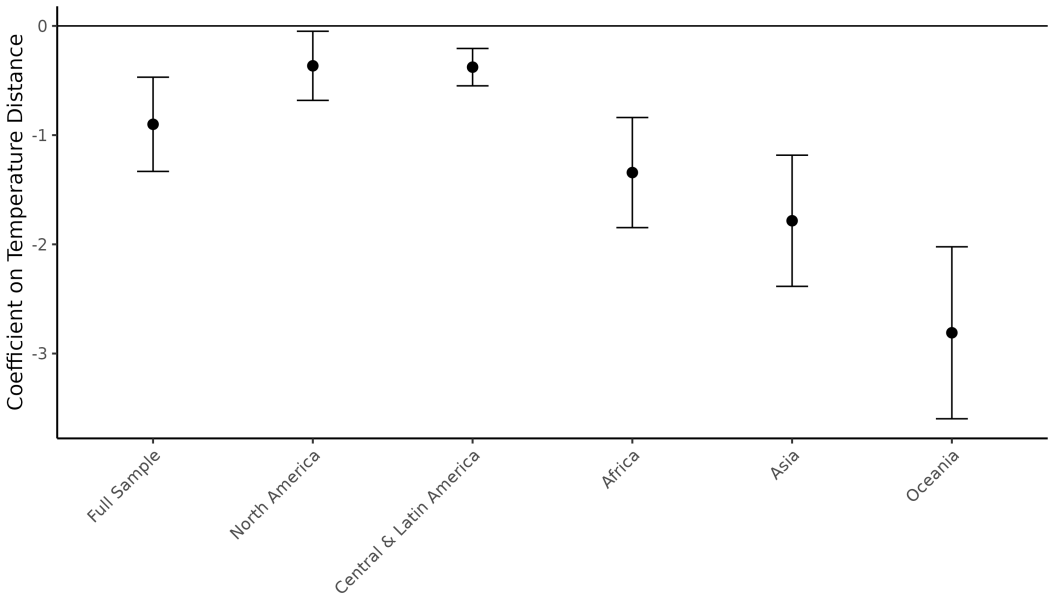
Notes: This figure reports estimates from the preferred specification (column 8 of Table 1) across different sample specifications. The basel specification is compared with samples that: (1) include only major colonial empires (France, Germany, Great Britain, Spain, Portugal, and Netherlands); (2) exclude origin cells with fewer than 500 settlers; (3) exclude origin cells with fewer than 1,000 settlers; (4) exclude Neo-Europe destinations (the United States, Canada, Australia, and New Zealand); and (5) exclude African destinations. The specification includes precipitat distance, geographic distance, geographic distance squared, Period × Country of Origin FE, and Period × 0.5° Destination Cell FE. Confide intervals at the 95% level are reported. Standard errors are clustered at the country-of-origin × 2.5° × 2.5° destination-cell level.

Figure A.5. Climate Distance and Migration Flows by Major European Colonial Empire



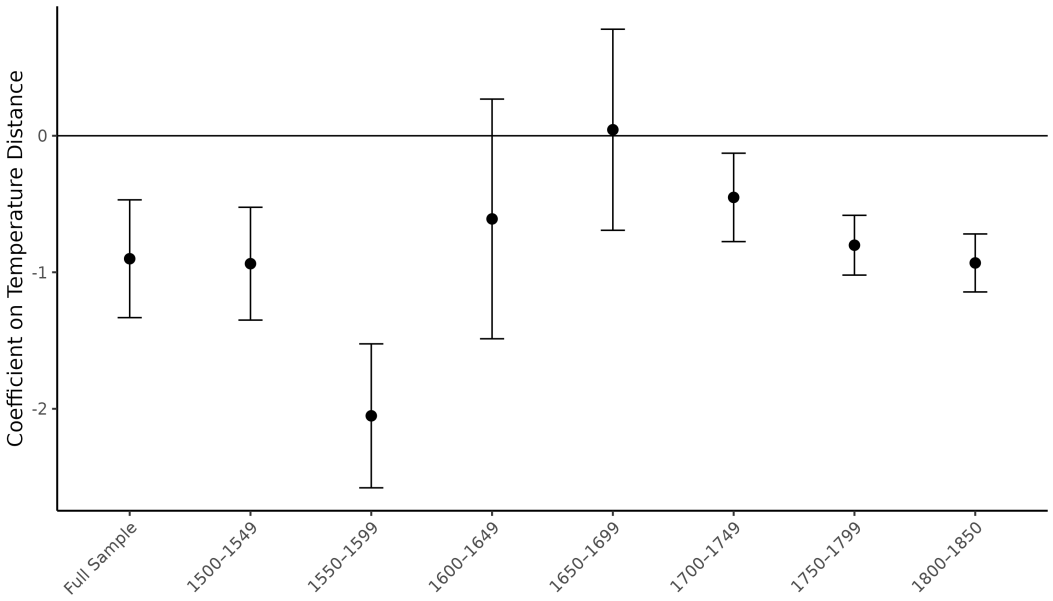
Notes: This figure reports estimates from the preferred specification (column 8 of Table 1) separately by major European colonial empire. To facilit comparisons, temperature distance is standardized within each subsample to have mean zero and standard deviation one. We focus on the six princi colonial empires: Spain, Portugal, France, Great Britain, Germany, and the Netherlands. The specification includes precipitation distance, geograp distance, geographic distance squared, Period × Country of Origin FE, and Period × 0.5° Destination Cell FE. Confidence intervals at the 95% le are reported. Standard errors are clustered at the country-of-origin × 2.5° × 2.5° destination-cell level.

Figure A.6. Climate Distance and Migration Flows by Continent of Destination



Notes: This figure reports estimates from the preferred specification (column 8 of Table 1) separately by continent of destination. To facilit comparisons, temperature distance is standardized within each continent subsample to have mean zero and unit standard deviation. The full-sam estimate (shown on the left) is negative and precisely estimated. Across all periods, the coefficient on temperature distance remains negative, thou its magnitude varies. The specification includes precipitation distance, geographic distance and its square, period × country-of-origin fixed effects, a period × 0.5° destination-cell fixed effects. Confidence intervals at the 95% level are reported. Standard errors are clustered at the country-of-ori × 2.5° × 2.5° destination-cell level.

Figure A.7. Climate Distance and Migration Flows by 50-Year Period



Notes: This figure reports estimates from the preferred specification (column 8 of Table 1) separately by 50-year period. To facilitate comparisc temperature distance is standardized within each period subsample to have mean zero and standard deviation one. The specification inclu precipitation distance, geographic distance, geographic distance squared, Period × Country of Origin FE, and Period × 0.5° Destination Cell FE. 9 confidence intervals shown. Confidence intervals at the 95% level are reported. Standard errors are clustered at the country-of-origin × 2.5° × 2 destination-cell level.

Table A.1. Number of Individuals by Country of Origin (1500–1850)

Country of Origin	N	%	Country of Origin	N	%
Great Britain	165,960	50.59	Estonia	205	0.06
Germany	57,433	17.51	Bosnia and Herzegovina	174	0.05
Ireland	27,075	8.25	Belarus	125	0.04
France	24,621	7.50	Romania	115	0.04
Netherlands	12,717	3.88	Russia	109	0.03
Norway	8,193	2.50	Latvia	106	0.03
Switzerland	6,516	1.99	Slovakia	86	0.03
Spain	5,271	1.61	Greece	69	0.02
Sweden	4,677	1.43	Croatia	38	0.01
Denmark	3,501	1.07	Bulgaria	31	0.01
Poland	3,015	0.92	Liechtenstein	21	0.01
Belgium	1,508	0.46	Moldova	20	0.01
Italy	1,281	0.39	Slovenia	19	0.01
Czech Republic	1,055	0.32	Republic of Serbia	14	0.00
Portugal	904	0.28	Malta	9	0.00
Austria	824	0.25	Macedonia	6	0.00
Luxembourg	744	0.23	Montenegro	5	0.00
Finland	522	0.16	Andorra	3	0.00
Lithuania	436	0.13	Albania	2	0.00
Ukraine	350	0.11	San Marino	2	0.00
Hungary	304	0.09	Kosovo	1	0.00
			Total	328,067	100%

Table A.2. Climate Distance and Migration Flows: Alternative Functional Forms

Dep. var.:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Temperature Distance	-0.166***	-0.167***		-0.104***	-0.106***
	(0.041)	(0.036)		(0.023)	(0.022)
Precipitation Distance	-0.002	-0.002		-0.009***	-0.009***
	(0.003)	(0.002)		(0.001)	(0.001)
Geographic Distance	1.145				
	(0.755)				
Log(Temperature Distance)			-0.583***		
			(0.129)		
Log(Precipitation Distance)			-0.000		
			(0.059)		
Log(Geographic Distance)		6.489***	6.369***		
		(1.572)	(1.380)		
Inland Origin Distance				-0.506**	
				(0.243)	
Sea Travel Distance				0.245*	
				(0.143)	
Log(Inland Origin Distance)					-0.679*
					(0.347)
Log(Sea Travel Distance)					2.900**
					(1.249)
Observations	13,786,220	13,786,220	13,786,220	9,702,235	9,702,235
R^2	0.133	0.139	0.153	0.051	0.052
Mean Temp. Distance	7.12	7.12	7.12	7.12	7.12
Mean Prec. Distance	38.87	38.87	38.87	38.87	38.87
Geog. Dist ^{2}	Yes				
Country of origin × period FE	Yes	Yes	Yes	Yes	Yes
Destination cell (0.5° × 0.5°) × period FE	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes all origin–destination pairs formed by European grid cells (origins) and overseas grid cells (destinations) at a 0.5° resolution, aggregated into 50-year periods between 1500 and 1850. The dependent variable, Number of Migrants, is the count of individuals in the FamiLinx database born in Europe and dying overseas, collapsed to the origin–destination–period cell. Column (1) reproduces the baseline specification to ease comparisons. The baseline already includes geographic distance and its square. Column (2) replaces these terms with the logarithm of geographic distance. Column (3) additionally log-transforms temperature and precipitation distance, consistent with the standard gravity specification (Anderson and Van Wincoop, 2003). Column (4) replaces geographic distance with three separate travel-cost components described in Section 3.4: inland distance from origin to coast, oceanic crossing distance, and inland distance from coast to destination. Column (5) takes the logarithm of these three travel-cost components. Standard errors, reported in parentheses, are clustered at the country of origin by 2.5° destination cell level. Significance levels: $***p < 0.01, **p < 0.05, *p < 0.1$

Table A.3. Climate Distance and Migration Flows: Geographic Controls

Dep. var.:	Number of Migrants					
	(1)	(2)	(3)	(4)	(5)	(6)
Temperature Distance	-0.166***	-0.215***	-0.120***	-0.114***	-0.103***	-0.104***
	(0.041)	(0.028)	(0.033)	(0.033)	(0.032)	(0.033)
Precipitation Distance	-0.002	-0.002	-0.000	-0.000	-0.001	-0.001
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Geographic Distance	1.145	1.787*	0.464	0.507	0.548	0.536
	(0.755)	(0.934)	(0.607)	(0.605)	(0.597)	(0.594)
Latitude Distance			-0.110***	-0.112***	-0.114***	-0.109***
			(0.018)	(0.018)	(0.018)	(0.017)
Ruggedness Distance				-0.001*	-0.000	-0.000
				(0.001)	(0.001)	(0.001)
Elevation Distance					-0.088***	-0.111***
					(0.015)	(0.015)
Both Coastal						-0.514***
						(0.185)
Observations	13,786,220	13,017,418	13,786,220	13,786,220	13,786,220	13,786,220
R^2	0.133	0.157	0.137	0.142	0.140	0.153
Mean Temp. Distance	7.12	7.12	7.12	7.12	7.12	7.12
Mean Prec. Distance	38.87	38.87	38.87	38.87	38.87	38.87
Mean Latitude Distance	32.48	32.48	32.48	32.48	32.48	32.48
Mean Elevation Distance	4.00	4.00	4.00	4.00	4.00	4.00
Mean Ruggedness Distance	138.52	138.52	138.52	138.52	138.52	138.52
Mean Both Coastal	0.20	0.20	0.20	0.20	0.20	0.20
Geog. Dist ^{2}	Yes	Yes	Yes	Yes	Yes	Yes
Country of origin × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination cell (0.5° × 0.5°) × period FE	Yes	Yes	Yes	Yes	Yes	Yes
Country of origin × Country of destination FE		Yes				

Notes: The sample includes all origin–destination pairs formed by European grid cells (origins) and overseas grid cells (destinations) at a 0.5° resolution, aggregated into 50-year periods between 1500 and 1850. The dependent variable, Number of Migrants, is the count of individuals in the FamiLinx database born in Europe and dying overseas, collapsed to the origin–destination–period cell. Temperature Distance is the absolute difference in average temperature between origin and destination grid cells. Precipitation Distance is the absolute difference in average precipitation. Geographic Distance is the great-circle distance between the centroids of the origin and destination grid cells, expressed in 10^3 km (thousands of km). Column 1 presents the baseline specification. Column 2 adds stringent origin-country by destination-country fixed effects, which absorb all time-invariant bilateral factors specific to a given country of origin–country of destination pair (such as historical ties, language proximity, or persistent migration networks). Column 3 adds latitude distance. Column 4 adds ruggedness distance. Column 5 adds elevation distance. Column 6 introduces a dummy indicating whether both origin and destination have coastal access.[18] The main results remain robust throughout. All specifications include Period × Country of Origin FE and Period × 0.5° Destination Cell FE. Standard errors, reported in parentheses, are clustered at the country of origin by 2.5° destination cell level. Significance levels: $***p < 0.01, **p < 0.05, *p < 0.1$

Table A.4. Climate Distance and Migration Flows: Soil Characteristics Robustness

Dep. var.:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Temperature Distance	-0.166***	-0.161***	-0.166***	-0.174***	-0.167***
	(0.041)	(0.043)	(0.041)	(0.040)	(0.039)
Precipitation Distance	-0.002	-0.002	-0.002	-0.003	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Geographic Distance	1.145	1.110	1.201	1.262*	1.355*
	(0.755)	(0.776)	(0.770)	(0.735)	(0.772)
Soil Distance (Raw)		-0.060			
		(0.084)			
Soil Distance (Resid. Levels)			0.122**		
			(0.051)		
Soil Distance (Resid. Distances)				0.130	
				(0.093)	
Clay Distance (Resid.)					0.050
					(0.043)
Organic Matter Distance (Resid.)					0.172***
					(0.032)
pH Distance (Resid.)					-0.011
					(0.021)
Water Content Distance (Resid.)					-0.284***
					(0.058)
Observations	13,786,220	13,747,859	13,344,163	13,747,859	13,344,163
R^2	0.133	0.133	0.136	0.128	0.140
Mean Temp. Distance	7.12	7.12	7.12	7.12	7.12
Mean Precip. Distance	38.87	38.87	38.87	38.87	38.87
Geog. Dist ^{2}	Yes	Yes	Yes	Yes	Yes
Country of origin × period FE	Yes	Yes	Yes	Yes	Yes
Destination cell (0.5° × 0.5°) × period FE	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes all origin–destination pairs formed by European grid cells (origins) and overseas grid cells (destinations) at a 0.5° resolution, aggregated into 50-year periods between 1500 and 1850. The dependent variable, Number of Migrants, is the count of individuals in the FamiLinx database born in Europe and dying overseas, collapsed to the origin–destination–period cell. Column 1 reproduces the baseline specification from the main robustness table. Column 2 adds raw soil distance, calculated as the Euclidean distance across four standardized soil characteristics: clay content, organic matter, pH, and water content. Column 3 uses soil distance computed from residuals after regressing soil characteristics on climate levels. Column 4 uses soil distance computed from residuals after regressing soil characteristic distances on climate distances. Column 5 includes all four individual residualized soil distances separately (clay, organic matter, pH, water content). The coefficient on temperature distance remains negative and statistically significant across all specifications, indicating that the climate matching result is robust to controlling for soil similarity between origin and destination. Standard errors, reported in parentheses, are clustered at the country of origin by 2.5° destination cell level. Significance levels: $***p < 0.01, **p < 0.05, *p < 0.1$

Table A.5. Climate Distance and Migration Flows: Alternative Climate Measures

Dep. var.:	Number of Migrants				
	(1)	(2)	(3)	(4)	(5)
Temperature Distance	-0.166***	-0.158***			
	(0.041)	(0.040)			
Precipitation Distance	-0.002	-0.002			
	(0.003)	(0.003)			
Temperature Range Distance		-0.078***			
		(0.014)			
Precipitation Range Distance		-0.005**			
		(0.002)			
Minimum Temperature Distance			-0.072***		-0.088***
			(0.015)		(0.013)
Maximum Temperature Distance				-0.110***	-0.144***
				(0.038)	(0.040)
Minimum Precipitation Distance			-0.001		0.003
			(0.003)		(0.003)
Maximum Precipitation Distance				-0.009***	-0.008***
				(0.002)	(0.001)
Geographic Distance	1.145	0.354	0.847	-0.074	-0.453
	(0.755)	(0.790)	(0.804)	(0.653)	(0.638)
Observations	13,786,220	13,786,220	13,828,512	13,828,512	13,828,512
R^2	0.133	0.137	0.117	0.120	0.123
Mean Temp. Distance	7.12	7.12	7.12	7.12	7.12
Mean Precip. Distance	38.87	38.87	38.87	38.87	38.87
Mean Temp. Range Distance	14.34	14.34	14.34	14.34	14.34
Mean Precip. Range Distance	50.58	50.58	50.58	50.58	50.58
Mean Min Temp. Distance	12.60	12.60	12.60	12.60	12.60
Mean Max Temp. Distance	6.47	6.47	6.47	6.47	6.47
Geog. Dist ^{2}	Yes	Yes	Yes	Yes	Yes
Country of origin × period FE	Yes	Yes	Yes	Yes	Yes
Destination cell (0.5° × 0.5°) × period FE	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes all origin–destination pairs formed by European grid cells (origins) and overseas grid cells (destinations) at a 0.5° resolution, aggregated into 50-year periods between 1500 and 1850. The dependent variable, Number of Migrants, is the count of individuals in the FamiLinx database born in Europe and dying overseas, collapsed to the origin–destination–period cell. Column 1 reproduces the baseline specification with mean annual temperature and precipitation distance. Column 2 adds temperature range distance (July–January difference) and precipitation range distance. Column 3 uses only minimum temperature distance (January) and minimum precipitation distance, replacing the mean measures. Column 4 uses only maximum temperature distance (July) and maximum precipitation distance. Column 5 includes both minimum and maximum temperature and precipitation distances (without the mean measures). The coefficients on seasonal temperature distances remain negative and statistically significant. Standard errors, reported in parentheses, are clustered at the country of origin by 2.5° destination cell level. Significance levels: $***p < 0.01, **p < 0.05, *p < 0.1$

Table A.6. Climate Distance and Mortality: Progressive Destination Controls

Dep. var.:	Age at Death				
	(1)	(2)	(3)	(4)	(5)
Temperature Distance	-0.170***	-0.144***	-0.103**	-0.102**	-0.110**
	(0.050)	(0.049)	(0.046)	(0.046)	(0.046)
Precipitation Distance	-0.001	-0.001	-0.001	-0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Geographic Distance	-1.292*	-0.448	-1.242	-1.232	-1.300
	(0.775)	(0.746)	(0.826)	(0.825)	(0.830)
Mean Temperature (d)		-0.295***	-0.122**	-0.121**	-0.101
		(0.066)	(0.059)	(0.059)	(0.061)
Mean Precipitation (d)		-0.002	0.001	0.001	-0.000
		(0.007)	(0.007)	(0.007)	(0.007)
Slope (d)			-0.052	-0.053	-0.053
			(0.270)	(0.270)	(0.272)
Log Elevation (d)			0.460***	0.461***	0.475***
			(0.118)	(0.118)	(0.116)
Ruggedness (d)			0.001	0.001	0.001
			(0.002)	(0.002)	(0.002)
Malaria Index (d)				-0.242	-0.185
				(0.473)	(0.474)
Log Pop 1500 (d)					-0.745
					(0.622)
Observations	312,326	312,326	312,316	312,316	312,155
R^2	0.076	0.076	0.077	0.077	0.077
Mean Temp. Distance	3.34	3.34	3.34	3.34	3.34
Mean Prec. Distance	29.23	29.23	29.23	29.23	29.23
Mean Age at Death	67.41	67.41	67.41	67.41	67.41
Geog. Dist ^{2}	Yes	Yes	Yes	Yes	Yes
Country of origin × period FE	Yes	Yes	Yes	Yes	Yes
Administrative level 1 × period FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable, Age at Death, is the age of death of individuals in the FamiLinx database born in Europe and dying overseas. Temperature Distance is the absolute difference in average temperature between origin and destination. Precipitation Distance is the absolute difference in average precipitation. Geographic Distance is the great-circle distance, expressed in 10^3 km. (d) indicates destination-level controls. All specifications include Period × Country of Origin FE, Period × Administrative Level 1 FE, and individual male dummy. Standard errors, reported in parentheses, are clustered at the country of origin by 2.5° × 2.5° destination-cell level. Significance levels: $***p < 0.01, **p < 0.05, *p < 0.1$

Table A.7. Climate Distance and Long-Run Development: Robustness Checks

Dep. var.:	Log(Luminosity + 0.01)						
	Baseline	Temp Bins	Log(N Mig.)	Log(PopD 2015)	Excl. Neo-Eur.	Excl. Africa	Min 50 Mig.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Temperature Distance	-0.057***	-0.048***	-0.037***	-0.019	-0.040*	-0.055***	-0.181***
	(0.015)	(0.016)	(0.014)	(0.012)	(0.024)	(0.016)	(0.060)
Precipitation Distance	-0.006***	-0.005***	-0.003***	-0.002**	-0.006**	-0.006***	-0.020***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.006)
Geographic Distance	0.078	0.101*	0.012	-0.009	0.074	0.079	0.128
	(0.056)	(0.055)	(0.051)	(0.049)	(0.070)	(0.070)	(0.225)
Mean Temp. Dest.	0.133***		0.119***	0.086***	0.047	0.139***	0.353***
	(0.024)		(0.023)	(0.020)	(0.043)	(0.025)	(0.079)
Temperature Bins at Destination (Reference: 15-20°C)							
Temp < 0°C		-0.891*					
		(0.468)					
Temp 0-10°C		-0.862***					
		(0.221)					
Temp 10-15°C		-0.529***					
		(0.133)					
Temp 20-25°C		-0.200					
		(0.178)					
Temp > 25°C		-0.541*					
		(0.316)					
Mean Precip. Dest.	0.004**	0.004**	0.003*	-0.001	0.004	0.004**	0.020***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.005)
Slope	0.012	-0.042	-0.001	0.009	0.083	0.006	0.051
	(0.077)	(0.075)	(0.070)	(0.062)	(0.107)	(0.082)	(0.231)
Log elevation	0.032	-0.021	0.078**	0.196***	0.056	0.038	0.189***
	(0.034)	(0.034)	(0.033)	(0.032)	(0.063)	(0.036)	(0.069)
Roughness	0.000	0.000	0.000	-0.000	0.000	0.001	0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)
Malaria Index	-0.016	0.013	-0.002	0.014	-0.051	-0.058	1.875**
	(0.055)	(0.051)	(0.056)	(0.047)	(0.069)	(0.072)	(0.821)
Log Pop. 1500 Dest.	0.411***	0.435***	0.280***		0.333***	0.433***	0.390***
	(0.045)	(0.046)	(0.040)		(0.088)	(0.048)	(0.092)
Log(N Migrants)			0.305***				
			(0.021)				
Log(Pop. Density 2015)				0.418***			
				(0.016)			
Absolute Latitude	0.024	-0.046**	0.033	0.021	-0.053	0.036	0.323***
	(0.025)	(0.023)	(0.023)	(0.020)	(0.048)	(0.027)	(0.084)
Observations	4,404	4,404	4,404	4,500	974	4,170	865
R^2	0.488	0.486	0.533	0.584	0.480	0.485	0.492
Mean Temp. Distance	6.36	6.36	6.36	6.36	6.36	6.36	6.36
Mean Prec. Distance	38.79	38.79	38.79	38.79	38.79	38.79	38.79

Notes: This table presents robustness checks for the main specification. All columns include the full set of controls and ADM1 fixed effects (column 5 from the main table). Column (1) replicates the baseline results. Column (2) replaces mean temperature with temperature bins at destination. Column (3) adds a control for log(number of migrants) to the destination cell. Column (4) adds a control for log(population density in 2015 + 0.01). Column (5) excludes neo-Europe destinations (USA, Canada, Australia, New Zealand). Column (6) excludes all African destinations. Column (7) restricts to cells that received at least 50 European migrants. The dependent variable is Log(Luminosity + 0.01), where luminosity is nighttime light intensity from VIIRS 2015 data. Temperature Distance is the weighted average absolute temperature difference between destination and European origins, where weights equal the number of migrants from each origin. Geographic Controls include slope, log(elevation + 0.01), and roughness. Climate Controls include mean temperature at destination (or temperature bins in column 2) and mean precipitation. Disease & Precolonial includes malaria ecology index (Kiszewski et al., 2004) and log(population density 1500). Standard errors, reported in parentheses, are clustered at the 2.5° × 2.5° grid cell level. Significance levels: $***p < 0.01, **p < 0.05, *p < 0.1$