NBER WORKING PAPER SERIES

ROOTS OF INEQUALITY

Oded Galor
Marc Klemp
Daniel C. Wainstock

Working Paper 31580
http://www.nber.org/papers/w31580

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2023, Revised November 2025

The authors wish to thank Carl-Johan Dalgaard, Andrew Foster, Luigi Guiso, Peter Hull, Fabio Mariani, Stelios Michalopoulos, Luca Pensieroso, Louis Putterman, and Uwe Sunde, as well as participants at the conferences "Deep Rooted Factors in Comparative Economic Development," Inequality, the Design of Markets and Redistributions (Villa Mondragone, Rome), and WIR (Harvard University), and at seminars held at AUEB (Athens), Brown University, the University of Copenhagen, the Hebrew University of Jerusalem, LUISS, EIEF (Rome), Ludwig Maximilian University, the University of Luxembourg, Pompeu Fabra University, the University of Oxford, and Tel-Aviv University, for helpful discussions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Roots of Inequality
Oded Galor, Marc Klemp, and Daniel C. Wainstock
NBER Working Paper No. 31580
August 2023, Revised November 2025
JEL No. O10, Z10

ABSTRACT

Why does inequality vary across societies? Why are some societies more unequal than others? We advance the hypothesis that in a market economy, where income differentials reflect variations in productive traits, a significant share of cross-societal differences in inequality may reflect enduring variation in the degree of diversity within societies, rooted in the prehistoric Out-of-Africa migration. Patterns of inequality within the U.S. population are consistent with this hypothesis, suggesting that disparities among groups originating from different ancestral societies are associated with the degree of diversity in productive traits within those societies, shaped during humanity's dispersal from Africa. Consistent with the proposed mechanism, populations whose ancestors originated closer to East Africa tend to exhibit greater dispersion in productive traits—education, ability, and labor supply—channels that appear to mediate the relationship between ancestral diversity and inequality.

Oded Galor
Brown University
Department of Economics
and CEPR
and also NBER
oded_galor@brown.edu

Marc Klemp
Brown University
Department of Economics
and Brown University
marc.klemp@econ.ku.dk

Daniel C. Wainstock
Brown University
daniel_crisostomo_wainstock@brown.edu

1 Introduction

Global inequality has widened significantly in recent decades, as the share of income held by the top 1% of the world population has reached 19%.[1] This pronounced disparity is echoed by an equally striking pattern — a profound variation in the degree of inequality across societies (Figure 1). Why does inequality vary between countries and regions? Why are some societies considerably more unequal than others?

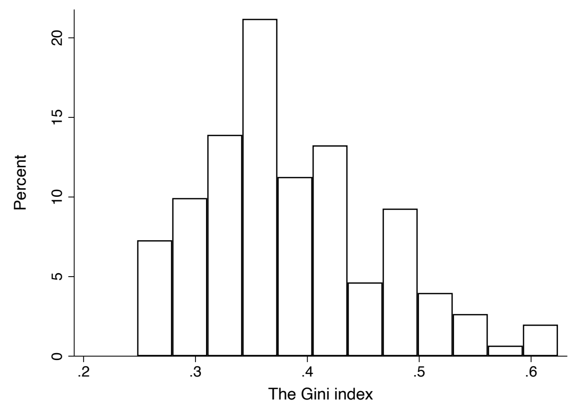


Figure 1: Differences in Income Inequality Across Countries

Notes: This figure depicts the histogram of the distribution of the Gini index of income inequality across countries during the time period 2000-2020 (World Bank Development Indicators).

Conventional wisdom suggests that cross-country variations in the degree of income inequality stem from the extent to which differences in market institutions, cultural predispositions toward egalitarianism, and the prevalence of inequality-mitigating policies affect the rewards to productive traits. These structural and cultural factors have operated alongside broader global forces—most notably technological change and globalization—whose unequal diffusion has further amplified cross-country disparities in inequality.[2]

Indeed, the mapping from productive traits to income inequality is more compressed in societies characterized by inequality-mitigating institutions or cultural traits, whereas inequality tends to be more pronounced in societies primarily driven by market forces (Figure 2(a)). Yet this prevailing perspective neglects a critical factor: the distribution of productive traits itself may vary systematically across societies, contributing significantly to cross-societal variation in income inequality, independently of cultural, institutional, or technological influences (Figure 2(b)).

[1] Chancel et al. (2022).

[2] Rosen (1981), Katz and Murphy (1992), Galor and Zeira (1993), Aghion et al. (1999), Galor and Moav (2000), Benabou (2000), De La Croix and Doepke (2003), Klor and Shayo (2010), and Piketty (2014).

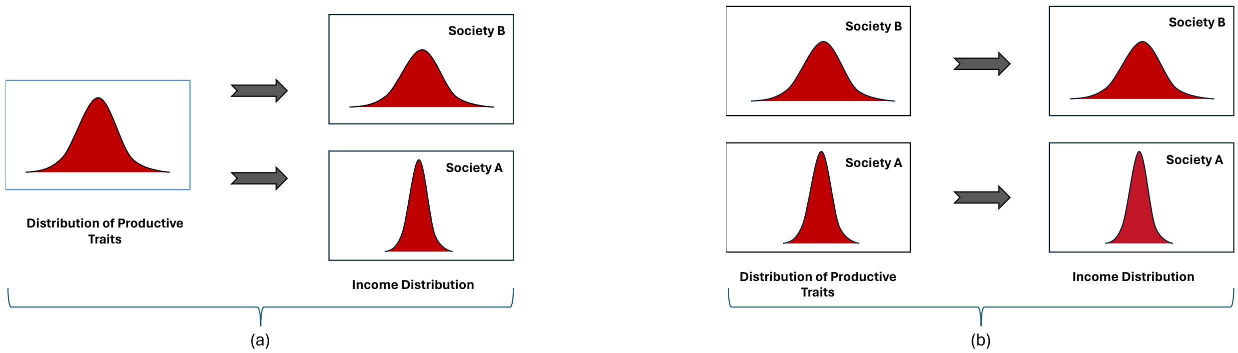


Figure 2: The Distribution of Productive Traits & Income Inequality

Notes: This figure illustrates two sources of inequality across societies. Panel (a) depicts how differences in inequality-mitigating cultural and institutional characteristics across societies that share an identical distribution of productive traits yield different degrees of income inequality within societies. Panel (b) shows how differences in the diversity of productive traits across societies yield different degrees of income inequality within each society, under otherwise identical cultural and institutional conditions.

This paper advances the novel hypothesis that, in a market economy where earning differentials reflect variations in productive traits, systematic differences in the diversity of these traits across societies, shaped by the prehistoric Out of Africa migration, have contributed significantly to the observed cross-societal differences in income inequality.

The prehistoric migration of Homo sapiens out of Africa represents one of the most transformative chapters in human history, laying the groundwork for the development of human settlements across the globe. The serial nature of this dispersal, amplified by population bottlenecks and genetic drift, led to a progressive reduction in diversity among populations that settled at greater migratory distances from Africa—a phenomenon known as the *Serial Founder Effect*. As humans migrated farther from their African origins, societies formed by their descendants exhibited a marked decline in cultural, linguistic, behavioral, and phenotypic diversity (Galor et al., 2024). Notably, this pattern is evident in the dispersion of predispositions toward education and ability—traits that lie at the heart of our hypothesis (Galor et al., 2025).

We hypothesize that, since market institutions reward individuals based on their cognitive and non-cognitive skills, as well as phenotypic and behavioral traits,[3] income inequality will be more pronounced among the descendants of populations with shorter weighted migratory distances from the cradle of humanity in Africa, as these populations exhibit greater dispersion in productive traits.

[3] Heckman et al. (2006), Case and Paxson (2008), Sunde et al. (2022).

In view of the highly plausible impact of population diversity on the emergence of institutional and cultural characteristics designed to exploit, mitigate, or foster the effects of human diversity,[4] conducting a conclusive empirical examination of the proposed hypothesis poses a significant challenge in a cross-country setting. In particular, the imposition of extractive, inequality-enhancing institutions during the colonial era in some of the most homogeneous countries (e.g., Bolivia, Ecuador, Guatemala, Paraguay, and Peru), and their enduring effects on inequality, obscures the underlying cross-country association between diversity in productive traits and income inequality.

Discerning the impact of diversity in productive traits on inequality can be fruitfully investigated through an exploration of the origins of variation in inequality among groups of individuals who, despite living in the same country, trace their origins to different ancestral homelands. These groups would be subject to the same economic incentives and political institutions, yet exhibit varying levels of ancestral population diversity. In this single-country context, the proposed hypothesis suggests that greater income inequality would arise among descendants of ancestral populations that resided closer to humanity's cradle in Africa and, consequently, exhibited higher diversity. In particular, the United States appears especially well-suited for examining this hypothesis. As a market economy, earnings differentials in the U.S. are likely to reflect variations in productive traits. Moreover, Americans trace their ancestry to nearly one hundred national origins, reflecting substantial variation in ancestral backgrounds. These virtues are further enhanced by the availability of individual-level data on earned income and ancestry for millions of U.S. inhabitants.

The empirical investigation of the proposed hypothesis examines the relationship between income inequality among individuals in the U.S. who trace their origins to the same ancestral homeland and the predicted variation in their productive traits, derived from their ancestry-adjusted migratory distance from East Africa to that homeland (Figure 3).

Figure 3: Chain of Causality

The empirical analysis establishes that groups of individuals who trace their origins to the same homelands, and whose ancestors resided closer to the cradle of humanity in Africa,

[4] Ashraf and Galor (2013b), Galor and Klemp (2023), Cook and Fletcher (2018), Arbatlı et al. (2020), and Posch et al. (2025).

indeed exhibit higher levels of inequality. This finding is drawn from various measures, including the Gini index of earned income, the share of income held by the top 1%, 5%, and 10% of earners in 2010, and repeated cross-sectional data on income distribution over the period 1980–2020. The association between diversity and inequality remains qualitatively similar within demographic bins that account for broad categories of educational attainment. The relationship between ancestral population diversity and inequality is considerable: in the baseline specification, moving from the lowest to the highest level of diversity among ancestral populations in the sample is associated with a 6.9 percentage-point higher Gini index (i.e., a 16% increase relative to the mean). This difference corresponds to a shift in the Gini index from the median to the 88th percentile of the inequality distribution. Remarkably, ancestral population diversity explains 12.5% of the overall income inequality in the US.

The findings remain qualitatively unaffected when accounting for: (a) confounding geographic characteristics that may correlate with migration distances from Africa and population diversity; (b) the possible influence of ancestral ethnolinguistic fragmentation, along with culture and institutions, on inequality among descendants in the US; and (c) US state and *Public Use Microdata Area* fixed effects, mitigating the potential impact of disproportionate ancestral migration to specific US regions. Moreover, the findings hold under alternative samples: (a) the CPS sample of second-generation migrants, avoiding self-reported ancestry and the potential differential impact of various migration waves; (b) alternative configurations of employment and age classifications; and (c) the exclusion of individuals with ancestral origins in Africa, Latin America, or the New World. Finally, the results remain robust to weighting of demographic bins based on the size of their ancestral populations.

Consistent with our hypothesis, the association between migratory distance from Africa and income inequality appears to be mediated by differences in the dispersion of key productive traits—educational attainment, labor supply, and residual ability. Individuals in the U.S. whose ancestors originated farther from the cradle of humanity in Africa tend to exhibit lower diversity in these traits. Additionally, consistent with the premise that greater interpersonal diversity increases the probability of a denser upper tail in the distribution of entrepreneurial skills, the analysis suggests that US residents whose ancestors lived closer to the cradle of humanity in Africa tend to be both more entrepreneurial and more unequal, indicating that trait dispersion may also mediate inequality through its influence on the prevalence of entrepreneurial talent.

These findings are consistent with the hypothesis that, in a market economy—where productive traits are differentially rewarded—greater diversity in such traits is associated with higher levels of inequality. Specifically, the association between diversity in productive traits and inequality appears stronger when the returns to these traits are elevated. Notably, the relationship between ancestral diversity and inequality is substantially more pronounced in the period from 1990 to 2020 than in the 1980s, a pattern that aligns with the sharp rise in the return to ability between 1980 and 1990 and the concurrent decline in the influence of labor unions.

Although the analysis focuses on individuals born in the U.S., the distribution of their productive traits may not fully reflect the original distribution within their ancestral homelands, due to selective migration patterns among their ancestors. Yet, while selection may have influenced the characteristics of individuals who chose to migrate—and, consequently, the traits brought by migrants to the U.S.—it is implausible that it resulted in group-level selection that systematically affected the within-group variance of these traits in the immigrant population. Furthermore, for selection to serve as an alternative explanation for the findings, it would need to produce a highly specific and improbable form of group selection—one that systematically reduced the variance of cultural traits among descendants of populations originating farther from the cradle of humanity in Africa, despite the absence of systematic variation in the diversity of productive traits across those ancestral homelands.

Examining the roots of inequality among descendants of Native American ancestral tribes—groups unaffected by selective migration over the past 10,000 years—further dispels concerns about the role of selective migration. Remarkably, the association between ancestral diversity and inequality remains qualitatively unchanged, despite the absence of apparent differential selective migration across Native American tribes.

2 The Out-of-Africa Migration and Population Diversity

The central mechanism underlying the proposed hypothesis is anchored in a prehistorical process that is substantiated by empirical findings across disciplines and reflects the gradual decline in human diversity as populations settled at greater distances from the African cradle of humanity. The prehistoric migration of Homo sapiens out of Africa was largely characterized by a stepwise expansion, whereby at each step, a subgroup of individuals left their ancestral settlement to establish a new colony farther away, carrying only a subset of the diversity of traits that had existed in their ancestral settlement. Due to the serial nature of this human dispersal, the resulting *Serial Founder Effect* led to a progressive reduction in the diversity of human populations that settled at greater migratory distances from Africa, amplified by population bottlenecks and genetic drift (Figure 4).

As humans migrated farther from Africa, cultural, linguistic, behavioral, genetic, and phenotypic diversity declined in the societies formed by their descendants, leading to a systematic compression in traits that underpin economic productivity (Figure 5).[5] Importantly, this enduring pattern encompasses traits that are central to economic performance, as the dispersion of predispositions toward education and cognitive ability also diminishes with migratory distance from Africa (Galor et al., 2025). This evidence, therefore, lends direct support to the core mechanism underlying our hypothesis, laying the foundation for understanding how the Out-of-Africa migration contributed to the cross-societal variation in the intensity of contemporary inequality.

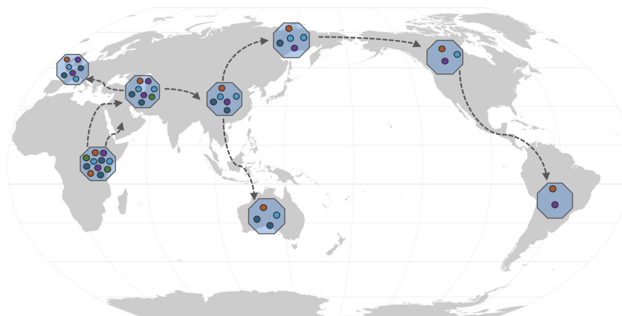


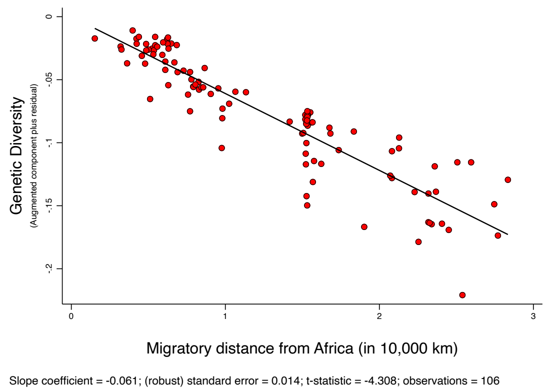
Figure 4: The Serial Founder Effect

Notes: An illustration of the declining diversity along the migratory routes Out-of-Africa.

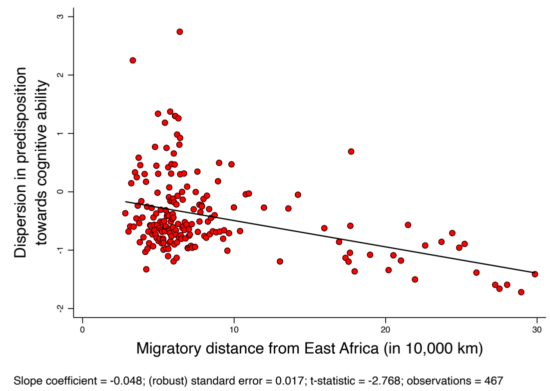
Figure 5 depicts this remarkably consistent pattern across multiple dimensions. Panel (a) documents a clear negative association between migratory distance and genetic diversity across indigenous ethnic groups. Panel (b) shows that diversity in genetic predispositions toward cognitive ability, estimated using ancient DNA in conjunction with modern genome-wide association methods, also declines with distance from Africa. Panels (c) and (d) demonstrate that folkloric and musical traditions, two deep-rooted forms of cultural expression, similarly exhibit reduced variation farther from the origin. Finally, panels (e) and (f) reveal that contemporary dispersion in cultural values, both within national populations and among second-generation migrants in Europe, is inversely related to the migratory distance of their ancestral origins. This pattern is qualitatively unaffected by the inclusion of continental fixed effects as well as controls for geographic confounders, and the removal of Africa from the sample.[6]

[5]The impact of the out-of-Africa migration on phenotypic, phonemic, and behavioral diversity among human populations has been extensively explored. See Manica et al. (2007), von Cramon-Taubadel and Lycett (2008), Betti et al. (2009), Atkinson (2011), Betti and Manica (2018), Ashraf et al. (2021).

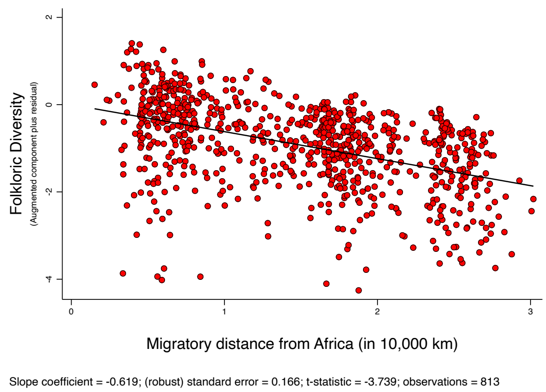
[6]Although there is some debate about the location of the cradle of humanity within Africa (e.g., Ragsdale et al. (2023)), the fact that global human dispersal began in the eastern part of the continent implies that the choice of origin has only a minor effect on predicted population diversity outside Africa. Moreover, while patterns within Africa are sensitive to this location, the estimated relationship between migratory distance and population diversity becomes even stronger when Africa is excluded from the sample.



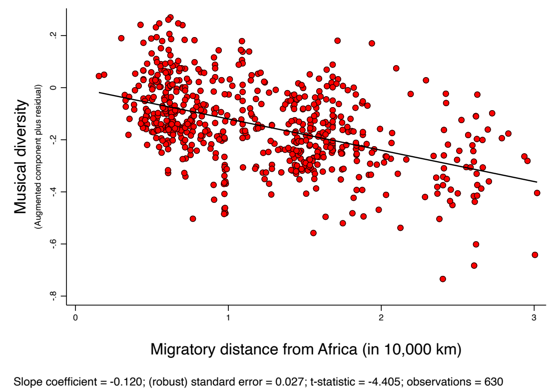
(a) Genetic diversity



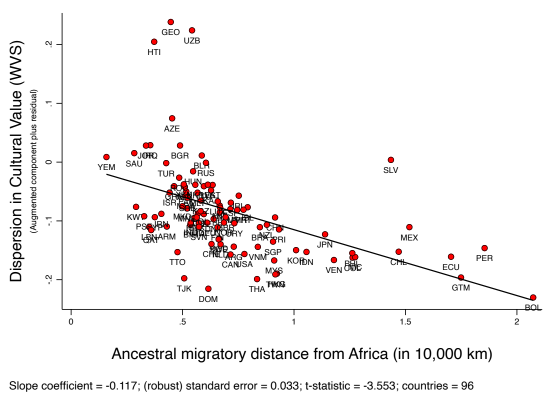
(b) Cognitive diversity



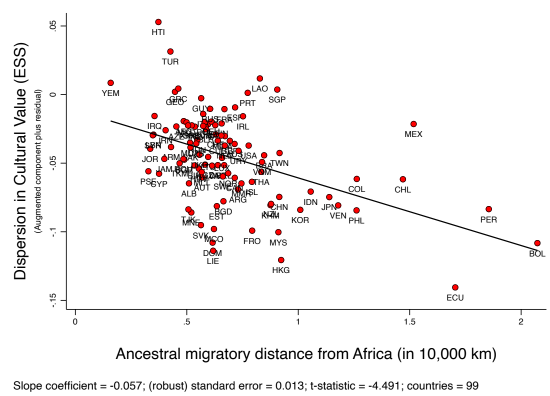
(c) Folkloric diversity



(d) Musical diversity



(e) National populations



(f) Second-Generation Migrants in Europe

Figure 5: Declining Diversity along the Migratory Routes out of Africa

Notes: This figure presents the reduction in population diversity among indigenous populations at greater migratory distances from Africa, as established in Galor et al. (2024) and Galor et al. (2025). The six scatterplots depicts the association between the (*ancestry-adjusted*) prehistoric migratory distance from the cradle of humanity in Africa and: (a) genetic diversity across indigenous ethnic groups; (b) diversity in genetic predisposition towards cognitive ability, leveraging recent advances in ancient DNA research and a state-of-the-art genome-wide association study, across artificial grid cells comprising ancient populations, (c) folkloric diversity across indigenous ethnic groups surveyed in Berezkin's Folklore and Mythology Catalogue; (d) musical diversity across indigenous ethnic groups surveyed in Lomax's Cantometrics; (e) contemporary dispersion in cultural values within countries surveyed in the WVS; (f) contemporary dispersion in cultural values among second-generation migrants in Europe surveyed in ESS.

3 Data and Empirical Strategy

The proposed hypothesis suggests that greater income inequality would be observed among ethnic groups within U.S. society whose ancestral homelands are situated at shorter weighted migratory distances from the cradle of humanity in Africa, and are therefore predicted to exhibit greater diversity in productive traits.

The empirical analysis tests this hypothesis by leveraging variations in income inequality across groups of U.S.-born individuals with distinct ancestral origins. It examines the effect of migratory distance from Africa, and thus predicted population diversity, on income inequality, drawing on data on income, ancestry, and demographic characteristics from the *American Community Survey* (ACS) 2010 and 2020 (5-year samples), as well as the U.S. Censuses of 1980, 1990, and 2000. These datasets comprise millions of individuals from over one hundred ancestral backgrounds.

3.1 Construction of Demographic Bins

The adult U.S.-born population is segmented into demographic bins defined by shared ancestral origin, sex, and age group.[7] This structure facilitates the examination of the impact of the Out-of-Africa migration, and the associated level of interpersonal diversity within societies, on inequality, while accounting for sex and age group.[8] Yet, the qualitative results remain unchanged when demographic bins are based solely on ancestral origin.

[7] To minimize measurement error in predicted diversity for individuals with multiple ancestries, we restrict our baseline analysis to the 55% of individuals in the ACS sample who report only a single ancestry. Expanding the analysis to include the full sample using either (i) migratory distance based on primary ancestry alone (Table C.7) or (ii) the average migratory distance of both ancestries (Table C.8) does not alter the qualitative findings, although the estimates are attenuated, likely due to increased measurement error. Moreover, self-reported ancestry appears to have a negligible impact on the analysis. Using data from second-generation migrants in the *Current Population Survey* (CPS), where the parental country of birth is directly observed, yields qualitatively similar results, provided the sample includes a sufficient number of ancestries (Table C.6).

[8] Including demographic bins with very few individuals may lead to measurement error in the computation of their Gini index (Figure D.1). Two approaches can mitigate this concern: (i) weighting bins by their population size, or (ii) imposing a minimum threshold on bin size. Due to the highly skewed distribution of the US population across ancestries, weighting bins by population size would amplify the influence of a few ancestral homelands with very large descendant populations (i.e., Germany, Ireland, and England), causing their patterns of inequality to dominate the analysis. Although the relationship between diversity and inequality remains qualitatively robust under weighted regression (Table C.9), this method is not included in the baseline analysis due to its distortionary effects, particularly when bins are further stratified. Instead, in light of the analysis in Figure D.1, we restrict the sample to demographic bins with at least 10 individuals—the minimum required to compute the share of the top 10%. Imposing a higher minimum bin size reduces the number of ancestral homelands while yielding only marginal gains in the stability of the estimates (Figure D.1).

3.2 Income Inequality

The Gini index of earned income serves as the primary measure of income inequality within each demographic bin,[9] complemented by the shares of income held by the top 1%, top 5%, and top 10% of the income distribution.[10] Income inequality is measured within each demographic bin, using data on US-born, working-age individuals who are in the labor force.

Importantly, the qualitative findings remain robust when the sample is instead restricted to: (i) employed individuals, (ii) individuals employed in the private sector, (iii) prime working-age individuals (i.e., ages 25–54), or (iv) full-time, year-round workers (i.e., those working at least 35 hours per week for at least 50 weeks per year).

3.3 Migratory Distance from the Cradle of Humanity in Africa

As established in Section 2, migratory distance from the cradle of humanity in Africa to each indigenous society captures the legacy of the out-of-Africa migration on the diversity of productive traits. Accordingly, this migratory distance serves as the pivotal independent variable in the empirical analysis.[11]

In constructing this variable for each ancestral country of origin, and thus for each demographic bin in the US, we take into account that modern nations often consist of populations with multiple ancestral origins. A country's migratory distance is therefore computed as the weighted average of the migratory distances from Africa of its ancestral populations, based on their proportional representation in the national population.[12] Consequently, the independent variable is the predicted population diversity within each demographic bin in

[9]Figure B.1 depicts the histogram of the level of inequality, as captured by the Gini index, across demographic bins.

[10]As shown in Table C.12, adjusting for topcoded income values does not affect our findings. In our sample, topcoded individuals account for 1.7% of observations. To address this, we apply a standard Pareto-imputation procedure, implemented separately for wage income and business income, as these components are topcoded independently and at distinct thresholds in the ACS. For each state-year cell, we estimate the Pareto shape parameter α using the top 1% of non-topcoded observations within each component. Topcoded values are then replaced with the expected value from the fitted Pareto distribution above the relevant topcoding threshold. Total earned income is reconstructed as the sum of these adjusted components.

[11]Restricting the analysis to predicted genetic diversity, as constructed by Ashraf and Galor (2013b), does not affect the qualitative results (Table C.21).

[12]This adjustment relies on the migration matrix of Putterman and Weil (2010), which maps contemporary populations to their ancestral homelands as of the year 1500. Potential concerns about measurement error in this adjustment are alleviated by the fact that the results remain qualitatively robust when the analysis is restricted to demographic bins corresponding to ancestral homelands in the Old World, where modern populations are predominantly native to their territory (Table C.14).

the US, as determined by the weighted migratory distance from East Africa to the homeland of the individuals in these demographic bins.[13]

3.4 Empirical Strategy

Leveraging variation in prehistoric migratory distance from East Africa to predict ancestral population diversity implies that our empirical strategy is immune to concerns about reverse causality, as contemporary income inequality in the US could not have influenced ancestral population diversity. However, to the extent that migratory distance from Africa is correlated with other ancestral determinants of income inequality in the US, the analysis may still be subject to omitted variable bias. Moreover, given the potential for selective migration into the US, the estimates may also be biased by selection.

First, migratory distance from Africa may be correlated with potential deep-rooted ancestral geographical characteristics that shaped societal diversity and, plausibly, income inequality. To account for the influence of these potentially confounding characteristics, we control for a range of ancestral geographical variables that could have shaped population diversity. In particular, we account for the potential impact of: (i) absolute distance from the equator, in light of its well-documented adverse effect on biodiversity; (ii) ecological diversity, due to its potential contribution to population diversity (Fenske (2014), Raz (2025)); and (iii) geographical isolation, as it tends to reduce both biodiversity and human diversity.

Second, migratory distance from Africa brought about a decline in the number of ethnic groups (Galor and Klemp, 2023) and reduced the degree of ethnolinguistic fragmentation (Ashraf and Galor, 2013a). Accordingly, it is a priori plausible that the influence of migratory distance from Africa on inequality operates through ethnic fragmentation rather than through diminished diversity in productive traits (Alesina et al. (2003), Montalvo and Reynal-Querol (2005), Desmet et al. (2009), Alesina et al. (2016a)). Under this alternative channel, migratory distance from Africa would affect inequality by fostering greater social cohesion and, arguably, reducing individualism and inequality, rather than by diminishing diversity in productive traits. To evaluate this potential alternative mechanism, we account for the confounding role of ancestral ethnic fragmentation, as captured by measures of ethnic fractionalization and ethnolinguistic fractionalization in the ancestral homeland of each demographic bin in the US.[14]

[13] The algorithm underlying the computation of these migratory distances is detailed in Appendix A.3.

[14] While some aspects of interpersonal diversity can be captured by indices of ethnolinguistic fractionalization and polarization, these measures predominantly reflect the proportional representation of ancestral groups in the population, disregarding the importance of the degree of interpersonal diversity *within* each ancestral group for the overall level of diversity at the national level.

Third, contemporary inequality in the United States could stem from the intergenerational persistence of inequality originating in ancestral homelands. This possibility would affect the interpretation of our findings in two key ways. First, if migratory distance from Africa is correlated spuriously with ancestral inequality, our results may inadvertently capture this confounding factor. Second, if migratory distance from Africa shaped ancestral inequality, it is possible that its influence on contemporary inequality in the US operates through the persistence of ancestral inequality rather than through interpersonal diversity in productive traits within the US.[15] To address these concerns, we incorporate various measures of inequality in ancestral homelands to assess their potential enduring impact on inequality among descendants in the US. While such controls could, in principle, absorb the effect of ancestral diversity in productive traits, this possibility is tempered by the fact that fully developed market economies were largely absent in many ancestral homelands. Consequently, historical inequality in these regions likely reflected institutional or cultural factors—such as caste systems or elite beyond the distribution of productive traits. Thus, controlling for ancestral inequality is unlikely to eliminate the observed relationship between interpersonal diversity in productive traits and contemporary inequality.

Fourth, the degree of inequality in the ancestral homelands of the US population may reflect the institutional and cultural characteristics prevalent in these regions. Inequality-mitigating institutions in an ancestral homeland may have reduced inequality in that historical context. Yet, since descendants now living in the United States are influenced by the institutional environment of the US, the institutional setup of the ancestral homeland may have mattered primarily through its impact on historical inequality and its potential persistence in shaping contemporary inequality among descendants in the US. Hence, to account for the potential role of ancestral institutions, it would be appropriate to control for ancestral inequality and examine whether ancestral population diversity continues to predict diversity in productive traits within the contemporary US population.

Furthermore, cultural characteristics of ancestral homelands are portable and can be transmitted by migrants and their descendants. In particular, cultural traits originating in ancestral homelands may significantly influence economic outcomes and inequality among their descendants in the US (Guiso et al., 2006). For instance, *Uncertainty Avoidance* may reduce entrepreneurship and income variability, *Individualism* may contribute to income disparity, and *Long-Term Orientation* may promote investment in physical and human capital,

[15] While this perspective aligns with our broader hypothesis, it suggests a distinct mechanism underlying our empirical results. Specifically, groups whose ancestors lived closer to the cradle of humanity in Africa might exhibit higher levels of contemporary inequality in the US not because of greater contemporary diversity in productive traits, but because their ancestors who migrated to the US carried higher levels of inequality.

as well as technological adoption, potentially increasing income inequality. To address this potential threat to our identification strategy, we account for the effects of these ancestral cultural factors on inequality across demographic bins in the US.

Fifth, the observed relationship between ancestral migratory distance and inequality among descendants of different ethnic groups in the US might be shaped in part by differential treatment, including systematic discrimination—particularly against individuals of African or Latin American descent. However, for discrimination to confound our estimates, it must be correlated both with inequality and with migratory distance in a way that mimics the effect of diversity in productive traits. That is, bias would arise only if individuals whose ancestral homelands are farther from Africa experience systematically different levels of discrimination and such discrimination directly reduces income inequality within those groups—despite no underlying differences in the dispersion of productive traits. This scenario seems unlikely. Nonetheless, we probe this possibility through three complementary robustness strategies: (i) excluding demographic bins associated with homelands in Africa or Latin America; (ii) including continental fixed effects to compare variation across ethnic homelands within continents rather than across them; and (iii) controlling for US state and micro-area fixed effects to account for regional variation in discriminatory environments, such as those prevalent in the American South.[16]

Sixth, disproportionate migration of ancestral groups to specific US regions (e.g., Scandinavians and Germans to the Midwest, Mexicans to the Southwest, and Asians to the West) could have influenced inequality within these groups due to differences in regional market returns rather than disparities in productive traits. However, such migratory patterns would only undermine our findings if, implausibly, descendants of ancestral homelands farther from humanity's cradle in Africa were disproportionately concentrated in regions with lower inequality or adopted local institutions that systematically reduced inequality, despite no underlying differences in productive traits. To further enhance the robustness of our identification strategy against this potential challenge, we incorporate fixed effects for US states and *Public Use Microdata Areas*, thereby accounting for regional characteristics that may have influenced inequality.

Finally, while the analysis exclusively examines individuals born in the US—who themselves did not migrate—their distribution of productive traits might not accurately reflect that of their ancestral homelands, due to selective migration patterns among their forebears. Although selection may have influenced which individuals migrated and, in turn, the traits introduced into the US population, it is highly unlikely to have resulted in group-level selection that systematically altered the variance of traits within immigrant groups. For

[16] We control for all 50 states and over 2,000 micro-areas.

selection to explain our findings, it would require an exceedingly implausible form of group selection—one that consistently reduced trait variance among descendants from ancestral homelands farther from the cradle of humanity in Africa, despite no corresponding differences in the diversity of productive traits across those regions. Nonetheless, to further assess the role of selection, we examine inequality among descendants of ancestral Native American tribes—populations that have not experienced selective migration over at least the past 10,000 years.

3.5 The Empirical Model

In line with our hypothesis, we model inequality within each demographic bin as a function of the population diversity associated with that bin, proxied by the weighted prehistoric migratory distance from Africa of its ancestral population. The model includes fixed effects for sex and age group in the US population, and controls for potentially confounding characteristics of the ancestral homeland—capturing geographical factors (proximity to the equator, ecological diversity, and degree of isolation) and ancestral attributes (ethnic fragmentation, historical inequality, and cultural and institutional characteristics).

In particular, we estimate the following OLS model:

$$G_{s,a,h} = \alpha + \beta D_h + \delta_s + \zeta_a + \theta X_h + \epsilon_h,$$

where the dependent variable, $G_{s,a,h}$, denotes the level of inequality in a demographic bin composed of individuals of the same sex s, age group a, and ancestral homeland h. The key independent variable, D_h, represents the ancestry-adjusted migratory distance from Africa to homeland h. The model includes fixed effects for sex, δ_s, and age group, ζ_a, and controls for a vector of characteristics of the ancestral homeland, X_h, comprising geographical factors, ethnic fragmentation, historical inequality, cultural norms, and institutional features. The coefficient of interest, β, is hypothesized to be negative.

Since the key independent variable is defined at the level of individuals' ancestral homelands across demographic bins, standard errors are clustered at the ancestral homeland level to account for within-group correlation.

4 Main Findings

4.1 Ancestral Population Diversity & Income Inequality

This section examines the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) multiple measures of income inequality among their descendants in the US—the Gini index and the income shares of the top 1%, top 5%, and top 10%. The analysis is based on income, ancestry, and demographic data from the 2010 ACS 5-year sample.[17]

Table I reports the baseline analysis, as depicted in Figure 6. The estimated coefficient in Column 1 suggests a highly statistically significant negative association between prehistoric migratory distance from Africa and our baseline measure of income inequality — the Gini index.[18] The estimates in Columns (2)-(4) reveal a similarly strong and statistically significant negative association under alternative measures of income inequality, namely the shares of income held by the top 1%, top 5%, and top 10% of the income distribution.[19]

[17]Due to potential distortions related to COVID-19, the 2020 ACS 5 year sample is excluded from the baseline analysis. It is, however, incorporated into the extended analysis of ancestral diversity and inequality using repeated cross sections from 1980 to 2020.

[18]Moreover, this baseline association remains highly statistically significant when we account for continent-of-origin fixed effects, thereby identifying the relationship between prehistoric migratory distance from Africa and inequality based on variation across ancestral homelands within continents (Column (2) in Table C.14).

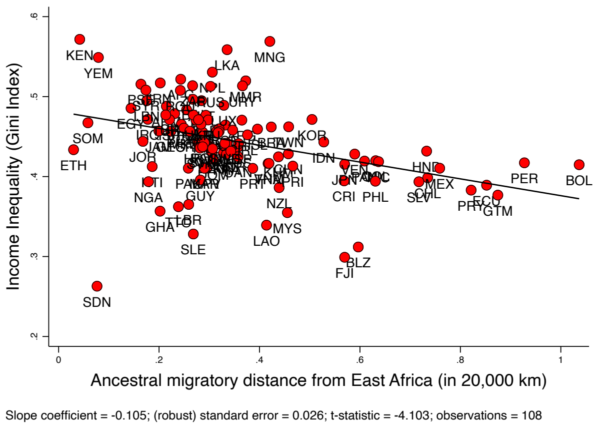
[19]Since computing the share of income held by the top 1% requires at least 100 individuals within a demographic bin, those bins with fewer than 100 individuals are dropped by construction, reducing the number of ancestral homelands by nearly one third.

Table I: Ancestral Diversity & Income Inequality

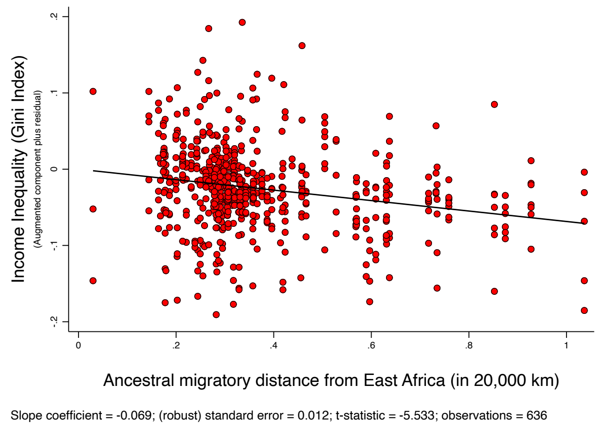
	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.069***	−0.023***	−0.053***	−0.056***
from the cradle of humanity	(0.012)	(0.0059)	(0.012)	(0.013)
Dep. var. mean	0.42	0.066	0.20	0.30
Individuals	1781548	1770406	1780511	1781548
Demographic bins	636	354	559	636
Ancestral homelands	95	61	87	95
Adjusted R^2	0.22	0.20	0.24	0.21

Notes: This table reports the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) various measures of inequality among their descendants in the US, based on the 2010 ACS 5 year sample. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

The association between an increase in prehistoric migratory distance from Africa and levels of inequality is sizable. In particular, a shift in the geographic origin of an ancestral population from the lowest to the highest ancestry-adjusted migratory distance from Africa (i.e., a 20,000 km increase) is associated with a 6.9 percentage point lower Gini index (i.e., a 16% reduction relative to the mean level of 0.42). This magnitude corresponds to a decrease in the Gini index from the median to the 12th percentile of the inequality distribution. Strikingly, diversity in productive traits—as implied by migratory distance from Africa—is associated with 13% of the variation in inequality within groups (Column 1, Table C.2).



(a) Raw Ancestry Bin



(b) Baseline Demographic Bins

Figure 6: Income Inequality & Ancestral Migratory Distance from East Africa

Notes: This figure depicts the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US. Panel (a) depicts the scatterplot of the association between income inequality and migratory distance from East Africa, irrespective of the inclusion of sex and age fixed-effects (Table C.1, Column (1)). Panel (b) depicts a (binned) scatterplot of the association between income inequality and migratory distance from East Africa in the baseline specification (Table I, Column 1).

These baseline findings remain qualitatively robust across a range of alternative specifications, data sets, ancestry measures, and sample restrictions, including:

- **Demographic groupings**: Analyzing individuals in the US grouped solely by ancestral origin, excluding sex and age (Table C.1), or using alternative permutations of demographic bins (Table C.2).
- **Employment classifications**: Considering variations such as (i) employed individuals, (ii) private sector workers, (iii) prime working-age individuals (25–54 years), or (iv) full-time, year-round workers defined as those working at least 35 hours per week for 50 weeks annually. (Table C.3).
- **Age and sex restrictions**: Employing alternative age classifications (Table C.4) or restricting samples to either males or females (Table C.5).
- **Ancestry measures**: Using data on the parental country of birth for second-generation migrants from the *Current Population Survey* (CPS) (Table C.6); including individuals who report a secondary ancestry, assigning the migratory distance of their primary ancestry (Table C.7); or averaging migratory distances across reported ancestries (Table C.8).

- **Weighted Regression**: Incorporating weighted regression based on population size (Table C.9).
- **Income types**: Considering earned income from either wage or self-employment (Table C.10).
- **Accounting for mean Income**: Controlling for the log mean income within each demographic bin (Table C.11).
- **Time of arrival to the US**: Focusing on descendants of populations arriving before 1850 (Table C.13) or on recent second-generation migrants (Table C.6).
- **Exclusions of world regions**: Excluding individuals with ancestral homelands in Africa, Latin America, or the New World (Table C.14).
- **Regional fixed effects**: Incorporating fixed effects for US states and for over 2,000 micro-areas (Table C.16–Table C.17).
- **Spatial dependence**: Accounting for spatial dependence using the Conley (1999) method (Table C.15).

4.2 Ancestral Diversity and Inequality across Time

In the absence of changes in the rewards to productive traits (e.g., changes in labor market institutions such as the decline in unionization rates), or in the composition of migrants, the quantitative impact of the prehistoric migratory distance from Africa, and thus predicted ancestral diversity, on inequality would be expected to remain stable over time.

Table II shows that the patterns established in Table I, based on data from the ACS 2010 5-year sample, are qualitatively unchagned in repeated cross section over the period 1980-2020, using the Censuses of 1980, 1990, and 2000, and the ACS 2010 and 2020 5-year samples. In particular, a shift in the geographic origin of an ancestral population from the lowest to the highest ancestry-adjusted migratory distance from Africa (i.e., a 20,000 km increase) is associated with a 6.5 percentage point lower Gini index (i.e., a 16% reduction relative to the mean level of 0.40). [20] The findings reinforce the hypothesis that in a market

[20]Extending the repeated cross-section analysis back to 1940—the first census year that includes income data—does not alter the qualitative results. However, this extension introduces several limitations. First, the 1940–1970 surveys lack information on ancestry, necessitating reliance on second-generation migrants and creating inconsistencies with the 1980–2020 sample. Second, the American labor market was less competitive during this earlier period (e.g., characterized by higher unionization rates and lower returns to ability). Third, the 1940 Census reports only wage income and omits business income, introducing an additional source of inconsistency across the sample period.

economy, where productive traits are differentially rewarded, diversity in these traits leads to greater inequality.

Table II: Ancestral Diversity & Inequality: Repeated Cross-Section, 1980-2020

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.065***	−0.019***	−0.051***	−0.057***
from the cradle of humanity	(0.012)	(0.0053)	(0.0073)	(0.0087)
Sample FE	✓	✓	✓	✓
Dep. var. mean	0.40	0.060	0.18	0.29
Individuals	8231170	8187412	8226177	8231170
Demographic bins	2852	1638	2493	2852
Ancestral homelands	108	66	97	108
Adjusted R^2	0.23	0.36	0.32	0.27

Notes: This table reports the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) various measures of inequality in the US over the period 1980–2020. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex, age-group, and sample period fixed-effects. The unit of observation is a demographic bin in a given survey period. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

Specifically, the estimated effect of diversity in productive traits on inequality becomes more pronounced when the returns to these traits are higher. Notably, given the five-fold increase in the return to ability in the US from 1980 to 1990, and the relative stability of these returns thereafter (Figure D.5), the impact of ancestral diversity on inequality is substantially larger and more significant over the period 1990–2020 relative to 1980. This amplification of the effect of diverse productive traits on economic inequality is further reinforced by the sharp decline in the influence of labor unions during this period.

4.3 Ancestral Diversity & Inequality within Education Categories

Interpersonal diversity may induce individuals within each demographic bin to sort into different educational categories. A priori, some of the impact of diversity on inequality may therefore reflect this sorting, potentially attenuating the impact of diversity on inequality.

Nevertheless, as established in Table III, the qualitative impact of interpersonal diversity on inequality still holds if demographic bins are further refined, accounting for four aggregate educational categories generated based on the IPUMS classification (i.e., high school or below, some college, college, and more than college).

Table III: Ancestral Diversity and Inequality within Educational Categories

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.081***	−0.014***	−0.046***	−0.056***
from the cradle of humanity	(0.0096)	(0.0047)	(0.0086)	(0.0074)
Education FE	✓	✓	✓	✓
Dep. var. mean	0.38	0.057	0.17	0.27
Individuals	1778218	1746052	1773501	1778218
Demographic bins	1841	948	1506	1841
Ancestral homelands	88	45	78	88
Adjusted R^2	0.14	0.078	0.16	0.16

Notes: This table reports the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) various measures of inequality in the US (based on the ACS 2010 5 year sample), accounting for individuals' educational attainment. The unit of observation is a bin defined by ancestry, sex, age group, and educational attainment. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex, age-group, and educational categories fixed-effects. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

In particular, as reported in Column 1, the estimated effect of migratory distance from Africa on the Gini index remains highly significant and stable even within educational categories. In particular, a shift in the geographic origin of an ancestral population from the lowest to the highest ancestry-adjusted migratory distance from Africa (i.e., a 20,000 km increase) is associated with a 8.1 percentage point lower Gini index (i.e., a 21% reduction relative to the mean level of 0.38).

4.4 The Importance of Within vs. Between Group Inequality

The testing of our hypothesis is inherently associated with the exploration of the impact of ancestral population diversity on *within-group* inequality among the descendants of these ancestral populations in the US. Importantly, this dimension of income inequality is the dominant component of overall income inequality. Inequality within groups of individuals originating from the same ancestral homelands is an order of magnitude greater than inequality between these groups. In particular, in the ACS 2010 5-year sample, within-group inequality accounts for 96% of the variation in overall income inequality in the US, whereas between-group inequality accounts for only 4% (Figure 7).

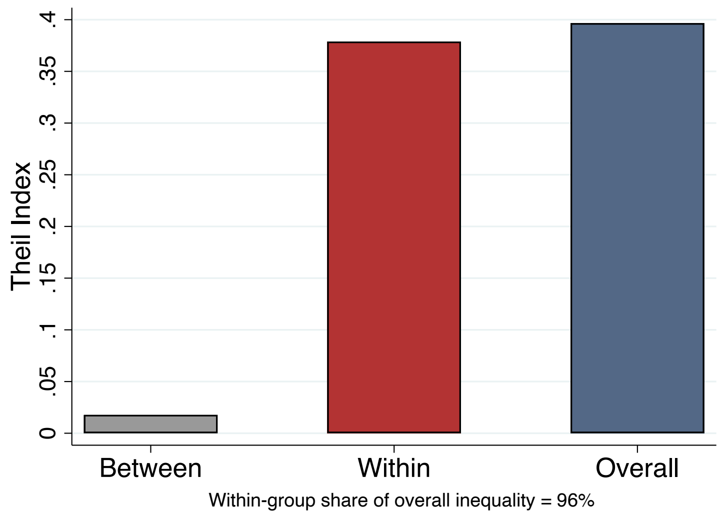


Figure 7: Decomposition of US Income Inequality

Notes: This figure depicts the Theil decomposition of income inequality (based on the ACS 2010 5-year sample) into within-ancestry and between-ancestry components.

The decomposition of the overall level of income inequality into the within-group and the between-group components suggests therefore that inequality *within* groups of individuals descended, for instance, from Europe or Africa is an order of magnitude larger than the inequality *between* the descendants of European and African populations; inequality that has plausibly been shaped by the persistent effects of slavery and the discrimination faced by the African American population.[21]

Remarkably, since diversity in productive traits alone accounts for 13% of the variation in within-group inequality (Column 1, Table C.2), and within-group inequality constitutes 96% of the overall inequality, diversity in productive traits explains an estimated 12.5% of the overall income inequality in the US.

[21] To fully capture the potential impact of slavery and the historical discrimination of African Americans on between-group inequality, this decomposition includes all individuals who report an ancestry, including reported ancestry that cannot be mapped into an ancestral national homeland (e.g., African Americans, Europeans, Latin Americans, White).

5 Ancestral Diversity & Inequality: Confounders

5.1 Geographical Characteristics in the Ancestral Homeland

Migratory distance out of Africa could be correlated with deep-rooted exogenous geographical determinants of societal interpersonal diversity, and the estimated impact of diversity on inequality may partly capture the influence of these underlying geographical factors.

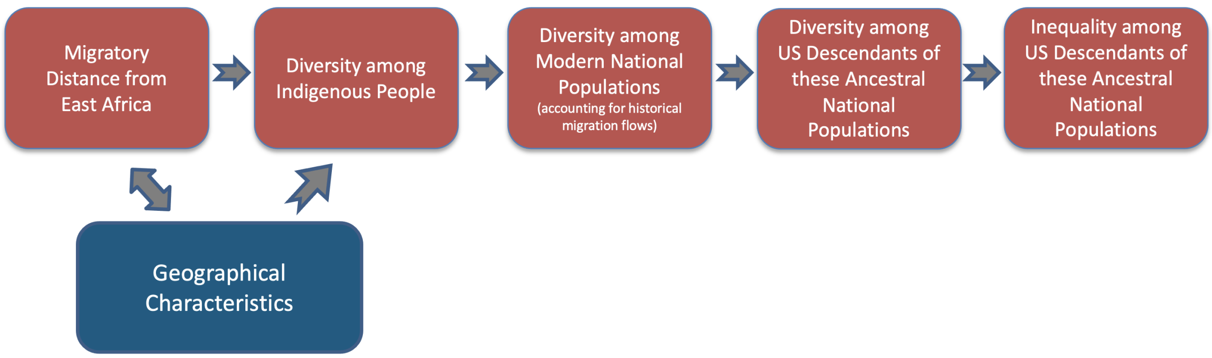


Figure 8: Geographical Factors as Omitted Variables in the Baseline Model

Hence, to mitigate this concern, we account for a range of potentially confounding ancestral geographic characteristics that may have also shaped diversity: (a) absolute distance from the equator and its well-documented adverse effect on biodiversity; (b) ecological diversity and its influence on population diversity; and (c) geographical isolation and its impact on biodiversity loss.[22]

[22]These ancestral geographic characteristics are ancestry-adjusted, reflecting the ancestral composition of the population in each ancestral homeland, and thus the geographical heritage of each segment of the population.

Table IV: Ancestral Diversity and Income Inequality: Accounting for Geographical Determinants of Diversity in the Ancestral Homeland

	Gini						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestral migratory distance	−0.069***		−0.070***		−0.071***		−0.068***
from the cradle of humanity	(0.012)		(0.014)		(0.013)		(0.012)
Ancestral absolute latitude		0.0039	−0.00090				
		(0.0035)	(0.0037)				
Ancestral caloric suitability (s.d.)				−0.00064	0.0034		
				(0.0033)	(0.0029)		
Ancestral caloric suitability (mean)				−0.0065	−0.0021		
				(0.0040)	(0.0036)		
Ancestral island						−0.014	−0.011
						(0.013)	(0.011)
Dep. var. mean	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Individuals	1781548	1781548	1781548	1781548	1781548	1781548	1781548
Demographic bins	636	636	636	636	636	636	636
Ancestral homelands	95	95	95	95	95	95	95
Adjusted R^2	0.22	0.18	0.22	0.18	0.22	0.18	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US remains robust when accounting for geographical determinants of diversity in the ancestral homeland. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. The coefficients for all controls are standardized. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

Reassuringly, Table IV suggests that the baseline results are qualitatively unaffected by the inclusion of these potential deep-rooted geographic determinants of societal interpersonal diversity. Columns 2–3 account for the potentially confounding effect of absolute latitude; Columns 4–5 account for ecological diversity, as captured by the mean and standard deviation of the Caloric Suitability Index;[23] and Columns 6–7 consider the potential influence of ancestral homeland isolation on the compression of traits, captured by a dummy variable indicating whether a group's ancestral origin is located on an island.

[23] This index reflects the potential calories per hectare per year of the most productive crop (Galor and Özak, 2016).

5.2 Ancestral Ethnic Fragmentation

Migratory distance from Africa has been shown to be correlated with a decline in the number of ethnic groups (Galor and Klemp 2023) and a decrease in the degree of ethnolinguistic fragmentation (Ashraf and Galor, 2013a). Hence, it is a priori plausible that the baseline results may, in part, reflect the impact of migratory distance on inequality through ethnic fragmentation rather than through diversity in productive traits. In this case, the effect of migratory distance on inequality could operate via greater social cohesion rather than via a reduction in the diversity of productive traits.

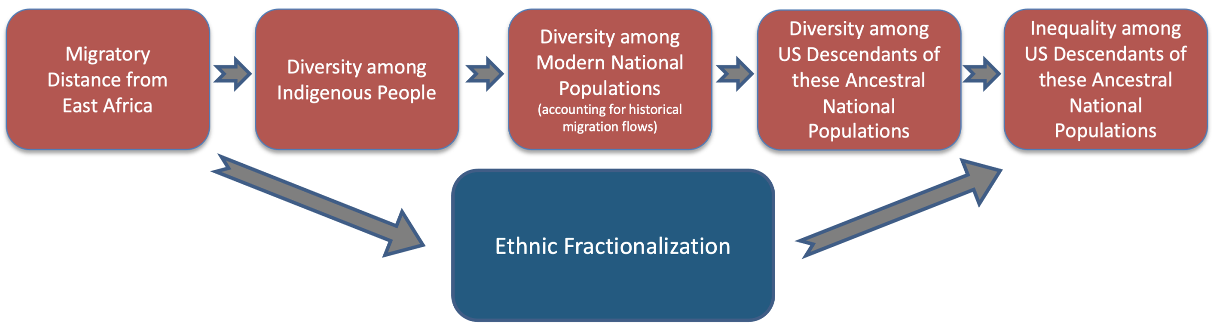


Figure 9: Ethnic Fractionalization as Omitted Variable in the Baseline Model

To explore this potential alternative channel, we account for the confounding effect of ancestral ethnic fragmentation, as captured by measures of ethnic fractionalization (Alesina et al. 2003) and ethnolinguistic fractionalization (Desmet et al., 2009) in the ancestral homeland of each demographic bin in the US.[24]

Table V suggests that these measures of ancestral ethnic fragmentation are either associated with lower inequality in the US (Column 2) or are orthogonal to the level of inequality (Column 5). Moreover, as reported in Columns 3 and 6, in regressions that include both interpersonal diversity—as captured by migratory distance from Africa—and the different measures of ancestral ethnic fragmentation, the coefficient on interpersonal diversity and its statistical significance remain largely unchanged. The evidence therefore suggests that the baseline results are unlikely to be driven by the impact of migration from Africa on ethnic fragmentation in the ancestral homelands.

[24]While some aspects of interpersonal diversity can be captured by indices of ethnolinguistic fractionalization, these measures predominantly reflect the proportional representation of ancestral groups in the population, abstracting by construction from the importance of the degree of interpersonal diversity within each ancestral group for the overall level of diversity at the national level. These measures of population diversity may therefore obscure the true impact of population diversity on inequality within nations.

Table V: Ancestral Diversity and Income Inequality: Accounting for the Impact of Ancestral Ethnic Fragmentation

	Gini					
	(1)	(2)	(3)	(4)	(5)	(6)
Ancestral migratory distance	−0.068***		−0.064***	−0.069***		−0.065***
from the cradle of humanity	(0.013)		(0.013)	(0.012)		(0.013)
Ancestral ethnic fractionalization		−0.0068**	−0.0047			
		(0.0034)	(0.0033)			
Ancestral ethnolinguistic fractionalization					−0.0071	−0.0052
					(0.0048)	(0.0042)
Dep. var. mean	0.42	0.42	0.42	0.42	0.42	0.42
Individuals	1755180	1755180	1755180	1737970	1737970	1737970
Demographic bins	607	607	607	620	620	620
Ancestral homelands	91	91	91	93	93	93
Adjusted R^2	0.22	0.19	0.23	0.22	0.18	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US remains robust when accounting for ancestral ethnic fragmentation. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. Despite the net gain of 2 ancestral homelands in Columns 4-6, the number of individuals declines since a sizable ancestral homeland is lost relative to Columns 1-3. The unit of observation is a demographic bin. The coefficients for all controls are standardized. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

5.3 Ancestral Inequality

Contemporary inequality in the United States could partly reflect the persistence of inequality inherited from ancestral homelands. This scenario would carry two important implications for the interpretation of our findings. First, if migratory distance from Africa is correlated with ancestral inequality for reasons unrelated to diversity in productive traits, our estimates may inadvertently capture this historical confounder. Second, to the extent that migratory distance shaped inequality in ancestral societies, its influence on current inequality in the US may operate through the intergenerational transmission of inequality rather than through the persistence of diversity in productive traits within the US.[25]

[25] While this interpretation remains consistent with our broader theoretical framework, it highlights a distinct channel. In particular, higher contemporary inequality among groups whose ancestors lived nearer to the cradle of humanity in Africa may reflect the inheritance of higher ancestral inequality rather than greater diversity in productive traits today.

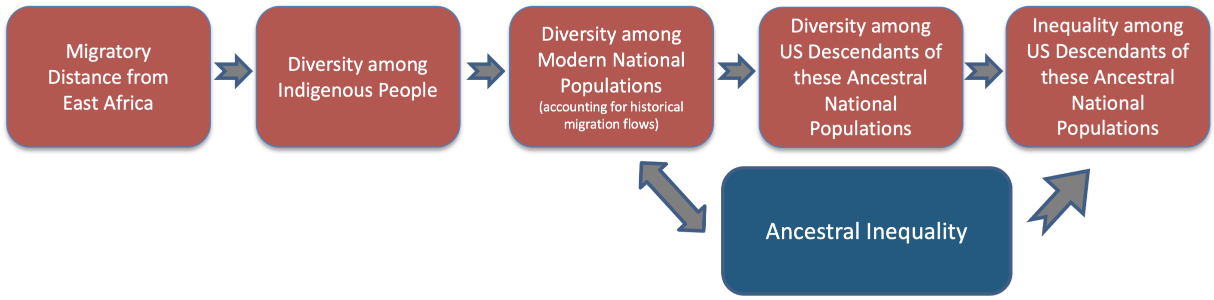


Figure 10: Ancestral Inequality as Omitted Variable in the Baseline Model

To address these concerns, we include several measures of historical inequality in ancestral homelands to assess their potential lasting influence on inequality among the descendants of these homelands in the US. While these controls could, in principle, absorb the effect of ancestral diversity in productive traits, this concern is mitigated by the fact that many ancestral societies lacked fully developed market economies. In such contexts, historical inequality likely reflected institutional or cultural arrangements beyond the distribution of productive traits. Accordingly, controlling for ancestral inequality is unlikely to eliminate the observed association between interpersonal diversity in productive traits and contemporary inequality.

Table VI presents evidence that the association between ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and income inequality among their descendants in the US remains robust when accounting for the degree of inequality in the ancestral homeland. Ancestral inequality is proxied by (i) ancestral class stratification (Columns 1–3), (ii) ancestral slavery (Columns 4–6), and (iii) the share of income held by the top 10% in the ancestral homeland in 1900 (Columns 7–9).[26] The results indicate that the estimated impact of migration from Africa on inequality remains intact when controlling for ancestral inequality, suggesting that the findings are unlikely to be driven by its intergenerational persistence (Columns 3, 6, and 9).[27]

[26] Data on historical Gini index are not widely available.

[27] The estimates in Table C.19 further suggest that a qualitatively similar pattern holds under alternative measures of ancestral inequality: (i) ancestral Gini in the period 1980–1999, as reported by the World Development Indicators, and (ii) ancestral ethnic inequality (Alesina et al., 2016b).

Table VI: Ancestral Diversity & Inequality: Accounting for Ancestral Inequality

	Gini								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ancestral migratory distance	−0.067***		−0.067***	−0.067***		−0.070***	−0.10***		−0.10***
from the cradle of humanity	(0.013)		(0.013)	(0.013)		(0.013)	(0.024)		(0.023)
Ancestral class stratification		0.0094	0.011						
		(0.024)	(0.019)						
Ancestral slavery					−0.0040	−0.0086			
					(0.0092)	(0.010)			
Ancestral share of income								0.14	0.091
held by the top 10% (1900)								(0.11)	(0.10)
Dep. var. mean	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43
Individuals	1778896	1778896	1778896	1778586	1778586	1778586	1227976	1227976	1227976
Demographic bins	621	621	621	613	613	613	154	154	154
Ancestral homelands	93	93	93	92	92	92	21	21	21
Adjusted R^2	0.22	0.18	0.22	0.22	0.18	0.22	0.29	0.21	0.29

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US remains robust when accounting for ancestral inequality. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

5.4 Ancestral Cultural and Institutional Factors

The level of inequality observed in the ancestral homelands of the US population may be indicative of enduring institutional and cultural features historically prevalent in those regions.

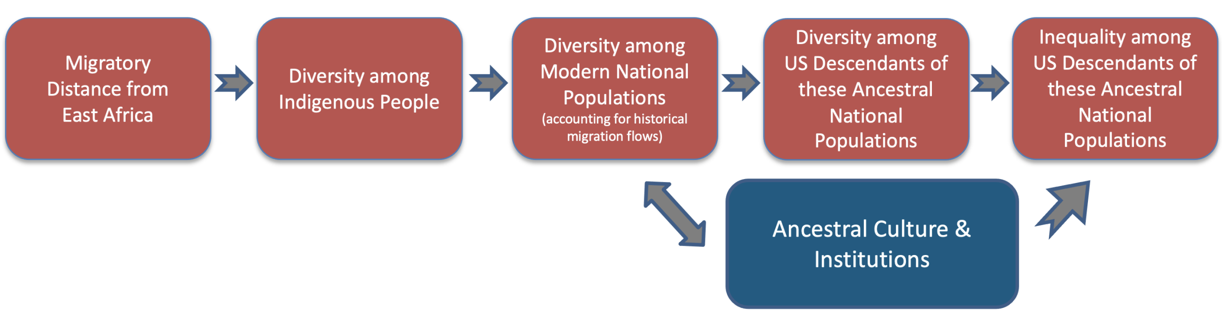


Figure 11: Ancestral Culture as Omitted Variable in the Baseline Model

The presence of inequality-mitigating institutions in ancestral homelands may have reduced inequality in the ancestral environment. Yet, the descendants of these homelands in the US are subject to the institutional characteristics of the US rather than those of their ancestral origins. The institutional structures in the ancestral homelands may still have mattered indirectly, as they could have influenced historical inequality and its potential intergenerational effects on current inequality among their descendants in the US. However, as shown in Table VI, ancestral inequality does not appear to have a persistent effect on income inequality today.

More directly, as implied by the analysis in Table C.20, ancestral institutions, such as the colonial legacy in ancestral homelands, the degree of jurisdictional hierarchy among the ethnic groups that compose national homelands, and the historical level of democracy in these nations as captured by the Polity V index in 1900, are not associated with the degree of inequality among the descendants of these populations in the US and do not affect the baseline findings.

Yet cultural characteristics in the ancestral homeland are inherently portable and could be transmitted across generations by migrants and their descendants. In particular, certain cultural traits present in some ancestral homelands may have significantly influenced inequality among their descendants in the US and could have mediated the effect of migratory distance from Africa on contemporary inequality. Specifically, (i) *Uncertainty Avoidance* may reduce entrepreneurship and narrow the distribution of earned income, (ii) *Long-Term Orientation* may promote investment in physical and human capital, as well as the adoption of new technologies, thereby increasing wage variability, and (iii) *Individualism* may contribute to greater income disparity.

Table VII suggests that *Uncertainty Avoidance*, *Long-Term Orientation*, and *Individualism* are not associated with the relationship between migratory distance from Africa and inequality. The estimated associations between each of these cultural traits and inequality are statistically indistinguishable from zero (Columns (2), (5), and (8)). Moreover, in regressions that include both migratory distance from Africa and each of these cultural factors (Columns (3), (6), and (9)), the coefficient on interpersonal diversity remains largely unchanged. These findings suggest that our estimates are unlikely to reflect the influence of the intergenerational transmission of these cultural traits from the ancestral homeland to the US.

Table VII: Ancestral Diversity & Inequality: Accounting for Ancestral Culture

	Gini								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ancestral migratory distance	−0.063***		−0.062***	−0.064***		−0.063***	−0.063***		−0.069***
from the cradle of humanity	(0.013)		(0.014)	(0.017)		(0.018)	(0.013)		(0.016)
Ancestral Uncertainty Avoidance		−0.0022	−0.00060						
		(0.0030)	(0.0026)						
Ancestral Long-term Orientation					0.0027	0.0018			
					(0.0036)	(0.0036)			
Ancestral Individualism								0.0018	−0.0029
								(0.0026)	(0.0029)
Dep. var. mean	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Individuals	1743728	1743728	1743728	1742163	1742163	1742163	1743728	1743728	1743728
Demographic bins	543	543	543	519	519	519	543	543	543
Ancestral homelands	78	78	78	74	74	74	78	78	78
Adjusted R^2	0.22	0.19	0.22	0.20	0.18	0.20	0.22	0.19	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US remains robust when accounting for cultural characteristics in the ancestral homelands that may be conducive to inequality, and does not appear to be mediated by them: (i) *Uncertainty Avoidance* (Columns (2)-(3)), (ii) *Long-Term Orientation* (Columns (5)-(6)), and *Individualism* (Columns (8)-(9)). Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. The coefficients for all controls are standardized. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

5.5 Sorting Across Space

The demand for and return to productive traits vary across US regions, prompting individuals to distribute themselves spatially according to their productive capabilities. Consequently, individuals from the same ancestral origin—whose ancestors lived closer to the cradle of humanity in Africa and are therefore more diverse—would be expected to exhibit greater geographical dispersion across regions within the US.

Indeed, Table VIII suggests that individuals from more diverse homelands tend to be more geographically dispersed within the US. Notably, greater migratory distance from Africa is associated with lower dispersion across US states (Column 2), while higher dispersion of an ancestral group is associated with greater inequality (Column 3), partially attenuating the association between migratory distance and inequality (Column 4).

In urban areas, where there is sizable demand for a broad spectrum of skills, the relationship between diversity and inequality is likely to mirror the overall, unconditional pattern observed in the baseline analysis. Reassuringly, Panel A of Table C.22 indicates that the results remain nearly identical when the analysis is restricted to individuals residing in urban areas. Conversely, in rural areas, where demand is concentrated in a narrower range of skills, the relationship between diversity and inequality is less likely to reflect the baseline pattern. Consistent with this expectation, Panel B of Table C.22 indicates that the results are weaker when the analysis is restricted to rural residents.

Table VIII: Ancestral Diversity & Income Inequality: Sorting Across States

	Dispersion in Residence (State)	Gini		
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.16***	−0.069***		−0.062***
from the cradle of humanity	(0.045)	(0.012)		(0.013)
Dispersion in Residence (State)			0.080***	0.041
			(0.030)	(0.031)
Dep. var. mean	0.86	0.42	0.42	0.42
Individuals	1781548	1781548	1781548	1781548
Demographic bins	636	636	636	636
Ancestral homelands	95	95	95	95
Adjusted R^2	0.098	0.22	0.19	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is partly mediated by the extent of sorting of these individuals across US states. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications account for sex and age-group. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

6 Overcoming Selective Migration – Native Americans

The study focuses solely on individuals born in the US, who did not migrate to the US themselves. However, their productive traits may not mirror those prevalent in their ancestral homelands, due to conceivable selective migration patterns among their ancestors, in reaction to market conditions in the US. Our findings, therefore, could partly reflect selection in productive traits rather than the impact of a systematic and historically entrenched dispersion of these traits.

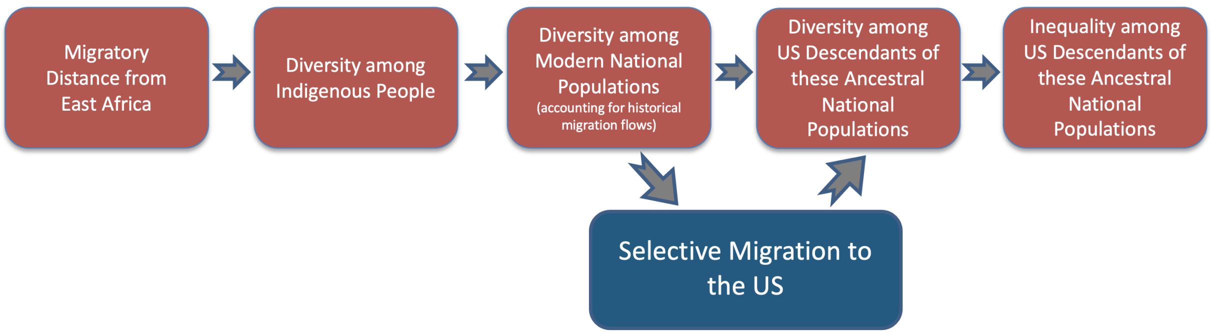


Figure 12: Selection in the Baseline Model

While selection may have shaped the characteristics of individuals who chose to migrate, thereby influencing the traits introduced into the US population, it is highly unlikely to have generated group-level selection that systematically compressed the variance of traits within immigrant groups originating farther from Africa. For selection to account our findings, it would require an exceedingly implausible form of group selection—one that consistently reduced trait variance among descendants of ancestral homelands farther from the cradle of humanity in Africa, despite no corresponding differences in the diversity of productive traits across those regions.

Nevertheless, despite the highly implausible role of selection in our findings, and to further address concerns about selection, we examine the roots of inequality among descendants of ancestral Native American tribes—groups that have not undergone selective migration over at least the past 10,000 years, which renders selection based on contemporary conditions in the US market infeasible.

Although Native Americans who arrived in the Americas before the submergence of the Bering land bridges due to postglacial sea-level rise more than 10,000 years ago may have been subject to selection into migration, there is evidence of a significant compression in the distribution of their traits following the crossing of the Bering Strait. As illustrated in Figure 13, there is a discernible negative association between prehistoric migratory distance from East Africa and folkloric diversity among 250 ethnic groups in North America.

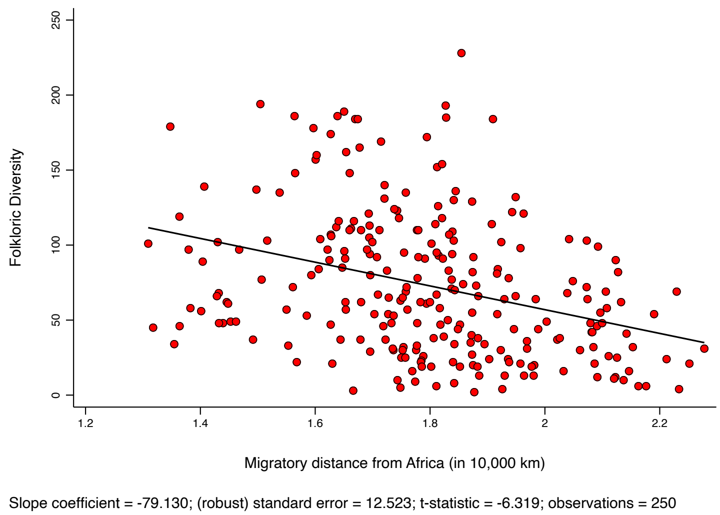


Figure 13: Declining Diversity along the Migratory Routes out-of-Africa in North America

Notes: This scatterplot depicts the empirical association between prehistoric migratory distance from East Africa and folkloric diversity across 250 ethnic groups in North America.

Thus, leveraging variation in migratory distance from Africa to the ancestral tribes of Native Americans, whose locations are shown in Figure 14, the study can address and mitigate concerns about the influence of selective migration in the post 1500 period.[28]

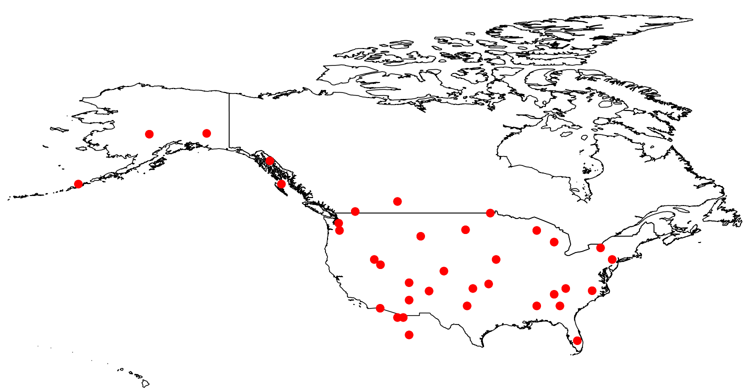


Figure 14: Locations of Native American Tribes

Notes: This figure depicts the locations of the ancestral Native American tribes in Table IX.

As reported in Table IX and depicted in Figure 15, the association between population diversity and contemporary inequality among Native Americans in the US remains sizable,

[28]Notably, the migratory distance from Alaska to Florida is nearly 8,200 km (i.e., 40% of the largest ancestry-adjusted distance from East Africa to any location on Earth), providing ample variation for a reliable estimation of the effect of diversity on inequality among Native Americans.

although the statistical significance is unavoidably lower due to the markedly smaller number of ancestral homelands. Native Americans originating from tribes located at greater migratory distances, and therefore predicted to be less diverse, are indeed less unequal: in the year 2010 (Column 1); in a repeated cross section over the period 1980 to 2020 (Column 2); and even within each of the various educational categories (Columns 3 and 4).

Table IX: Ancestral Diversity & Inequality: Native Americans

	Gini			
	Baseline		Within-education	
	ACS 5YR 2010	Repeated cross-section	ACS 5YR 2010	Repeated cross-section
	(1)	(2)	(3)	(4)
Ancestral migratory distance from East Africa	−0.15*	−0.16***	−0.14*	−0.14***
	(0.079)	(0.054)	(0.076)	(0.050)
Sample FE		✓		✓
Education FE			✓	✓
Dep. var. mean	0.41	0.41	0.37	0.37
Individuals	22524	82424	20706	76772
Demographic bins	246	808	456	1569
Ancestral homelands	32	36	32	36
Adjusted R^2	0.050	0.044	0.17	0.17

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa and (ii) income inequality holds within a sample of Native Americans. Migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin in a given survey period. Heteroskedasticity-robust standard errors (clustered at the tribe-level) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

Moreover, as shown in Table C.24, these findings remain qualitatively robust across a variety of alternative specifications, including: (i) grouping individuals in the US solely by ancestral origin or according to alternative permutations of sex and age; (ii) estimating weighted regressions based on population size; and (iii) incorporating fixed effects for US states and 170 micro-areas.

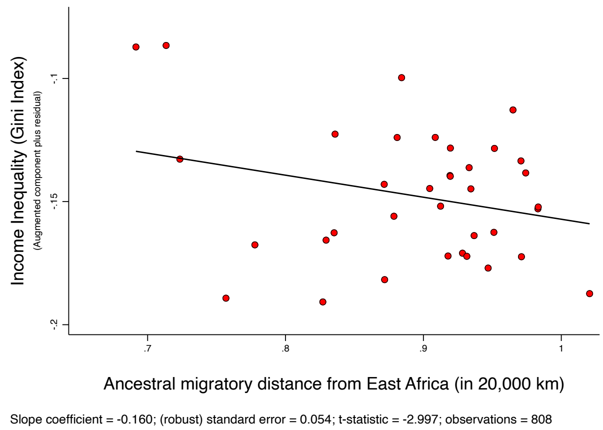


Figure 15: Migratory Distance from Africa & Inequality: Native Americans

Notes: This figure depicts a binned scatterplot of the association between income inequality among groups of individuals in the United States who share ancestry from the same Native American tribe, and the migratory distance from East Africa of their ancestral tribe.

7 Mediating Channels

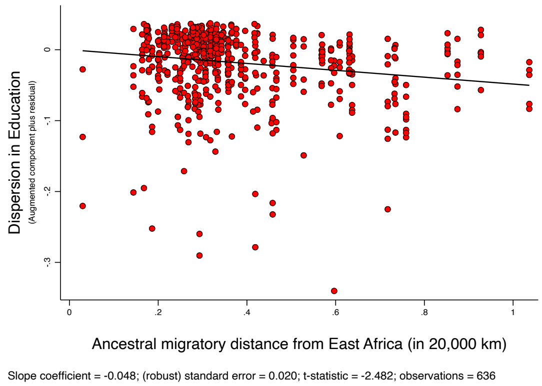
This section explores the mechanism through which a shorter prehistoric migratory distance from Africa, and the consequent increase in ancestral diversity, has contributed to higher levels of inequality. As implied by the proposed hypothesis, the effect of migratory distance on inequality operates through its influence on diversity in key productive traits. Specifically, we examine three complementary dimensions: (i) diversity in educational attainment, (ii) variation in labor supply, and (iii) dispersion in unobserved productivity-related traits, proxied by residual earnings inequality. These channels demonstrate how ancestral diversity has shaped the distribution of economic outcomes in the contemporary US population. Moreover, in line with the premise that greater interpersonal diversity raises the likelihood of a thicker upper tail in the distribution of entrepreneurial capabilities, the findings reveal that individuals in the U.S. with ancestral origins closer to the cradle of humanity in Africa tend to exhibit both higher levels of entrepreneurship and greater inequality, suggesting that diversity in productive traits may mediate inequality through its influence on the concentration of entrepreneurial talent

7.1 Dispersion in Education and Work Effort

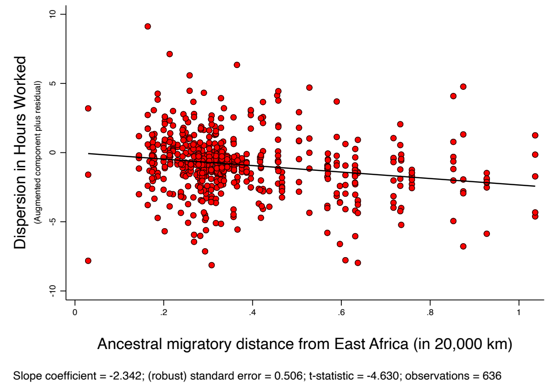
We begin by examining two observable dimensions of productive traits: educational attainment and labor supply. Our analysis reveals that individuals in the US whose ancestral populations originated closer to the cradle of humanity in Africa exhibit (i) greater educational diversity and (ii) more varied labor supply. Moreover, dispersion in these traits is associated with higher levels of income inequality, thereby mediating the impact of migratory distance from the cradle of humanity in Africa on inequality.

In line with the proposed hypothesis, demographic bins of US inhabitants whose ancestors originated closer to the cradle of humanity in Africa, and are therefore more diverse, exhibit: (i) greater dispersion in educational attainment, and (ii) greater variation in hours worked, plausibly reflecting a broader range of predispositions toward labor and leisure. Table X presents these mediating regressions, revealing a negative and statistically significant association between migratory distance from Africa and the dispersion in educational attainment and hours worked (Columns 1 and 2). These patterns, depicted in Figure 16, further validate the argument that the *Out-of-Africa Migration* and the associated *Serial Founder Effect* generated a compression in traits that has persisted to the present day.[29]

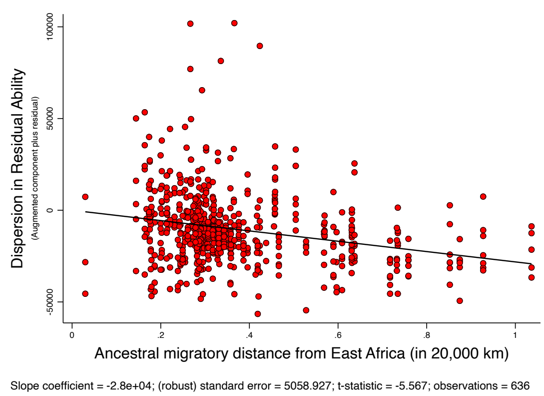
[29] Dispersion in these variables within each demographic bin is measured by: (i) one minus the Herfindahl Index of educational attainment, where education levels are aggregated into four categories based on the IPUMS classification: high school or below, some college, college, and more than college, and (ii) the standard deviation of hours worked.



(a) Education



(b) Labor Supply



(c) Residual Ability

Figure 16: Migratory Distance from East Africa to Ancestral Homelands and Dispersion in Education, Work Effort, and Residual Ability among their US Descendants

Notes: This figure depicts the association between migratory distance from East Africa to the ancestral homelands of US inhabitants and dispersion in traits among individuals in the US who trace their origins to these ancestral populations.

As expected, greater dispersion in these productive traits—education and hours worked—is positively and statistically significantly associated with income inequality, as measured by the Gini index (Columns 5 and 7) . Moreover, consistent with the view that these are indeed mediating channels, the point estimate of the effect of migratory distance from Africa on inequality (Column 4) declines when controlling for dispersion in education (Column 6) and hours worked (Column 8).[30]

[30] While, a priori, past income inequality might influence the dispersion of education, the evidence suggests that inequality during the period when education was formed is not significantly associated with educational dispersion. In particular, inequality in the 1980, 1990 and 2000 show no association educational dispersion in 2010 (Table C.25).

7.2 Dispersion in Residual Ability

To further investigate the role of ancestral diversity in shaping income inequality, we examine variation in unobserved productive traits—such as cognitive and non-cognitive abilities—using a proxy based on the dispersion of income residuals. Residual income are first estimated for each individual using an earnings regression that controls for observed characteristics, including age, sex, education, occupation, race, total weeks worked, usual hours worked per week, and area of residence. The standard deviation of these residuals is then computed within each demographic bin to capture the extent of latent heterogeneity in productivity. Greater residual dispersion plausibly reflects a broader range of unobserved traits influencing individual earnings within the group.[31]

In line with the proposed hypothesis, as depcited in Figure 16, individuals in the US whose ancestors originated closer to the cradle of humanity in Africa exhibit significantly greater dispersion in income residuals, suggesting a broader distribution of unobserved productive traits within these groups (Column 3). Consistent with this interpretation, residual dispersion is positively and statistically significantly associated with income inequality, as measured by the Gini index (Column 9). Furthermore, the estimated effect of migratory distance from Africa on inequality (Column 4) declines once residual dispersion is accounted for (Column 10), consistent with the notion that unobserved trait diversity mediates the link between ancestral origins and contemporary inequality.

[31]Given the substantial measurement error inherent in ancient DNA data, the dispersion in genetic predisposition toward cognitive ability cannot be reliably used as a mediator for the effect of ancestral migratory distance on contemporary inequality. In particular, postmortem degradation and sparse genomic coverage introduce significant error into polygenic score estimates of ancient individuals' predispositions toward cognitive ability. This variable can nonetheless be used to assess the impact of migratory distance from Africa on cognitive predisposition, as non-systematic measurement error in the outcome variable reduces statistical significance but does not bias the estimated coefficient. By contrast, using a noisy regressor, such as cognitive diversity, to explain contemporary inequality would attenuate the estimate toward zero.

Table X: Mediating Channels: Dispersion in Education, Work Effort & Residual Ability

	Dispersion in Education	Dispersion in Hours Worked	Dispersion in Residual Ability	Gini						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ancestral migratory distance	−0.048**	−2.34***	−1.23***	−0.069***		−0.063***		−0.055***		−0.021*
from the cradle of humanity	(0.020)	(0.51)	(0.22)	(0.012)		(0.013)		(0.013)		(0.011)
Dispersion in Education					0.16***	0.12**				
					(0.057)	(0.059)				
Dispersion in Hours Worked							0.0069***	0.0059***		
							(0.0014)	(0.0014)		
Dispersion in Residual Ability									0.040***	0.039***
									(0.0027)	(0.0028)
Dep. var. mean	0.69	11.0	2.0e-10	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Individuals	1781548	1781548	1781548	1781548	1781548	1781548	1781548	1781548	1781548	1781548
Demographic bins	636	636	636	636	636	636	636	636	636	636
Ancestral homelands	95	95	95	95	95	95	95	95	95	95
Adjusted R^2	0.040	0.061	0.31	0.22	0.19	0.23	0.23	0.26	0.50	0.51

Notes: This table explores potential mediating channels that may underlie the association between prehistoric migratory distance from Africa and income inequality, as captured by the Gini index. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

7.3 Entreprenership

Motivated by the premise that greater interpersonal diversity increases the likelihood of a denser upper tail in the distribution of skills conducive to entrepreneurship, we hypothesize that individuals whose ancestors originated from more diverse homelands are more likely to become entrepreneurs. Consistent with this prediction, the analysis reveals that demographic bins of U.S.-born individuals whose ancestors resided, on average, closer to the cradle of humanity in Africa, and who therefore exhibit greater ancestral diversity, display significantly higher rates of entrepreneurship.

Table XI: Mediating Channels: Entreprenership

	% Entrepreneurs	Gini		
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.053*** (0.017)	−0.069*** (0.012)		−0.056*** (0.013)
% Entrepreneurs			0.28*** (0.085)	0.23** (0.090)
Dep. var. mean	0.048	0.42	0.42	0.42
Individuals	1781548	1781548	1781548	1781548
Demographic bins	636	636	636	636
Ancestral homelands	95	95	95	95
Adjusted R^2	0.28	0.22	0.21	0.24

Notes: This table explores the role of entrepreneurship as a potential mediating channel that may underlie the association between prehistoric migratory distance from Africa and income inequality, as captured by the Gini index. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

Table XI presents this additional mediating regression. It first documents a negative and statistically significant association between migratory distance from Africa and the fraction of entrepreneurs (Column 1). This pattern, as depicted in Figure 17, further validates the argument that the Out-of-Africa Migration and the associated Serial Founder Effect generated a compression in the distribution of skills conducive to entrepreneurship. Importantly, the prevalence of entrepreneurship is positively associated with income inequality (Column 3), suggesting that entrepreneurial activity may constitute a potential mediating channel through which ancestral diversity affects inequality.[32] Notably, once entrepreneurship is introduced as a mediating variable, the point estimates of the effect of migratory distance on inequality decline relative to the baseline reduced-form specification (Column 4), consistent with the potential mediating role of entrepreneurship.

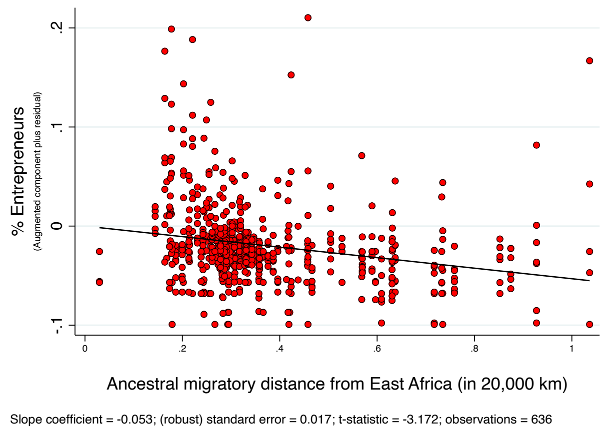


Figure 17: Migratory Distance from East Africa to Ancestral Homelands and Entreprenership among their US Descendants

Notes: This figure depicts the association between migratory distance from East Africa to the ancestral homelands of US inhabitants and the share of entrepreneurs among individuals in the US who trace their origins to these ancestral populations.

8 Concluding Remarks

We advance a novel hypothesis that, in a market economy where earning differentials reflect variations in productive traits across individuals, a significant component of the differences in inequality across societies can be attributed to ancestral variation in productive traits shaped during the prehistoric Out-of-Africa migration.

Exploring the roots of inequality within the US population, and leveraging rich micro-data on millions of US-born individuals spanning nearly one hundred ancestral origins, we find supporting evidence for our hypothesis. The findings hold across various samples, including a Native American sample that consists exclusively of individuals who have not been subjected to selective migration into the US in the post-1500 period. Moreover, the results are robust to the inclusion of potentially confounding geographical characteristics that may be correlated with migratory distance from Africa, and to the potentially confounding impact

[32]Following prevailing wisdom, we define entrepreneurs as incorporated self-employed individuals and compute their share within each demographic bin. Incorporated businesses are more likely to engage in activities requiring nonroutine cognitive skills.

of ancestral ethnolinguistic fragmentation, inequality, and cultural and institutional characteristics—factors that may be associated with ancestral diversity. Diversity in productive traits accounts for an astounding 12.5% of the overall inequality in the US.

Notably, the association between ancestral diversity and income inequality is mediated by variation in key productive traits. Individuals in the U.S. whose ancestors originated closer to the cradle of humanity in Africa exhibit greater heterogeneity across multiple dimensions of productivity: (i) higher dispersion in educational attainment, (ii) more varied labor supply as reflected in hours worked, and (iii) broader residual ability, as proxied by income residuals. This heightened diversity in both observed and unobserved traits is, in turn, associated with higher levels of income inequality—supporting the hypothesis that ancestral variation in productive traits, shaped during the prehistoric exodus of humans from Africa, underlies contemporary disparities in economic outcomes. Moreover, in line with the premise that greater interpersonal diversity raises the likelihood of a denser upper tail in the distribution of entrepreneurial skills, the analysis indicates that U.S. residents with ancestral origins closer to the cradle of humanity in Africa tend to exhibit both higher rates of entrepreneurship and greater income inequality, consistent with the hypothesis that dispersion in productive traits may also mediate inequality through its effect on the prevalence of entrepreneurial talent.

The findings suggest that implementations of growth-enhancing diversity policies ought to be considered in the context of the trade-off between growth and inequality. Fostering diversity and thus growth in societies that are relatively homogeneous would be associated with greater inequality, whereas fostering social cohesion in societies that are highly diverse may promote growth while mitigating inequality.

References

Aghion, P., E. Caroli, and C. Garcia-Penalosa (1999): "Inequality and economic growth: the perspective of the new growth theories," *Journal of Economic literature*, 37, 1615–1660.

Alesina, A., J. Harnoss, and H. Rapoport (2016a): "Birthplace diversity and economic prosperity," *Journal of Economic Growth*, 21, 101–138.

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

Arbatli, C. E., Q. H. Ashraf, O. Galor, and M. Klemp (2020): "Diversity and conflict," *Econometrica*, 88, 727–797.

Ashraf, Q. and O. Galor (2013a): "Genetic diversity and the origins of cultural fragmentation," *American Economic Review*, 103, 528–533.

Ashraf, Q. and O. Galor (2013b): "The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, 103, 1–46.

Ashraf, Q. H., O. Galor, and M. Klemp (2021): "The ancient origins of the wealth of nations," in *The Handbook of Historical Economics*, Elsevier, 675–717.

Atkinson, Q. D. (2011): "Phonemic diversity supports a serial founder effect model of language expansion from Africa," *Science*, 332, 346–349.

Benabou, R. (2000): "Unequal societies: Income distribution and the social contract," *American Economic Review*, 91, 96–129.

Betti, L., F. Balloux, W. Amos, T. Hanihara, and A. Manica (2009): "Distance from Africa, not climate, explains within-population phenotypic diversity in humans," *Proceedings of the Royal Society B: Biological Sciences*, 276, 809–814.

Betti, L. and A. Manica (2018): "Human variation in the shape of the birth canal is significant and geographically structured," *Proceedings of the Royal Society B*, 285.

Case, A. and C. Paxson (2008): "Stature and status: Height, ability, and labor market outcomes," *Journal of political Economy*, 116, 499–532.

Chancel, L., T. Piketty, E. Saez, and G. Zucman (2022): "World inequality report 2022," .

Conley, T. G. (1999): "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 92, 1–45.

Cook, C. J. and J. M. Fletcher (2018): "High-school genetic diversity and later-life student outcomes: micro-level evidence from the Wisconsin Longitudinal Study," *Journal of Economic Growth*, 23, 307–339.

De La Croix, D. and M. Doepke (2003): "Inequality and growth: why differential fertility matters," *American Economic Review*, 93, 1091–1113.

Desmet, K., S. Weber, and I. Ortuño-Ortín (2009): "Linguistic diversity and redistribution," *Journal of the European Economic Association*, 7, 1291–1318.

Fenske, J. (2014): "Ecology, Trade, and States in Pre-Colonial Africa," *Journal of the European Economic Association*, 12, 612–640.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, D. Backman, A. Chen, G. Cooper, S. Richards, M. Schouweiler, and M. Westberry (2024): "IPUMS CPS: Version 12.0 [dataset]," *Minneapolis, MN: IPUMS.*

Forkel, R. et al. (2022): "Glottocodes: Identifiers linking families, languages and dialects to comprehensive reference information," *Semantic Web*, 13.

Galor, O. and M. Klemp (2023): "The Roots of Autocracy," *Mimeo, Brown University, Department of Economics.*

Galor, O., M. Klemp, and D. C. Wainstock (2024): "Roots of Cultural Diversity," *National Bureau of Economic Research.*

Galor, O., M. Klemp, and D. C. Wainstock (2025): "The Prehistoric Dispersal from Africa and the Compression of Within-Group Productive Trait," *Mimeo, Brown University, Department of Economics.*

Galor, O. and O. Moav (2000): "Ability-biased technological transition, wage inequality, and economic growth," *The quarterly journal of economics*, 115, 469–497.

Galor, O. and O. Özak (2016): "The Agricultural Origins of Time Preference," *American Economic Review*, 106, 3064–3103.

Galor, O. and J. Zeira (1993): "Income distribution and macroeconomics," *The review of economic studies*, 60, 35–52.

Guiso, L., P. Sapienza, and L. Zingales (2006): "Does culture affect economic outcomes?" *Journal of Economic perspectives*, 20, 23–48.

Hofstede, G., G. J. Hofstede, and M. Minkov (2010): "Cultures and Organizations: Software of the Mind. Revised and expanded 3rd Edition," *N.-Y.: McGraw-Hill*, 560.

Katz, L. F. and K. M. Murphy (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107, 35–78.

Kirby, K. R., R. D. Gray, S. J. Greenhill, F. M. Jordan, S. Gomes-Ng, H.-J. Bibiko, D. E. Blasi, C. A. Botero, C. Bowern, C. R. Ember, et al. (2016): "D-PLACE: A global database of cultural, linguistic and environmental diversity," *PloS one*, 11.

Klor, E. F. and M. Shayo (2010): "Social identity and preferences over redistribution," *Journal of Public Economics*, 94, 269–278.

Manica, A., W. Amos, F. Balloux, and T. Hanihara (2007): "The effect of ancient population bottlenecks on human phenotypic variation," *Nature*, 448, 346–348.

Marshall, M. G. and T. R. Gurr (2020): "Polity 5: Political Regime Characteristics and Transitions, 1800-2018," *Center for Systemic Peace.*

Montalvo, J. G. and M. Reynal-Querol (2005): "Ethnic polarization, potential conflict, and civil wars," *American economic review*, 95, 796–816.

Murdock, G. P. (1967): "Ethnographic atlas: a summary," *Ethnology*, 109–236.

Piketty, T. (2014): "Capital in the twenty-first century," *Harvard University Press*.

Posch, M., J. Schulz, and J. Henrich (2025): "How Cultural Diversity Drives Innovation," *Journal of Political Economy*.

Putterman, L. and D. N. Weil (2010): "Post-1500 Population Flows and the Long Run Determinants of Economic Growth and Inequality," *Quarterly Journal of Economics*, 125, 1627–1682.

Ragsdale, A. P., T. D. Weaver, E. G. Atkinson, E. G. Hoal, M. Möller, B. M. Henn, and S. Gravel (2023): "A weakly structured stem for human origins in Africa," *Nature*, 1–9.

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

Rosen, S. (1981): "The economics of superstars," *The American economic review*, 71, 845–858.

Ruggles, S., S. Flood, M. Sobek, D. Backman, A. Chen, G. Cooper, S. Richards, R. Rogers, and M. Schouweiler (2023): "IPUMS USA: Version 14.0 [dataset]," *Minneapolis, MN: IPUMS*.

Sunde, U., T. Dohmen, B. Enke, A. Falk, D. Huffman, and G. Meyerheim (2022): "Patience and comparative development," *The Review of Economic Studies*, 89, 2806–2840.

von Cramon-Taubadel, N. and S. J. Lycett (2008): "Brief communication: human cranial variation fits iterative founder effect model with African origin," *American Journal of Physical Anthropology*, 136, 108–113.

Supplemental Appendix

Contents

A Variable Definitions and Sources

A.1 Ancestral Homeland

- **Self-reported ancestry of the US population:** We follow the coding of the variable ”ancestr1d” (i.e., detailed ancestry, first response) in IPUMS USA to match the self-reported ancestry to a modern national boundary,[33] where the set of nations is based on the classification of the World Bank Development Indicators.[34] If IPUMS does not match an ancestry to a modern national boundary we establish the following assignment criteria: (i) if the ancestry is assigned unambiguously by historical sources to a unique modern national homeland, then we follow this assignment (e.g., Cornish and Manx as part of the United Kingdom), (ii) if the ancestry is a former nation that split up (e.g., Czechslovakia and Yugoslavia), we match the ancestry to the contemporary country of the historical capital, (iii) if the ancestry is an ethnic group (that is not mapped by IPUMS to a modern nation), we use the assignment provided by the Ethnographic Atlas (e.g., Kurds and Lapps), (iv) if the group is not in the Ethnographic Atlas (e.g., Cossack), then we match it to the closest capital of a contemporary country where this group is currently located, (v) individuals who report an ancestry which can not be mapped into an a unique ancestral homeland (e.g., African Americans, Europeans, Latin Americans, White) are excluded from the analysis. Data Source: Authors' assignment based on Ruggles et al. (2023).

- **Native American Tribe:** We follow the coding of the variable ”tribed” (i.e., detailed tribe) in IPUMS USA to match Native Americans to their ancestral homeland. [35] Data Source: Authors' assignment based on Ruggles et al. (2023).

- **Parental country of birth:** We follow the coding of the variable ”fbpl” (i.e., father's birthplace) in IPUMS CPS to match individuals to their ancestral homeland according to their paternal lineage.[36] Data Source: Authors' assignment based on Flood et al. (2024).

[33]Self-reported ancestry seems to have a negligible impact on the analysis. Using data from second-generation migrants in the *Current Population Survey* (CPS), where the parental country of birth is directly observed, yields qualitatively similar results, as long as the sample retains a sufficient number of ancestries (Table C.6). Additionally, 45% of individuals in our ACS sample report primary and secondary ancestries. Analyzing (i) individuals identifying only a primary ancestry (Table C.7) or (ii) the average migratory distances of both ancestries (Table C.8) does not alter the findings.

[34]Hong Kong and Taiwan are classified as additional clusters while Macau is not following the convention (e.g., Putterman and Weil (2010)).

[35]We consider (i) Pueblo and Pueblo-Hopi, (ii) Tlingit-Haida and Haida, and (iii) Eskimo, Yupik, and Inupiat to be the same ancestral homeland.

[36]The repeated cross-section CPS sample of second-generation migrants is significantly smaller than the ACS sample, encompassing approximately 100,000 individuals from 74 ancestral homelands, compared to nearly 8 million individuals from 108 ancestral homelands in the ACS. This smaller sample size increases the susceptibility of CPS estimates to measurement error and reduces their representativeness, particularly when further stratifying bins. To address these limitations, we define ancestry based on paternal lineage, as using maternal lineage reduces the number of ancestral homelands to 72, and requiring identical paternal and maternal lineages further narrows it to 49. Importantly, our qualitative results remain consistent regardless of whether ancestry is defined by maternal lineage or by identical paternal and maternal lineages. Moreover, we rely on the repeated cross-section design of the CPS precisely to mitigate the limitations imposed by the small sample size, as it allows us to pool information across years and reduce the impact of measurement errors.

A.2 Income Inequality

- **Gini**: The Gini index of earned income within each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).
- **Top 1%**: The share of earned income held by the top 1% within each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).
- **Top 5%**: The share of earned income held by the top 5% within each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).
- **Top 10%**: The share of earned income held by the top 10% within each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).

A.3 Migratory Distance from the Cradle of Humanity

In alignment with archaeological evidence, the migratory path permits crossings of the Bab-el-Mandeb Strait, the Strait of Hormuz, and the Bering Strait while excluding the crossing of the Strait of Gibraltar and the Himalayan mountain range. Due to the limited capacity of early humans to traverse large bodies of water, migratory paths were confined to land masses. However, considering that sea levels at the time of the exodus during the Ice Age were nearly 100 meters lower than today, along with the ability of humans to cross shallow water, we permit travel within 100 km of the shoreline. Varying the buffer to 25 km or 50 km does not alter the results. Moreover, given the uncertain dynamics of the Serial Founder Effect over long maritime distances, it is assumed that oceanic travel to remote islands has no additional impact on diversity. Yet, permitting such an impact does not qualitatively affect our findings.

- **Modern National Populations:** The migratory distance to each country is defined as the shortest traversable path from Omo I (Ethiopia)—the site of the earliest known Homo sapiens remains in East Africa—to the country's modern capital city. Data Source: Authors' computations.[37]
- **Native American Tribes:** The migratory distance to each tribe is defined as the shortest traversable path from Omo I (Ethiopia)—the site of the earliest known Homo sapiens remains in East Africa—to the pair of coordinates of the corresponding tribe on Ethnographic Atlas or

[37] Since the ancestral homeland may consist of population which are themselves from different ancestries, the ancestry-adjusted migratory distance from Africa to the ancestral homeland captures the weighted average of the migratory distances from Africa of each of these ancestral populations, accounting for the proportional representation of these deeper ancestral populations in the ancestral homeland, using the migration matrix of Putterman and Weil (2010). If the ancestral homeland is not in the matrix, we keep the unadjusted migratory distance only if the homeland is in the Old World given the drastic changes in the composition of populations of the New World in the post-1500 period. To address potential concerns about measurement error, we demonstrate that our findings remain qualitatively robust even when restricting the analysis to demographic bins corresponding to ancestral homelands in the Old World, where the modern population is predominantly native to its territory (Table C.14).

Glottolog.[38] Data Source: Authors' computation based on Murdock (1967), Kirby et al. (2016), and Forkel et al. (2022).

A.4 Fixed-Effects

- **Sex**: Each individual's sex. Data Source: Ruggles et al. (2023).
- **Age group**: Each individual's age group: 15-24, 25-34, 35-44, 45-54, or 55-64. Data Source: Authors' computation based on Ruggles et al. (2023).
- **Sample**: The repeated cross-section include five different samples: Censuses of 1980, 1990, and 2000 as well as ACS 5-year samples of 2010 and 2020. Data Source: Ruggles et al. (2023).
- **Continental Fixed-Effects**: Dummy variables capturing the location of each ancestral homeland of the US population in either: Africa, Asia, Europe, Americas, or Oceania. Data Source: Authors' assignment.

A.5 Size of demographic bin

- **Size of demographic bin**: Number of individuals in a demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).

A.6 Ancestral Geographical Controls

- **Absolute latitude**: The absolute value of the latitude of the geodesic centroid of each ancestral homeland of the US population. Data Source: Authors' computation.
- **Ecological Diversity**: Standard deviation and mean of caloric suitability within the territory of each ancestral homeland. Data Source: Authors' computation based on Galor and Özak (2016).
- **Island**: A dummy variable that captures whether each ancestral homeland of the US population is located on a small island country as defined by the United Nations M49 classification. Data Source: United Nations Statistics Division.

A.7 Ancestral Ethnic Fragmentation Controls

- **Ethnic Fractionalization**: The index captures the probability that two individuals in a country share the same ethnicity. Data Source: Alesina et al. (2003).

[38]The coordinates of Alaskan Athabaskan are the average of all Central Alaska-Yukon Athabaskan or Southern Alaskan Athabaskan groups in Glottolog. The coordinates of Sioux are the average of Lakota and Dakota in Glottolog. The coordinates of Eskimo are the average of Yupik and Inupiat groups at Alaska in Glottolog. The coordinates of Sioux are the average of Yuman groups in Glottolog. The coordinates of Pueblo are the average of all groups in Glottolog which are in either the Keresan family, the Kiowa-Tanoan family, Zuni, or Hopi.

- **Ethnolinguistic Fractionalization**: The index captures the probability that two individuals in a country share the same ethnicity, weighted by their linguistic distance. This is also known as the Greenberg index. Data Source: Desmet et al. (2009).

A.8 Ancestral Inequality Controls

- **Class stratification**: The degree and type of class differentiation, excluding purely political and religious statuses. The original variable records ethnic groups as belonging to one of the following categories: (1) absence of significant class distinctions among freemen, (2) wealth distinctions, (3) elite stratification, (4) dual stratification, and (5) complex stratification. Using this information, we take as evidence of class stratification if the original variable takes on the value of 2, 3, 4, or 5. Data Source: Murdock (1967) and Giuliano and Nunn (2016).

- **Slavery**: The forms and prevalence of slave status, treated quite independently of both class and caste status. The original variable records ethnic groups as belonging to one of the following categories: (1) absence or near absence of slavery, (2) incipient or nonhereditary slavery (i.e., where slave status is temporary and not transmitted to the children of slaves), (3) slavery reported but not identified as hereditary or nonhereditary, and (4) hereditary slavery present and of at least modest social significance. Using this information, we take as evidence of slavery if the original variable takes on the value of 2, 3, or 4. Data Source: Murdock (1967) and Giuliano and Nunn (2016).

- **Share of income held by the top 10%**: The share of income held by the top 10% during in 1900. Data Source: Chancel et al. (2022).

- **Gini**: The Gini index during the time period 1980-1999. Data Source: World Bank Development Indicators.

- **Ethnic Inequality**: The Gini index of mean luminosity per capita across ethnic homelands (GREG) within a given country. Data Source: Alesina et al. (2016b).

A.9 Ancestral Cultural Controls

- **Uncertainty Avoidance**: The degree to which individuals avoid ambiguity and uncertainty. Data Source: Hofstede et al. (2010).

- **Long-Term Orientation**: The fostering of virtues oriented towards future rewards, in particular, perseverance and thrift. Data Source: Hofstede et al. (2010).

- **Individualism**: The degree of interdependence a society maintains among its members. The importance placed on attaining personal goals. Source: Hofstede et al. (2010).

- **Early settlers**: Ancestral homelands which accounted for a significant share of the US population in 1850, which is the oldest full-count individual-level Census available. We define early settlers as the top 10 foreign birthplaces at that time. Data Source: Authors' computation based on Ruggles et al. (2023).

A.10 Ancestral Institutional Controls

- **Colonial legacy**: Ancestral homelands which have been subjected to European colonization in the post-1500 era. Data Source: Arbatlı et al. (2020).
- **Jurisdictional Hierarchy**: The number of jurisdictional levels beyond the local community, ranging from 1 for stateless societies, through 2 or 3 for petty and larger paramount chiefdoms or their equivalent, to 4 or 5 for large states. Polities imposed recently by colonial regimes are excluded. Data Source: Murdock (1967) and Giuliano and Nunn (2016).
- **Polity score**: Polity V Project provides a polity score based on the subtraction of the autocracy score from the democracy score. Data Source: Marshall and Gurr (2020).

A.11 Mediating Channels

- **Dispersion in Hours Worked**: The standard deviation of hours worked within each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).
- **Dispersion in Education**: 1 - [The Herfindahl index of educational categories (i.e., high school or below, some college, college, and more than college)] within each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).
- **Dispersion in Residual Ability**: Income residuals are obtained from a saturated earnings regression controlling for age, sex, race, educational attainment, occupation, total weeks worked, usual hours worked per week, and micro-area fixed effects. The within-bin dispersion of these residuals is measured by the standard deviation of residual earnings for each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).
- **Entreprenership**: Share of incorporated self-employed individuals within each demographic bin. Data Source: Authors' computation based on Ruggles et al. (2023).

B Summary Statistics

Table B.1: Summary Statistics

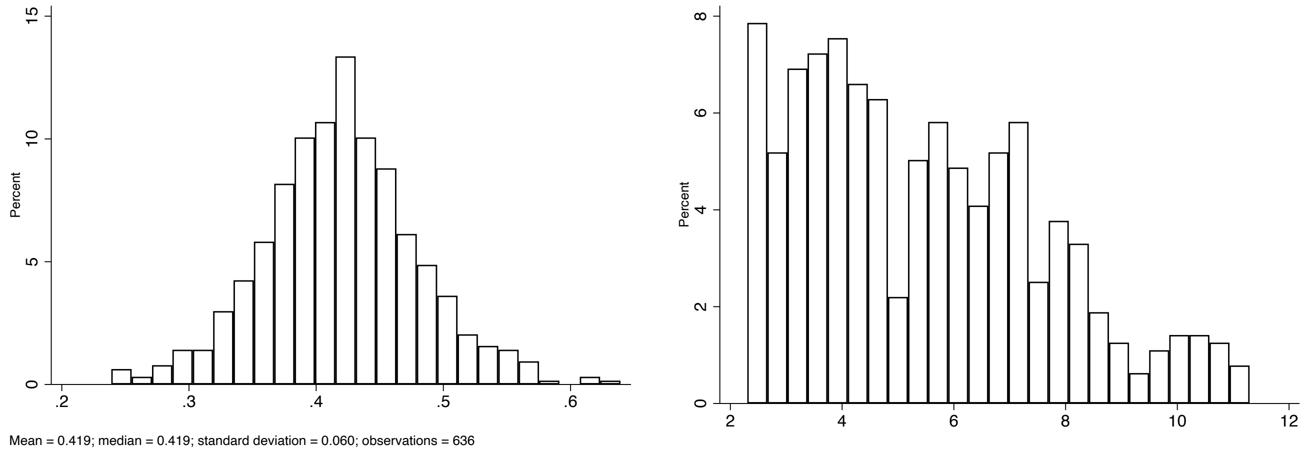
	Mean	SD	Median	Min	Max	N
A. Dependent variables						
Gini index	0.42	0.06	0.42	0.24	0.64	636
Share of income held by the top 1%	0.07	0.02	0.07	0.03	0.14	354
Share of income held by the top 5%	0.20	0.05	0.19	0.10	0.41	559
Share of income held by the top 10%	0.30	0.06	0.30	0.17	0.53	636
B. Independent variables						
Ancestral migratory distance from East Africa	7,449	3,695	6,392	595	20,730	95
C. Size of demographic bin						
Size of demographic bin	2,801	8,950	182	10	80,326	636
D. Ancestral geography						
Absolute latitude	36	14	36	6	64	95
Caloric suitability (s.d.)	1,688	922	1,644	0	3,994	95
Caloric suitability (mean)	6,787	2,465	7,653	0	10,108	95
Island	0.00	0.02	0	0	0	95

Notes: The table provides for all variables used in the data analysis the mean, the standard deviation (SD), the median, the minimum value (MIN), the maximum value (MAX), and the number of observations (N).

Table B.2: Summary Statistics

	Mean	SD	Median	Min	Max	N
E. Ancestral ethnic fragmentation						
Ethnic fractionalization	0.37	0.23	0.37	0	0.85	91
Ethnolinguistic fractionalization	0.14	0.16	0.08	0	0.65	93
F. Ancestral inequality						
Class stratification	0.96	0.12	1	0	1	93
Slavery	0.27	0.40	0.01	0	1	92
Share of income held by the top 10% (1900)	0.51	0.05	0.51	0	0.58	21
Gini index (1980-1999)	0.40	0.11	0.36	0	0.59	66
Ethnic inequality	0.41	0.24	0.41	0	0.95	91
G. Ancestral Cultural Factors						
Uncertainty avoidance	68	21	68	13	112	78
Long term orientation	46	24	46	4	100	74
Individualism	41	23	36	6	90	78
Early settlers	0.11	0.31	0	0	1	95
H. Ancestral Institutions						
Colonial legacy	0.52	0.50	1	0	1	89
Jurisdictional hierarchy	2.93	0.53	3	1	4	93
Polity score (1900)	-0.87	5.97	-2	-10	10	47
I. Mediating Channels						
Dispersion in Hours Worked	11.0	2.1	11.1	4	21.6	636
Dispersion in Education	0.69	0.05	0.71	0.36	0.75	636
Dispersion in Residual Ability	45,473	22,956	39,041	9,629	161,290	636
% Entrepreneurship	0.05	0.05	0.03	0.00	0.30	636

Notes: The table provides for all variables used in the data analysis the mean, the standard deviation (SD), the median, the minimum value (MIN), the maximum value (MAX), and the number of observations (N).



(a) Gini index (b) Log Bin size

Figure B.1: The Structure of Demographic Bins

Notes: This figure depicts the histograms of: (a) inequality across demographic bins as captured by the Gini index, and (b) the distribution of the log number of individuals in each demographic bin.

C Robustness Checks - Tables

C.1 Raw Ancestry Bins

Table C.1: Robustness to Raw Ancestry Bins

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.11***	−0.022***	−0.084***	−0.088***
from the cradle of humanity	(0.026)	(0.0078)	(0.020)	(0.023)
Dep. var. mean	0.44	0.073	0.22	0.32
Individuals	1782313	1781351	1782215	1782313
Demographic bins	108	83	102	108
Ancestral homelands	108	83	102	108
Adjusted R^2	0.13	0.090	0.12	0.10

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US holds within raw demographic bins defined exclusively by ancestry. Ancestry-adjusted migratory distance is measured in units of 20,000 km. The unit of observation is a bin at the level of ancestry. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.2 Alternative Demographic Bins

Table C.2: Robustness to Alternative Demographic Bins

	Gini			
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.11*** (0.026)	−0.092*** (0.020)	−0.080*** (0.014)	−0.069*** (0.012)
Sex FE		✓		✓
Age FE			✓	✓
Dep. var. mean	0.44	0.43	0.43	0.42
Individuals	1782313	1782227	1782012	1781548
Demographic bins	108	204	354	636
Ancestral homelands	108	106	104	95
Adjusted R^2	0.13	0.11	0.29	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US holds when sex and age fixed effects are accounted for separately. The unit of observation is a bin at the level of: (a) ancestry, (b) ancestry and sex, (c) ancestry and age-group, (d) baseline demographic bin. Ancestry-adjusted migratory distance is measured in units of 20,000 km. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.3 Employment Status and Working Age

Table C.3: Robustness to Alternative Employment Status and Working Age

	Gini				
	Labor Force	Employed	Private Sector	Prime Working Age	Full-Time
	(1)	(2)	(3)	(4)	(5)
Ancestral migratory distance from the cradle of humanity	−0.069*** (0.012)	−0.066*** (0.013)	−0.082*** (0.015)	−0.072*** (0.014)	−0.063*** (0.017)
Dep. var. mean	0.42	0.40	0.42	0.41	0.35
Individuals	1781548	1699703	1247275	1371829	1273130
Demographic bins	636	630	589	510	590
Ancestral homelands	95	94	90	95	90
Adjusted R^2	0.22	0.26	0.28	0.20	0.34

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected when the specification is estimated on the sample of: (i) all working age individuals, (ii) working age individuals in the labor force, (iii) employed working age individuals, (iv) employed in the private sector, or (v) only full time workers. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications include sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity robust standard errors (clustered at the ancestral origins of the US population) is reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.4 Classifications of Age Group

Table C.4: Robustness to Alternative Classifications of Age Group

	Gini		
Split in groups of:	20-year	10-year	5-year
	(1)	(2)	(3)
Ancestral migratory distance from the cradle of humanity	−0.071*** (0.017)	−0.069*** (0.012)	−0.071*** (0.013)
Dep. var. mean	0.43	0.42	0.41
Individuals	1782029	1781548	1780531
Demographic bins	361	636	1109
Ancestral homelands	104	95	90
Adjusted R^2	0.12	0.22	0.24

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected by alternative age groupings. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications include sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity robust standard errors (clustered at the ancestral origins of the US population) is reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.5 Men vs. Women

Table C.5: The Impact of Diversity on Inequality: Men vs. Women

	Gini		
	(1)	(2)	(3)
Ancestral migratory distance from the cradle of humanity	−0.069*** (0.012)	−0.072*** (0.015)	−0.065*** (0.017)
Men	✓	✓	
Women	✓		✓
Dep. var. mean	0.42	0.43	0.41
Individuals	1781548	979668	801880
Demographic bins	636	322	314
Ancestral homelands	95	95	91
Adjusted R^2	0.22	0.19	0.24

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US, as measured by the Gini index, is similar in samples consisting exclusively of either men or women (Columns (1)–(3)), and is not markedly affected by the extent of their employment (Columns (4)–(6)). Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

It should be noted that the effect of diversity on inequality is larger for men among individuals that are fully employed and larger for women among all individuals.

C.6 Second-Generation Migrants

Table C.6: Second-Generation Migrants

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.052***	−0.0076	−0.042**	−0.036***
	(0.015)	(0.014)	(0.016)	(0.014)
Dep. var. mean	0.41	0.069	0.19	0.29
Individuals	108993	73735	102201	108993
Demographic bins	1402	249	909	1402
Ancestral homelands	74	15	52	74
Adjusted R^2	0.017	0.084	0.039	0.019

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected when the hypothesis is tested using a repeated cross-section of second-generation migrants in the Current Population Survey (1994–2024). Migratory distance is measured in units of 20,000 km. All specifications account for survey year fixed-effects. The unit of observation is a bin at the level of ancestry. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.7 Individuals Reporting Multiple Ancestries - Distance to Primary Ancestry

Table C.7: Inclusion of Individuals Reporting Secondary Ancestry – Migratory Distance to the Primary Ancestry

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.054***	−0.019***	−0.056***	−0.053***
	(0.014)	(0.0048)	(0.011)	(0.016)
Dep. var. mean	0.42	0.066	0.20	0.30
Individuals	3279700	3268894	3278746	3279700
Demographic bins	666	408	596	666
Ancestral homelands	102	74	90	102
Adjusted R^2	0.18	0.15	0.27	0.20

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is qualitatively unaffected when individuals reporting a secondary ancestry are included and assigned to demographic bins based on their primary ancestry. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.8 Individuals Reporting Multiple Ancestries – Weighted Migratory Distance to the Primary and Secondary Ancestries

Table C.8: Inclusion of Individuals Reporting Secondary Ancestry – Weighted Migratory Distance to the Primary and Secondary Ancestries

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.071*** (0.013)	−0.022*** (0.0049)	−0.052*** (0.0084)	−0.067*** (0.010)
Dep. var. mean	0.41	0.065	0.19	0.29
Individuals	3334570	3255002	3320655	3334570
Demographic bins	3814	1370	2797	3814
Ancestral homelands	653	219	454	653
Adjusted R^2	0.092	0.077	0.14	0.11

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is qualitatively unaffected when individuals reporting a secondary ancestry are included, assigned to demographic bins based on their mixed ancestry, and ancestral diversity for each bin is predicted using the average prehistoric migratory distance of the two ancestries. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.9 Weighted Regression

Table C.9: Weighted Regression

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.080*** (0.017)	−0.034*** (0.0024)	−0.10*** (0.011)	−0.10*** (0.013)
Dep. var. mean	0.44	0.073	0.22	0.32
Individuals	1782337	1781351	1782215	1782313
Demographic bins	111	83	102	108
Ancestral homelands	111	83	102	108
Adjusted R^2	0.28	0.63	0.55	0.46

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is robust to an alternative estimation method that accounts for the differential prevalence of each ancestral group in the US—namely, a weighted least squares regression in which weights are based on group size in the sample. Migratory distance is measured in units of 20,000 km. The unit of observation is a bin at the level of ancestry. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.10 Separating Earned Income into Wage and Business Income

Table C.10: Accounting for either Wage or Business Income

	Gini		
	Wage	Business Income	Earned Income
	(1)	(2)	(3)
Ancestral migratory distance from the cradle of humanity	−0.067*** (0.015)	−0.098*** (0.028)	−0.069*** (0.012)
Dep. var. mean	0.40	0.70	0.42
Individuals	1567708	212773	1781548
Demographic bins	617	372	636
Ancestral homelands	93	66	95
Adjusted R^2	0.13	0.16	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected when the analysis is restricted to wage workers (Column (1)) or self-employed individuals, with inequality computed based on their business income (Column (2)). Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.11 Mean Income

Table C.11: Accounting for Mean Income

	Gini		
	(1)	(2)	(3)
Ancestral migratory distance from the cradle of humanity	−0.081*** (0.0096)		−0.072*** (0.0099)
Log mean income		0.068*** (0.012)	0.060*** (0.012)
Dep. var. mean	0.38	0.38	0.38
Individuals	1778218	1778218	1778218
Demographic bins	1841	1841	1841
Ancestral homelands	88	88	88
Adjusted R^2	0.14	0.13	0.16

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected by the mean income in the demographic bin. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex, age-group, and educational categories fixed-effects. The unit of observation is a bin at the level of ancestry, sex, age-group, and educational attainment. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.12 Topcoded Individuals

Table C.12: Ancestral Diversity & Income Inequality: Pareto-Adjustment of Top-coded Incomes

	GINI	TOP 1%	TOP 5%	TOP 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.074***	−0.026***	−0.060***	−0.063***
from the cradle of humanity	(0.013)	(0.0068)	(0.014)	(0.014)
Dep. var. mean	0.42	0.071	0.20	0.31
Individuals	1781548	1770406	1780511	1781548
Demographic bins	636	354	559	636
Ancestral homelands	95	61	87	95
Adjusted R^2	0.23	0.21	0.27	0.23

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected when a standard Pareto-imputation procedure is applied to account for topcoded income. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.13 Time since Settlement in the US

Table C.13: Accounting for Time since Settlement in the US

	GINI			
	INCL. EARLY SETTLERS			EXCL. EARLY SETTLERS
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.069***		−0.068***	−0.071***
from the cradle of humanity	(0.012)		(0.012)	(0.013)
Early settlers		−0.0080*	−0.0075	
		(0.0047)	(0.0046)	
Dep. var. mean	0.42	0.42	0.42	0.42
Individuals	1781548	1781548	1781548	324040
Demographic bins	636	636	636	556
Ancestral homelands	95	95	95	85
Adjusted R^2	0.22	0.18	0.22	0.21

*Notes:*This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants is not driven by ancestral homelands that accounted for a significant share of the US population in 1850 (e.g., Ireland, Germany, and the United Kingdom). Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.14 Continent-Specific Factors

Table C.14: Robustness to Continent-Specific Factors

	Gini				
	(1)	(2)	(3)	(4)	(5)
Ancestral migratory distance from the cradle of humanity	−0.069*** (0.012)	−0.069*** (0.018)	−0.068*** (0.018)	−0.096*** (0.034)	−0.096*** (0.034)
Continent FE		✓	✓	✓	✓
Africa			✗		
Latin America				✗	
New World					✗
Dep. var. mean	0.42	0.42	0.42	0.42	0.42
Individuals	1781548	1781548	1780090	1600075	1569521
Demographic bins	636	636	602	477	461
Ancestral homelands	95	95	88	71	69
Adjusted R^2	0.22	0.23	0.25	0.21	0.21

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected when the analysis exploits variation within continents and excludes individuals whose ancestral homelands are in Africa, Latin America and the Caribbean, or the New World. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.15 Conley's Spatial Correlation

Table C.15: Robustness to Conley's Spatial Correlation

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.069*** (0.012)	−0.023*** (0.0061)	−0.053*** (0.011)	−0.056*** (0.012)
Dep. var. mean	0.42	0.066	0.20	0.30
Individuals	1781548	1770406	1780511	1781548
Demographic bins	636	354	559	636
Ancestral homelands	95	61	87	95

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US remains statistically significant when accounting for spatial autocorrelation across ancestral homelands using Conley's method. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. Conley standard errors (500 km cutoff) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

C.16 Accounting for State-Specific Factors

Table C.16: Robustness to State Fixed-Effects

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.10***	−0.0083	−0.070***	−0.089***
from the cradle of humanity	(0.019)	(0.0054)	(0.011)	(0.015)
Dep. var. mean	0.42	0.069	0.21	0.31
Individuals	1776332	1733360	1770588	1776332
Demographic bins	2098	906	1686	2098
Ancestral homelands	94	58	83	94
Adjusted R^2	0.069	0.17	0.097	0.063

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US remains statistically significant when the analysis includes US state fixed effects. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a bin at the level of ancestry and state. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.17 Accounting for Micro Area-Specific Factors

Table C.17: Robustness to Micro-Areas Fixed-Effects

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance	−0.061***	−0.016***	−0.052***	−0.062***
from the cradle of humanity	(0.022)	(0.0025)	(0.011)	(0.017)
Dep. var. mean	0.39	0.063	0.18	0.28
Individuals	1639233	1083817	1540980	1639233
Demographic bins	19863	4756	13040	19863
Ancestral homelands	62	25	48	62
Adjusted R^2	0.13	0.12	0.15	0.13

Notes: his table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US remains statistically significant when the analysis includes fixed effects for US micro-areas. Ancestry-adjusted migratory distance is measured in units of 20,000 km. The unit of observation is a bin at the level of ancestry and over 2000 micro-areas. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.18 Exclusion of any Decade

Table C.18: Repeated Cross-Section: Robustness to Exclusion of any Decade

	Gini				
Exclude:	1980	1990	2000	2010	2020
	(1)	(2)	(3)	(4)	(5)
Ancestral migratory distance	−0.069***	−0.067***	−0.059***	−0.064***	−0.064***
from the cradle of humanity	(0.012)	(0.011)	(0.012)	(0.014)	(0.015)
Dep. var. mean	0.41	0.41	0.40	0.40	0.40
Individuals	6612353	6351379	6626686	6449622	6884640
Demographic bins	2394	2359	2295	2216	2144
Ancestral homelands	106	108	108	108	97
Adjusted R^2	0.20	0.24	0.24	0.22	0.23

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US, as measured by the Gini index, is not driven by any particular decade and remains stable and statistically significant when any single decade is excluded from the analysis. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.19 Ancestral Inequality

Table C.19 explores the effect of ancestral inequality on the Gini index in each demographic bin in the US, accounting for the ancestral inequality, as proxied by: (i) the Gini index and (ii) ancestral ethnic inequality. The results indicate that the estimated effects of ancestral inequality are insignificantly different than zero, and even slightly negative (Column (2) and (4)). Ancestral inequality therefore does not appear to persist. The finding further suggests that the estimated impact of the migration from Africa on inequality does not capture the impact of the persistence of ancestral inequality (Column (3) and (6)).

Table C.19: Accounting for Ancestral Inequality

	Gini					
	(1)	(2)	(3)	(4)	(5)	(6)
Ancestral migratory distance from the cradle of humanity	−0.075***		−0.084***	−0.068***		−0.069***
	(0.013)		(0.019)	(0.013)		(0.014)
Ancestral Gini (1980-1999)		−0.079**	0.025			
		(0.032)	(0.037)			
Ancestral ethnic inequality					−0.012	0.0041
					(0.015)	(0.015)
Dep. var. mean	0.42	0.42	0.42	0.42	0.42	0.42
Individuals	1607465	1607465	1607465	1755180	1755180	1755180
Demographic bins	424	424	424	607	607	607
Ancestral homelands	66	66	66	91	91	91
Adjusted R^2	0.26	0.22	0.26	0.22	0.18	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US, as measured by the Gini index, is unaffected by ancestral inequality—captured by ancestral Gini coefficients over the period 1980–1999 (Columns (2)–(3)) and by ancestral ethnic inequality (Columns (5)–(6)). Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.20 Ancestral Institutions

In view of the impact of European colonial settlements on the human capital, technology, institutions, and racism in the colonies, the societal and income inequality that these settlements have generated may have persisted to the descendants of these colonies in the US. Thus, we account for the potentially confounding effect of the colonial legacy in ancestral homelands. Moreover, we account for the degree of jurisdictional hierarchy among the ethnic groups that compose national homelands, and the historical level of democracy in these nations as captured by the Polity V index in 1900. The estimates suggest that these institutional dimensions do not have an impact on the baseline findings.

Table C.20: Accounting for Ancestral Institutions

	Gini								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ancestral migratory distance from the cradle of humanity	−0.065*** (0.013)		−0.062*** (0.016)	−0.067*** (0.013)		−0.067*** (0.012)	−0.063*** (0.012)		−0.062*** (0.012)
Ancestral colonial legacy		−0.010 (0.0064)	−0.0027 (0.0069)						
Ancestral jurisdictional hierarchy					−0.0014 (0.0060)	−0.0022 (0.0057)			
Ancestral polity score (1900)								−0.0051 (0.0035)	−0.0048* (0.0028)
Dep. var. mean	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Individuals	1748765	1748765	1748765	1778896	1778896	1778896	1364930	1364930	1364930
Demographic bins	595	595	595	621	621	621	337	337	337
Ancestral homelands	89	89	89	93	93	93	47	47	47
Adjusted R^2	0.22	0.19	0.22	0.22	0.18	0.22	0.29	0.24	0.29

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US, as measured by the Gini index, is qualitatively unaffected by ancestral institutions, as captured by: (i) ancestral colonial legacy, (ii) ancestral jurisdictional hierarchy, and (iii) ancestral polity score in 1900. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.21 Predicted Genetic Diversity

Table C.21: Predicted Genetic Diversity

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral predicted genetic diversity	0.43*** (0.086)	0.15*** (0.035)	0.34*** (0.078)	0.38*** (0.087)
Dep. var. mean	0.42	0.066	0.19	0.30
Individuals	1780310	1769879	1779320	1780310
Demographic bins	615	351	542	615
Ancestral homelands	92	59	84	92
Adjusted R^2	0.22	0.22	0.26	0.22

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US is unaffected when ancestral diversity is proxied by the ancestry-adjusted predicted genetic diversity of each ancestral homeland. All specifications account for sex and age group fixed effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.22 Rural-Urban Divide

Table C.22: Rural-Urban Divide

	A. Urban			
	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.075*** (0.013)	−0.023*** (0.0048)	−0.055*** (0.012)	−0.066*** (0.014)
Dep. var. mean	0.42	0.065	0.20	0.30
Individuals	1283490	1272583	1282495	1283490
Demographic bins	617	343	545	617
Ancestral homelands	93	58	86	93
Adjusted R^2	0.22	0.18	0.25	0.22
	B. Rural			
	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.054 (0.044)	−0.0052 (0.0099)	−0.031 (0.024)	−0.053 (0.032)
Dep. var. mean	0.39	0.061	0.18	0.28
Individuals	344676	339717	344092	344676
Demographic bins	309	178	268	309
Ancestral homelands	49	28	39	49
Adjusted R^2	0.14	0.056	0.18	0.12

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US holds when accounting for the rural–urban divide. To classify urban residence, we use the "metro" variable from IPUMS USA, which indicates metropolitan status. Individuals are coded as living in an urban area if they reside in a metropolitan area. Those for whom metropolitan status is indeterminable are excluded from the analysis. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications account for sex and age-group. The unit of observation is a bin at the level of ancestry and an urban dummy. Panel (A) includes only rural bins. Panel (B) includes only urban bins. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.23 Including African Americans and Native Americans

Table C.23: Including African Americans and Native Americans

	Gini	Top 1%	Top 5%	Top 10%
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.057***	−0.013*	−0.043***	−0.047***
	(0.014)	(0.0070)	(0.012)	(0.012)
Dep. var. mean	0.42	0.066	0.20	0.30
Individuals	2268937	2257795	2267900	2268937
Demographic bins	652	370	575	652
Ancestral homelands	97	63	89	97
Adjusted R^2	0.21	0.17	0.23	0.20

Notes: This table presents evidence that the association between (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) various measures of income inequality among their descendants in the US remains qualitatively unaffected when African Americans and Native Americans are included in the analysis. The migratory distance for African Americans is defined as the average distance from the cradle of humanity to Angola, the Democratic Republic of the Congo, the Republic of the Congo, and Gabon. The migratory distance for Native Americans is defined as the distance from the cradle of humanity to the territory of the present-day United States. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.24 Native American Tribes and Alternative Specifications

Table C.24: Robustness to Native Americans Analysis

	Gini						
	(1) WLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
Ancestral migratory distance from East Africa	−0.20***	−0.16***	−0.16***	−0.16***	−0.16***	−0.18**	−0.35***
	(0.032)	(0.044)	(0.050)	(0.043)	(0.054)	(0.072)	(0.087)
Sample FE	✓	✓	✓	✓	✓	✓	✓
Sex FE			✓		✓		
Age FE				✓	✓		
State FE						✓	
Micro-Area FE							✓
Dep. var. mean	0.45	0.45	0.43	0.43	0.41	0.41	0.40
Individuals	82748	82748	82748	82713	82424	76179	43900
Demographic bins	107	107	214	423	808	683	607
Ancestral homelands	36	36	36	36	36	36	32
Adjusted R^2	0.24	0.12	0.097	0.067	0.044	0.047	0.17

Notes: This table presents evidence that the association between (i) prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of Native American tribes and (ii) income inequality among their descendants in the US is: (a) robust to an alternative estimation method which accounts for the differential prevalence of each tribe in the US (i.e. a weighted least squares regression where weights are based on the size of each tribe in our sample), (b) holds unconditionally, and in particular irrespective of the inclusion of sex and age fixed-effects, (c) holds conditional on sex and irrespective of the inclusion of age fixed-effects, (d) holds conditional on age and irrespective of the inclusion of sex fixed-effects, (e) remains significant if the analysis is conducted with US states' fixed-effects, and (f) remains significant if the analysis is conducted with US micro-area' fixed-effects. This analysis is based on a repeated cross-section. The unit of observation is a bin in a given survey period at the level of: (i-ii) tribe, (iii) tribe and sex, (iv) tribe and age-group, (v) baseline demographic bin, (vi) tribe and state, (vii) tribe and over 170 micro-areas bin. Migratory distance is measured in units of 20,000 km. Heteroskedasticity-robust standard errors (clustered at the tribe-level) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.25 Dispersion in Education Accounting for Lagged Inequality

Table C.25: Dispersion in Education Accounting for Lagged Inequality

	Dispersion in Education			
	25-64 yrs old		25-34 yrs old	
	(1)	(2)	(3)	(4)
Ancestral migratory distance from the cradle of humanity	−0.10***	−0.10***	−0.091***	−0.096***
	(0.030)	(0.032)	(0.021)	(0.021)
Gini (1980)		−0.00085		−0.024
		(0.059)		(0.054)
Gini (1990)		0.075		−0.033
		(0.053)		(0.075)
Gini (2000)		−0.044		−0.037
		(0.050)		(0.059)
Dep. var. mean	0.70	0.70	0.69	0.69
Individuals	1768694	1768694	362065	362065
Demographic bins	431	431	124	124
Ancestral homelands	66	66	66	66
Adjusted R^2	0.18	0.19	0.12	0.11

Notes: This table demonstrates that the role of human capital dispersion as a mediating channel remains robust after accounting for lagged inequality in previous decades. Ancestry-adjusted migratory distance is measured in units of 20,000 km. All specifications accounts for sex and age-group fixed-effects. The unit of observation is a demographic bin. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level.* Significant at the 10 percent level.

C.26 Group Size

Table C.26: Ancestral Diversity, Income Inequality, and Group Size

	Ancestral migratory distance		Gini	
	(1)	(2)	(3)	(4)
Group size	0.0013		−0.00034	
	(0.010)		(0.0019)	
ln(Group size)		0.014		0.00053
		(0.016)		(0.0060)
Dep. var. mean	7252.7	7252.7	0.44	0.44
Individuals	1782313	1782313	1782313	1782313
Demographic bins	108	108	108	108
Ancestral homelands	108	108	108	108
Adjusted R^2	−0.0094	−0.0041	−0.0094	−0.0093

Notes: This table establishes that (i) ancestry-adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population and (ii) income inequality among their descendants in the US are both uncorrelated with the size of each ancestral group in the US—proxied by the group size in the sample. Migratory distance is measured in km. The unit of observation is a bin at the level of ancestry. The independent variables—group size and ln(group size)—are standardized. Heteroskedasticity-robust standard errors (clustered at the ancestral origins of the bin) are reported in parentheses.

D Robustness Checks - Figures

D.1 Minimum Size of Demographic Bins

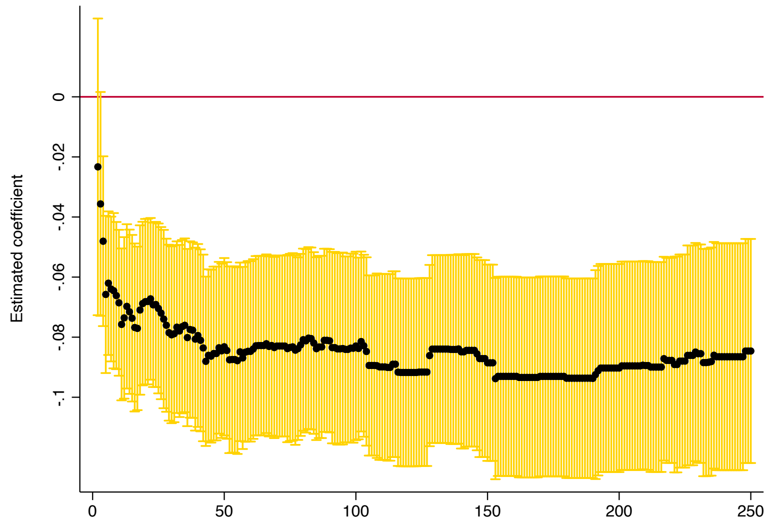


Figure D.1: Robustness to Minimum Size of Demographic Bins

Notes: This figure depicts the changes in estimated coefficient in our baseline specification, as we restrict the sample to demographic bins to include a minimum bin size and varying level from 2 to 250.

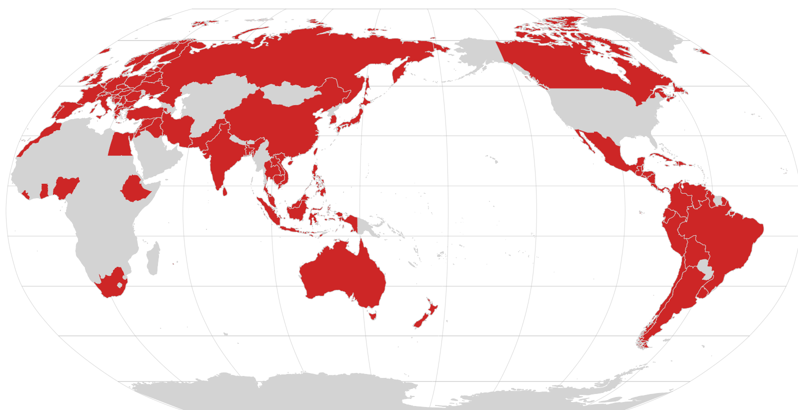
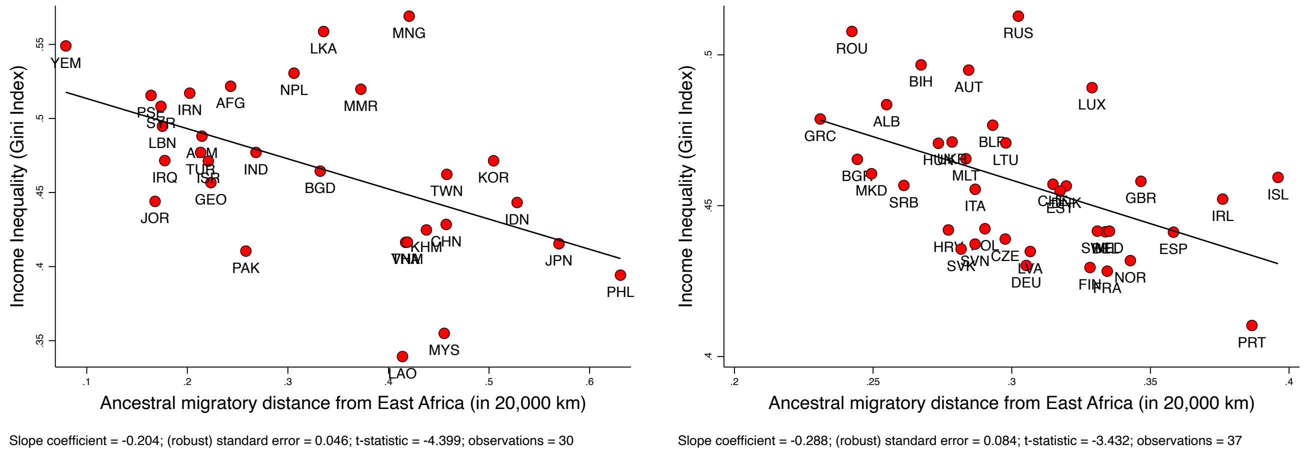


Figure D.2: Ancestral Homelands Included in the Analysis

Notes: This figure highlights in shaded red all the countries included in our analysis (Table I, Column(1)).



(a) Asia (b) Europe

Figure D.3: Income Inequality & Ancestral Migratory Distance from East Africa for Asia and Europe

Notes: This figure depicts the association between ancestral population diversity (predicted by the ancestry adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population) and income inequality among the descendants of these populations in the US, irrespective of the inclusion of sex, and age fixed-effects.

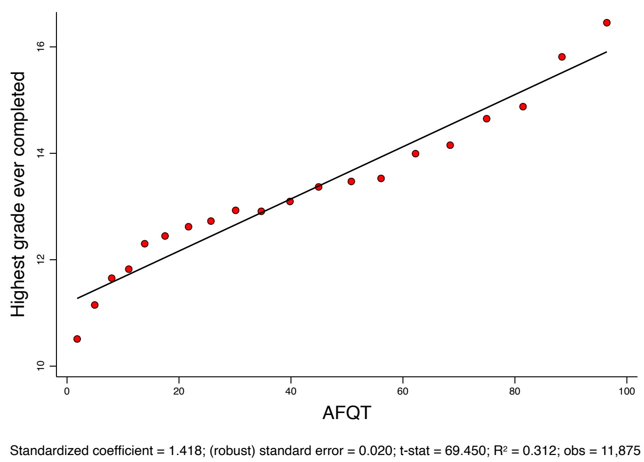


Figure D.4: Educational Attainment and Cognitive Ability

Notes: This figure depicts the binned scatterplot of the association between highest grade ever completed and AFQT score for a nationally representative sample of individuals from the National Longitudinal Survey of the Youth 1979.

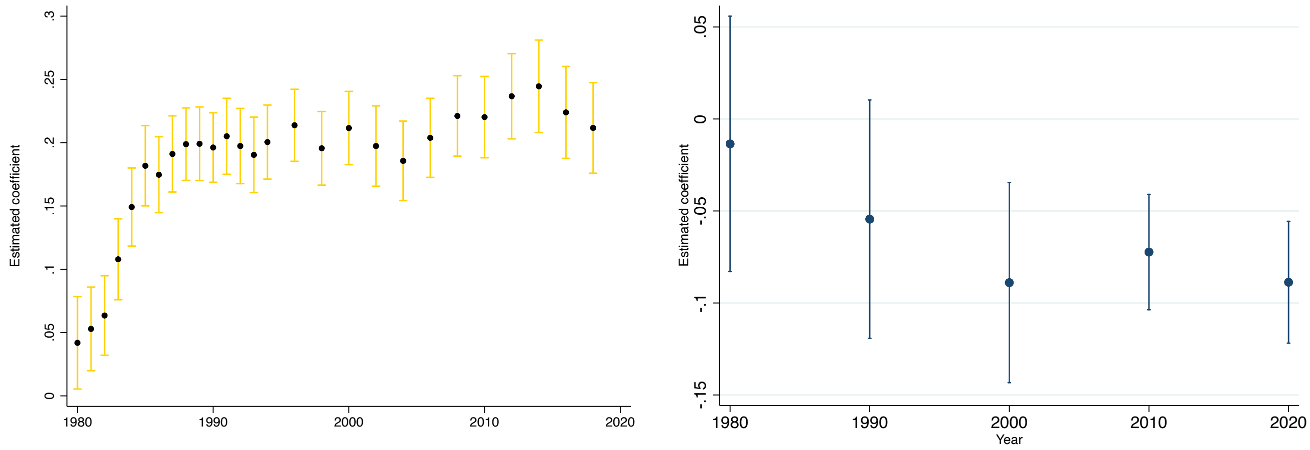


Figure D.5: The Returns to Ability and The Impact of Migratory Distance from East Africa on Income Inequality Among US Inhabitants Over Time

Notes: The left panel shows the evolution of the returns to ability (AFQT) in the US over time from 1979 to 2020. The right panel depicts the association between ancestral diversity (predicted by the ancestry adjusted prehistoric migratory distance from the cradle of humanity in Africa to the ancestral homelands of the US population) and inequality, as measured by the Gini index, across groups of individuals in the Unites States originated from the same ancestral background by decade using the common sample of ancestries across all decades. Ancestry-adjusted migratory distance is measured in units of 20,000 km.